Student Satisfaction Surveys and Nonresponse: Ignorable Survey, Ignorable Nonresponse

by

Luc Boyer

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor in Philosophy
in
Sociology

Waterloo, Ontario, Canada, 2009

Author's Declaration

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

In light of dwindling survey response rates, and increased costs in securing cooperation, nonignorable nonresponse, where the variables of interest are related to the decision whether to respond to a survey request, is a mounting concern. Topic saliency is strongly suspected as one of the main factors of nonignorability. With an increasing reliance on satisfaction exit surveys to measure how university alumni qualify their experiences during their degree program, it is uncertain whether satisfaction is sufficiently salient, for some alumni, to generate distinguishable satisfaction scores between respondents and nonrespondents.

This thesis explores whether, to what extent, and why nonresponse to student satisfaction surveys makes any difference to our understanding of student university experiences. A modified version of Michalos' multiple discrepancies theory was utilized as the conceptual framework to ascertain which aspects of the student experience are likely to be nonignorable, and which are likely to be ignorable. In recognition of the hierarchical structure of educational organizations, the thesis explores the impact of alumnus and departmental characteristics on nonresponse error. The impact of survey protocols on nonresponse error is also explored.

Nonignorable nonresponse was investigated using a multi-method approach. Quantitative analyses were based on a combined dataset gathered by the Graduate Student Exit Survey, conducted at each convocation over a period of three years. These data were compared against basic enrolment variables, departmental characteristics, and the public version of Statistic Canada's National Graduate Survey. Analyses were conducted to ascertain whether nonresponse is nonignorable at the descriptive and analytical levels (form resistant hypothesis). Qualitative analyses were based on nine cognitive interviews from both recent and soon-to-be alumni.

Results were severely weakened by external and internal validity issues, and are therefore indicative but not conclusive. The findings suggest that nonrespondents are different from respondents, satisfaction intensity is weakly related to response rate, and that the ensuing nonresponse error in the marginals can be classified, albeit not fully, as missing at random. The form resistant hypothesis remains unaffected for variations in response rates. Cognitive interviews confirmed the presence of measurement errors that further weakens the case for nonignorability. An inadvertent methodological alignment of response pool homogeneity, a misspecified conceptual model, measurement error (dilution), and a non-salient, bureaucratically-inspired, survey topic are proposed as the likely reasons for the findings of ignorability. Methodological and organizational implications of the results are also discussed.

Acknowledgements

If it is true that it takes a village to raise a child, it is also true that it took an entire department to raise this sociologist. There are, however, specific individuals who must be recognized for their distinctive contributions in the successful completion of my doctoral program. I wish to acknowledge their efforts in making this dissertation possible.

I want to send a sincere word of thanks to Prof. Rick Helmes Hayes, and Prof Augie Fleras for contributing greatly to my professional socialization. I also wish to acknowledge departmental lighthouses, Ilona Kosa and Luanne McGinley, who have both consistently listened and attended to so many graduate students, at times preventing us from crashing on the rocky shores of the university bureaucracy. I wish to extend a special note of gratitude to Dr. Jacob Sivak and Penny Pudifin, whose unwavering belief in graduate students gave meaning to my work at the Graduate Studies Office. Finally, I wish to extend my gratitude to both members of my thesis committee, Prof Keith Warriner and Prof Janice Aurini, for their support, advice and recommendations.

I would also like to thank my supervisor, Prof John Goyder, for taking on what must have seemed an interminable project. Thank you for providing me numerous opportunities to hone my professional skills, as a research assistant to the GSES, and through an AAPOR conference presentation, publications, and undergraduate teaching. You have my sincere gratitude for respecting my recursive learning process. Thank you for respecting my desire to learn on my own and to learn how to deal with the scrapes and bruises that come with doing research on nonrespondents. Your unwavering support has pushed me to learn beyond what I thought possible. Be assured that none of it will be in vain.

Of course, being a doctoral student does not shield one from the realities of life. At times, these realities have reminded me of the artificiality of graduate student life, and the need to attend to the ones most important in our lives. At other times, the relentless necessity to disassociate myself from others in order to cope with the sustained concentration such work requires meant cancelled vacations, outings, and family gatherings. My gratitude cannot begin to acknowledge the unconditional support and selflessness, exhibited all through these long years, of my mother, who passed away in 2004, and my partner, Reid. Yes, I am done now.

Table of Contents

List of Tables

List of Figures

Chapter 1
Satisfaction Surveys and Nonignorable Nonresponse

> When polls were first launched, we had told the world that the new system of sampling was superior to that followed by the *Literary Digest*. In fact, we were emboldened to say in print as early as July of 1936 that the *Literary Digest* would be wrong in November, and we were foolhardy enough to predict just what the *Digest* would find from its post-card poll of many millions. Our predictions about the *Literary Digest* poll results came true. The *Digest* was not only wrong, but its error was almost exactly what we said it would be. The sampling polls were on the right side, and by this very fact could lay claim to superiority over the straw vote methods which had prevailed up to that time. (Gallup 1957: 23-24)

1.1 Introduction

A lot of ink has been shed over the years about the quantity and quality of polls in public life. George Gallup, a strong proponent of polls applied to all spheres of public life, was certainly no stranger to this ongoing debate. It was Gallup, in 1936, who exposed the dangers associated with straw polls, and demonstrated the validity of scientific sampling methods applied to election forecasting. With a quota sample of a mere three thousands, as opposed to the Literary Digest's ten million ballots, Gallup was able to accurately predict the Roosevelt landslide in 1936, whereas the Literary Digest predicted the Republican Landon. Further analysis of the Literary Digest results pointed to a combined cause stemming as much from a biased initial sample that predominantly recruited from car and telephone owners, as from nonrespondents who favoured Democrats (Babbie, 1998; Squire, 1988). Emboldened by the success of the new sampling technique, the issue of the day for survey methodologists was the relative merit of quota versus probability sampling. In the aftermath of the election of 1948, for which Gallup's quota biased samples incorrectly predicted Dewey's victory, researchers soldiered on to develop new and better sampling methods

(Gallup, 1951; Gallup 1953, 1965); probability sampling, already utilized in government surveys in the mid-thirties, became the norm for survey practice (Frankel & Frankel, 1987).

Over the last decade, however, the challenge surrounding polls and surveys has no longer been of selection bias introduced by an inadequate sampling design; random sampling methodology when executed correctly offers adequate protection against fieldworker bias. The emerging issue, for a majority of survey researchers, is self-selection bias: the grim reality today is the difficulty in meeting sampling targets of four hundred to one thousand individuals – even when cash payments are offered. Sampled individuals are increasingly more difficult to contact, and less and less inclined to cooperate with a survey request. With the increasing reliance on surveys and polls as a means to accountability and representation in public institutions[1], survey practionners are placed, as Tourangeau (2004) points out, in a historically familiar but untenable position: as with the Literary Digest poll, there exists the real possibility that survey estimates will not only be inaccurate, but "spectacularly" so[2]. The purpose of this thesis is to explore whether, to what extent, and why nonresponse to student satisfaction surveys, used extensively within educational institutions, makes any difference to our understanding of student university experiences.

1.2 Ignorable and Nonignorable Nonresponse

It is not that survey researchers were unaware of the problem posed by nonresponse for survey statistics. Well before the dominance of probability sampling in survey practice,

[1] See for example Sudman and Bradburn (1987) on the growth of public opinion research.
[2] The error may come to light as a conflation between precision and accuracy of the survey estimate.

the basic calculation for nonresponse bias[3], as well as the basic methodological outline to assess the size of nonresponse bias, was already laid out (Cochran, 1953; Deming, 1944). The issue is an over-reliance on, and unwillingness to question, the basic premise that high response rates[4] are necessarily desirable, and indicative of better quality estimates (Clark & Boser, 1995). As Groves points out: "survey research has no more useful measures of nonresponse errors now than it did at its beginnings. Response rates have tended to be treated as proxy measure of nonresponse bias" (Groves, 1987: S161). With the growing decline in response rates, whether for household (Atrostic, Bates, Burt, & Silberstein, 2001; De Heer, 1999), telephone (Curtin, Presser, & Singer, 2005; Steeh, 1981; Steeh, Firgis, Cannon, & DeWitt, 2001), academic (Baruch, 1999), or topically-focused mail surveys (Connelly, Brown, & Decker, 2003), there is a worrying trend that could see the potential for large increases in nonresponse biases[5]. Recent empirical studies (Curtin, Presser, & Singer, 2000; Groves, 2006b; Keeter, Miller, Kohut, Groves, & Presser, 2000) of nonresponse bias have assuaged those fears somewhat by debunking the original premise; low response rates are no more indicative of nonresponse error than higher ones. Nevertheless, there remains the stark reality that one can never be sure when nonresponse is source of nonresponse error, and when it is not.

[3] Bias of the respondent mean is classically defined as the nonresponse rate multiplied by the difference in respondent and nonrespondent means (Groves & Couper, 1998: 3). More recent definitions of bias based on probabilistic theory can be found in (Groves, 2006b; Lessler & Kalsbeek, 1992).

[4] Definitions of response rates abound. They will vary depending on how completed returned surveys, partially filled-out returned surveys, ineligible cases, refusals, non-contacted and other forms of non-cooperation are entered in the calculation. Following Goyder (1987: 9-10), response rates can be derived from the returned surveys count divided by the difference between the original sample count and the number of confirmed ineligibles. More detailed formulas can be found in AAPOR (2006: 32-34)'s *Standard Definitions Manual*

[5] It should be noted that reports of decline in response rates are not always unanimous. There are studies such as Cummings, Savitz and Konrad (2001) which reports constant response rates over a ten year period ranging from 1985 to 1995. Connelly, Brown and Decker (2003) reports the same effect for a study period ranging from 1977 to 1987. Only when the period under study was extended were any significant trends were found.

The contribution of Rubin (1987) to the problem of nonresponse was to fix ideas on two types of nonresponse: ignorable and nonignorable. Nonresponse will be ignorable, that is to say it will not affect inferences made from survey results, if the differences between nonrespondents and respondents are essentially random with respect to the survey variables of interest. Despite the fact that a given population group may be underrepresented in the respondent sample, the latter will be biased but nonresponse can still be deemed ignorable. If, differences between respondents and nonrespondents are essentially random within the underrepresented group, the biased respondent sample can be corrected (reweighted) to reflect population values. However, nonresponse is nonignorable if "respondent and nonrespondent with exactly the same values of variables observed for both have systematically different values of variables missing for the nonrespondent" (Rubin, 1987: 202). Respondents and nonrespondents are not only systematically and significantly different from each other; the probability of cooperating with a survey request is contingent upon the values of the survey variables of interest[6]. The value of Rubin's conceptual distinction is to forgo the search for predictors of response rates, to forgo worries about bias caused by under or overrepresentation; the issue is whether the causes of nonresponse are related to the survey variables of interest – related to the phenomena under study.

It is with this clarification in mind that a growing body of research has centered on topic saliency as the central cause for self-selection bias. When sampled individuals are interested or otherwise involved with a given topic covered by a survey, research has shown that they are much more likely to cooperate with a survey request (Heberlein & Baumgartner,

[6] The result of this self-selection on the distribution of values will be a "truncated sample" (Winship & Mare, 1992)

1978). It is also more likely that these individuals, presumably more knowledgeable about the given topic, will hold different values on the survey variables of interest. As a consequence, respondents are likely to be those who find the topic salient and likely to hold values on the survey variables of interest different than nonrespondents; nonresponse is nonignorable because the cause of nonresponse is related to the values of the survey variables. Topic saliency is one of the mechanisms through which nonignorability operates.

What has plagued much of the research on topic saliency, however, is the inability to define precisely what one means by topic saliency, before the survey is administered and in terms of the sampled populations' own criteria. The difficulty in doing this does not lie in the presumption that topic saliency is essentially the private inner subjective world of a particular individual and, as a consequence, cannot be determined in advance; nor does it lie in the fact that there are as many topics as there are surveys. The issue is to provide topic saliency with a rationale, a content from which saliency is to be ascertained by the sampled individuals. If such a rationale has intersubjectivity, that is to say, forms the background from which a collectivity of individuals assesses the saliency of a given topic, only then may we be able to move away from the triteness of "topic interest" and "topic involvement", and begin to understand how topic saliency actually operates. It is our contention that the saliency of satisfaction surveys, if this is at all possible, is likely to be ascertained from the individuals' relationship with the organization from which the services were provided.

1.3 Saliency of Organizational Relationships

Why should we expect nonignorable differences between respondents and nonrespondents to satisfaction surveys? The answer lies, we would hypothesize, in the nature of our modern relationship with public institutions and organizations. Whether it is due to a lack of time, knowledge or both, we regularly relinquish control to public administrators, professionals and civil servants. Modern life prevents us from participating in all organizational decisions that may affect our lives. Day-to-day decisions about goods and services provided by organizations are routinely taken for granted. We, in effect, relinquish control over consumption (Etzioni, 1958). Satisfaction surveys disturb this relationship: "we expect to be able to trust authorities to do their work appropriately and well. […] we want safe airplanes and food, not the chance to participate in meat inspections and airline safety" (Warren, 1996: 49). As a result, satisfaction surveys are simply not salient to our lives. Only when the products and services fail to meet safety standards, do we pay attention to the offending organizations and institutions. To paraphrase Warren (1996), topic saliency operates at the margins of trust. Being satisfied is a strong impetus for not responding to satisfaction surveys.

To complicate matters, however, satisfaction surveys are bound with a conception of descriptive representation[7]. Their aim, akin to public opinion polls, is to collect information on the thoughts and impressions of the quality of a given service in a private or a public institution. In effect, citizens, consumers or clients are given the opportunity to voice for

[7] We are drawing and extending from Brehm (1993) and Berinsky (2004)'s analysis of nonresponse to public opinion polls.

themselves how they understood their experience at the supermarket, the hospital, or the welfare office. They are given an authority of voice: "the authority of voice does not mean that accounts of experiences and interests are accepted without question but, rather, presumes individuals are the best representatives of themselves" (Warren, 1996: 50)[8]. Through the active usage of satisfaction surveys, this authority of voice serves two important latent functions. First, organizations, whether private or public, are given the opportunity to open up new channels of communication, and to highlight to the consumers, citizens or clients what is deemed important to management (Kraut, 1996). Second, it assuages the bureaucratic divide by recreating severed links between consumption and control[9]. Citizens, clients and consumers, as recipients of the services provided by a given organization, are given the opportunity to inform management as to the quality and type of services they might expect in the future. Such surveys are often a key element of a wider client-centered management style[10].

Taken together, this means that many of us will concentrate our energies, and allocate time to organizational relationships that do matter. We become part of an "attentive public" that observes, reads, and essentially stays abreast of the organizations' policies and inner workings (J. Miller, 1983: 22-32; Roseneau, 1974: 98-103). This is particularly true of a

[8] The growth of opinion surveys may be linked to a desire for greater authenticity: to speak about one's student experience directly to the University umediated and uninterpreted by department officials. One should not however, conflate this desire for authenticity with truth. Without necessarily falling into delibitating solipsism, it should be said that opinions and attitudes are, to a large extent, a product of one's social context, and as a result largely self-interested. The gathering of opinions about students' experiences, whether through departmental officials or from students themselves is perspectival.

[9] See for example Gruber (1987) and Etzioni (1958) on the problem of control, and Katz and Danet (1973) for seminal works on the relationship between officials and clients.

[10] See Dinsdale and Marson (1999) for a discussion of strategies used to improve services in the public sector. See Foltz (1996) for a discussion of citizen surveys for public administration.

public that has transmuted its organizational relationship into a relationship of loyalty and citizenship behaviour. This public, quite unlike those who trust, is much more likely to be attentive, knowledgeable and motivated to respond to satisfaction surveys, and to report higher levels of satisfaction than most. The question is whether these individuals, who are poised to respond to organizational surveys, are nothing more than an interest group, or are representative of the general public that an organization serves. Stated more plainly, if the opinions and attitudes of the "attentive" respondents are no different than nonrespondents on the questions posed by the survey, nonresponse is ignorable, otherwise it is nonignorable. The danger, of course, is the latter: organizations are likely to poll the very individuals who are the most central to institutional norms.

Here, it should be clear to public administrators that survey nonresponse to student satisfaction surveys is not simply a problem for just survey researchers but to all parties involved. It is likely that students involved in all aspects of their degree program, central to the norms of the institution, are much more likely to answer student satisfaction surveys. If the satisfaction survey has recruited a wide range of experiential-bases from the alumni pool, to the extent that differences between respondents and nonrespondents vary only randomly, then nonresponse is ignorable. Second, if, as most of the survey literature would suggest, some individuals are more inclined to respond, nonresponse is ignorable provided that the underrepresented group does not base its decision whether to cooperate with a satisfaction survey on its satisfaction. However, as we have alluded already, there is a distinct possibility that students will base their decision to cooperate on the basis of their satisfaction with their degree program; nonresponse would be nonignorable. Empirical results on nonresponse bias

in satisfaction surveys are mixed: some show a small satisfaction bias, others a small dissatisfaction bias, still others no bias. There is still no conclusive evidence that would settle the matter.

1.4 Research Agenda

We need to understand whether, and under what conditions, and why nonresponse to satisfaction surveys can be deemed ignorable or nonignorable. The thesis focuses on two core questions[11].

Q1: Which aspects of the student experience are likely to be nonignorable, and which are most likely to be ignorable?

This question, drawing from the literature on topic saliency, seeks to identify which aspects of the university experience are salient enough to generate distinguishable respondent and nonrespondent attitudes, and which aspects yield similar answers from respondents and nonrespondents. Some survey variables, relating to salient aspects of the alumni's university experience, may be prone to nonresponse error, while other variables, less central to their experience, may not. This means we need to test the boundaries of ignorability by identifying salient experiences and evaluate for the presence of nonresponse error under a wide range of response rates. Since satisfaction surveys have both a descriptive and an analytical usage, it is necessary to assess the impact of nonresponse on two simultaneous fronts: first, an assessment must be made as to whether respondents' average and variance represent accurately the distribution of university experiences in the student population; second, an

[11] This research thesis aligns itself with studies of bias measurement from which these questions were derived. See for example Stoop (2005)

assessment must be conducted to establish if nonresponse alters the hypothesized relationship between variables affecting university experiences.

Q2: Why would satisfaction (or dissatisfaction) with their organizational experience cause alumni to systematically respond or decline to respond to a satisfaction survey?

Inquiries about ignorability or nonignorability must lead to a better understanding of survey cooperation behaviour. As Groves, Presser and Dipko (2004) rightly pointed out, theories of survey cooperation must be able to differentiate between response behaviours conducive to ignorability from those producing nonignorability. Because satisfaction can easily be reduced to a utilitarian explanation of survey cooperation[12], we need to acquire a greater vocabulary around the reasons for survey nonresponse. In particular, we need to trace the potential effects of an individual's organizational relationship on nonresponse error. As suggested earlier, we need to ascertain whether those who respond to satisfaction surveys because of their loyalty to the institution[13], share similar experiences to those who abstain from responding because they are essentially satisfied (trust).

In the chapters to follow, our inquiry will attempt to address these two core questions. Chapter two depicts the current state of survey research surrounding the ignorability issue generally, and as it applies to satisfaction surveys in particular. The review will provide a rationale as to why ignorability is perceived as a legitimate approach to the problem of

[12] The conceptualization of satisfaction-based behaviour can easily degenerate into a debate between maximizers and satisficers (Iyengar, Wells, & Schwarts, 2006; Schwarts et al., 2002). We agree with Michalos (1973) that greater vocabulary about rational action is required.

[13] Schiltz makes precisely this point on the effect of organizational experience on nonresponse error: "those who came through the baccalaureate experience with an affection for the institution and quite probably those who are satisfied with their jobs will be more likely to answer" (Schiltz, 1988: 70). The author makes a further point that if that would the case, alumni surveys should exhibit higher returns than the surveys on the general public.

nonresponse. This will be followed by conceptual and empirical studies that will challenge this approach. Topic saliency will be specifically reviewed as a potential source of nonignorability in satisfaction surveys. Conceptual hypotheses will be devised to challenge the ignorability approach, and to assess when, how and why nonresponse may be deemed nonignorable.

Chapter three outlines the methodological approaches employed to test these conceptual hypotheses. The methodology adopts a multi-method approach to the study of nonresponse. A quantitative approach will assess the degree of bias on satisfaction variables, and evaluate which aspects of the university experience are most salient, and therefore most conducive to nonignorable nonresponse. We will base our analysis on data collected from the University of Waterloo's Graduate Student Exit Survey, as well as Statistics Canada's National Graduate Survey. These analyses will be complemented by a qualitative approach which aims to uncover the cognitive dimensions of topic saliency and their effect on nonresponse error.

Chapters four through seven proceed with the analysis of nonresponse. Chapter four investigates the influence of student characteristics such as enrolment variables and program discrepancies on nonresponse; nonresponse bias estimates are calculated from both alumni population data and Statistics Canada's alumni survey. Chapter five analyzes whether variations in departmental response rates are related to variations in the model elaborated in the previous chapter; the potential for hierarchically-clustered biases will be systematically tested against departmental records. Chapter six tests for the presence of bias when survey implementation, in terms of final reminder formats and mailing dates, is systematically

varied. Chapter seven gathers alumni's think-aloud reflections as they answer the graduate student survey; these data are analyzed with reference to topic saliency and measurement error. Finally, chapter eight synthesizes the findings from all methodological approaches, evaluates the conceptual hypotheses elaborated in chapter two, and offers recommendations for the implementation of future student satisfaction surveys.

In presenting the results of our inquiry, we must caution the reader that our findings will not yield definitive answers to the extent of resolving once and for all the issues surrounding nonignorability. The methodological difficulties encountered during this research mitigate against our ability to present strong conclusions. The methodological stratagems used to circumvent the obvious fact that nonresponse studies imply studying individuals who by definition are not responding, are generally weak in validity and demand extra-survey data that are generally difficult to obtain. The extant data for this research was limited by the inability to secure access to the restricted portion of Statistics Canada databases, the inability to extract large amounts of variables from alumni student records, the inability to crosslink anonymous unmarked surveys with student records, and ironically the inability to generate sufficient interest from alumni to respond to fieldwork interviews. Coupled with possible satisfaction ceiling effects associated with a university voted the best Canadian university by Maclean's, the results of this study remain tentative. Despite these methodological shortcomings, this case study should be understood as a necessary first step in clarifying the boundaries of ignorability at the individual, departmental, and cognitive levels.

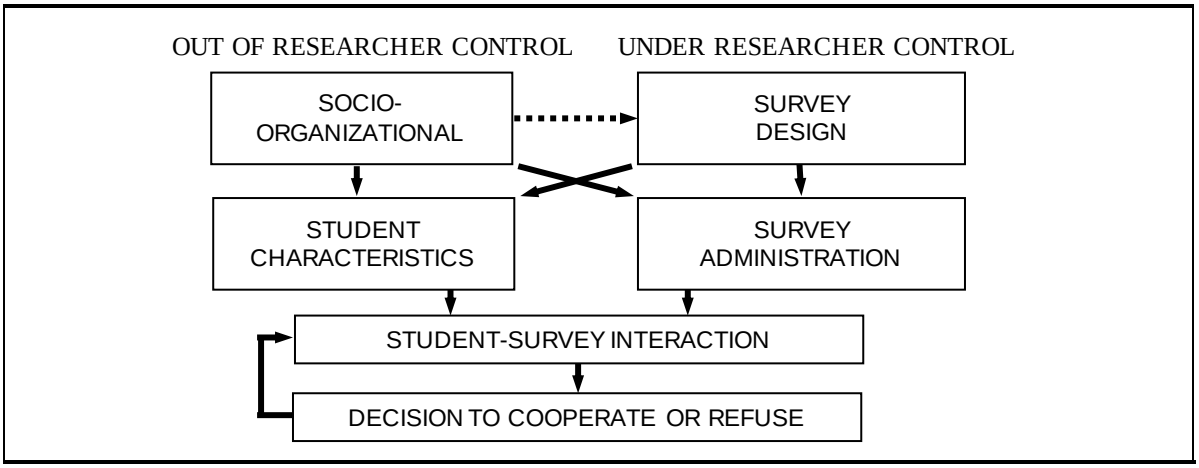
Chapter 2
Literature Review: Topic Saliency and Ignorability

2.1 Introduction

This chapter assembles known components of nonresponse error attributable to topic saliency and considers their impact on satisfaction bias. The literature will be analyzed to ascertain the state of current knowledge on the ignorability issue in relation to topic saliency. In the course of gathering the literature on nonignorable nonresponse to student satisfaction surveys, it became clear that the number of empirical studies specifically dealing with satisfaction was extremely limited. It is common to find studies providing correlates of response rates but rarely will they inquire about the effects of such variations on survey variables. To acquire a more comprehensive understanding of the effect of response rates on nonresponse error, we have proceeded with three concomitant strategies. First, the search was broadened to include satisfaction bias studies from other organizations that administer employee surveys, health surveys, and hotel surveys. Admittedly, the net publication count was still limited. We suspect that most studies of nonresponse bias remain unpublished internal reports. Second, student satisfaction studies and nonresponse studies were conjointly scanned in order to extract any potential "common-cause" variables (Groves, 2006b). These variables are not indications of bias; they only serve to delineate potential avenues for further empirical research. Third, the literature was organized around Groves and Couper (1998)'s

survey cooperation conceptual framework[14] (see Figure 2-1 below). For each conceptual block, the current state of our knowledge on nonresponse error (e.g. topic saliency) in satisfaction surveys will be assessed.

Figure 2-1 Survey Cooperation Conceptual Framework



Source: (Groves & Couper, 1998: 30)

The first section traces the linkages between nonresponse and ignorability. It will first show how nonresponse error can be conceived as ignorable or nonignorable; empirical findings on ignorability will follow. The second section will extend the discussion of

[14] Two important modifications were made to Groves and Couper (1998)'s model, both of which, we believe, are implicitly acknowledged by the authors. The first modification to the model is the establishment of a causal relationship between social environment and survey design. While it is true that survey researchers do have a range of methods to choose from, giving the appearance of freedom, they have consistently been aware of, and adjusted themselves to, particular social conditions. Survey design has a history (Frankel & Frankel, 1987; Tourangeau, 2004). In addition, the organizational context in which survey researchers find themselves may, depending on the autonomy given to the researcher, have a serious curtailing effect on the freedom to choose one design over another. What appears to be "under the researcher's control" must be properly understood within its organizational context as well. This is true for the survey administration portion of the survey; it is equally true for the choice of survey design. A second modification to the model is the placement of social exchange theories in the "decision to cooperate or refuse" conceptual block. Social exchange, like its other theoretical counterpart, leverage-saliency, is a theory of survey cooperation decision-making. It is, we believe, a complement to the leverage-saliency theory of survey cooperation (Goyder, Boyer, & Martinelli, 2006).

ignorability to topic saliency explaining how satisfaction relates to saliency. The third section will conduct a review of findings that will attempt to trace the presence, direction and magnitude of satisfaction bias. In addition, a search of socio-demographic factors, organizational context, survey design determinants and cognitive biases related to survey interaction will be conducted. The fourth section will review theoretical models that will elucidate how one's organizational relationship may influence the decision to participate in an organizational survey. The chapter concludes the literature review by assembling theoretical propositions, empirical findings, and unexplored areas of research. Together, they will form the basic hypotheses in response to the research questions elaborated in our introduction.

2.1.1 Nonresponse and Ignorability

The case for ignorability receives credibility from a series of studies that demonstrate that nonresponse, particularly at higher levels, does not generate the expected levels of bias in survey variables. Whereby it was generally assumed that low response rates are indicative of a biased dataset, current readings on nonresponse challenges this taken-for-granted assumption, and renders a more complex look at nonresponse error. This section begins by tracing the linkages between nonresponse and nonresponse error; it will be followed by an explanation as to how and to what extent nonresponse could be construed as ignorable.

2.1.2 Nonresponse and Nonresponse Error

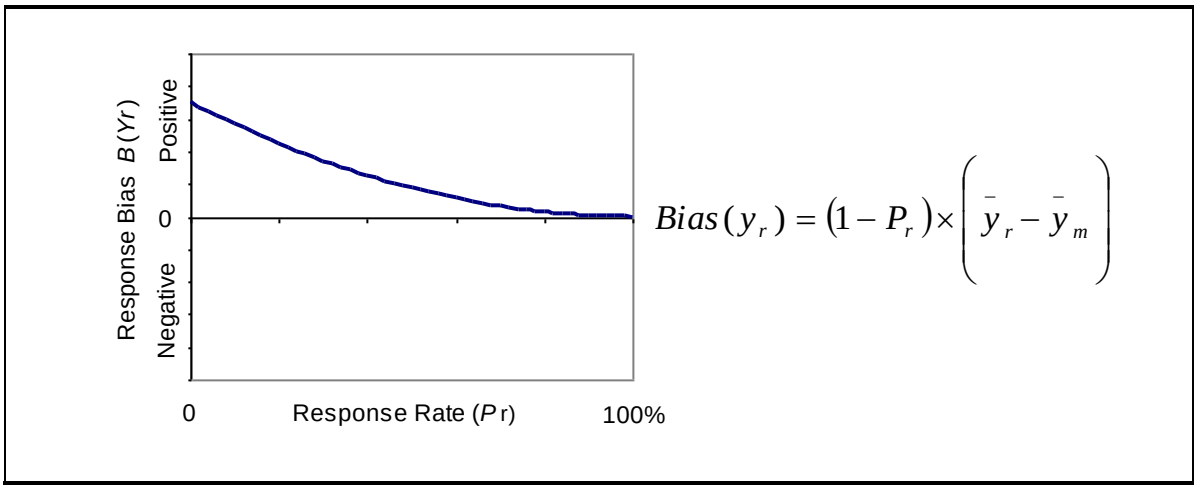
It is no surprise to survey researchers that not all individuals sampled from the population of interest will respond to a given survey request. Following (Groves & Couper,

1998), nonresponse can be classified as one of three types: nonresponse can occur due to non-contact, to refusals, or through other health or particular conditions of the sampled individual. Sources of non-contact may come in the form of incorrect mailing addresses, non-working phone numbers, never-at-home individuals, or otherwise unreachable households (AAPOR, 2006). Nonresponse may also occur for a host of other reasons pertaining to the sampled individuals: he or she may have physical or mental health issues, literacy, language or translation issues, or may even be deceased. Finally, nonresponse may be caused by individuals who, despite being contacted, refuse to cooperate and flatly resist any attempts to engage them in responding to a given survey. The magnitude of nonresponse will vary depending on the respective size of non-contacts, refusals, and other conditions of sampled individuals.

Nonresponse may be indicative of nonresponse error. A truncated dataset (e.g. less than 100% response rate) may be said to exhibit (non-sampling) nonresponse error if "the values of statistics computed based only respondents data differ from those based on the entire sample data" (Groves, Fowler et al., 2004: 59). When a portion of the sampled population groups are more susceptible to responding while others remain nonrespondents, a bias will be introduced if these nonresponding groups are also related to the main variables of interest present in the survey. Some population groups will be over or under represented in the respondent pool, and this will bias the survey variable marginals positively or negatively. In essence, the respondents' pool will hold different values on the survey variables of interest than that of the entire sample.

We can deduce the resultant bias in the respondent's mean from the classic formula[15] in Figure 2-2. Nonresponse bias (in the respondent mean) is the product of two factors: 1) the nonresponse rate (1-P_r) and 2) the mean differences between the respondent (y_r) and nonrespondents (y_m) on a given survey variable of interest. While this mean difference may occur at any response rates, higher rates will diminish the "risk of nonresponse bias" (Groves, Fowler et al., 2004: 59); said differently, the nonresponse rate (1-P_r) amplifies the mean differences between the respondent (y_r) and nonrespondents (y_m). It is precisely in light of this classic formula on respondent bias (Groves, Fowler et al., 2004), that the current decline in response rate can be seen as so alarming.

Figure 2-2 Nonresponse Positive Bias as a Function of Response Error



The recognition of the problem of nonresponse error has often led to a desire to increase the response rate, and this for two reasons. First, as the formula suggests, an increase in response rate will diminish the impact of any mean differences between respondents and nonrespondents. It is presupposed that, as the response rate increases, one is drawing from a

[15] Here we have chosen the classic formula of respondent bias

more diversified response pool of respondents, and the initial bias is eventually diluted. As can be seen in Figure 2-2, an increase in response rate reduces the (in this case positive) initial bias. Second, it is often assumed that low response rates are also indicative of more biased responses: those who respond more readily to a survey request may hold different values on the variables of interests than those who are said to be more resistant. It is this particular assumption that has been the more enduring among survey researchers, and the more difficult to assess. As a result, it seemed always preferable to increase response rates to their maximum values in the hope of circumventing any initial bias on variables that are absent from the sampling frame (ex: satisfaction scores, etc), and thus untestable for bias. Unfortunately, response strategies that may have seemed to work in the past do not seem to halt the free fall in response rates. Moreover, it is more costly than ever to reach an ever-growing body of recalcitrant individuals.

However, a series of empirical studies has provided some evidence that nonresponse rates may not be indicative of the quality of the survey. Variations in response rates, as a consequence of changes in the survey design, or through greater recruitment efforts, do not necessarily effect any change on the survey variables of interest. Thus, Keeter et al (2000) opinion poll found that a change of response rate from 36% to 61% produced variation less than nine percentage points on fourteen of ninety-one opinion questions. Differential response pattern were detected amongst demographic variables but no test was conducted against any of the opinion questions. Curtin et al (2000) comparative study between early respondents and those who required additional recruitment efforts, revealed small but significant differences on the index of consumer sentiment – albeit using larger than normal

samples. Using election results as a benchmark to assess the accuracy of exit polls, Merkle and Edelman (2002) also found no substantial bias despite influences on response rates such as interviewer and voter's age. In contrast, Parashos, Morgan and Messer (2005), Teitler, Reichman and Sprachman (2003) and Groves (Groves, 1989; 2006b) found significant nonresponse biases showing the uncertainty and unpredictability of bias in relation to nonresponse.

2.1.3 Nonresponse Error: Ignorable and Nonignorable

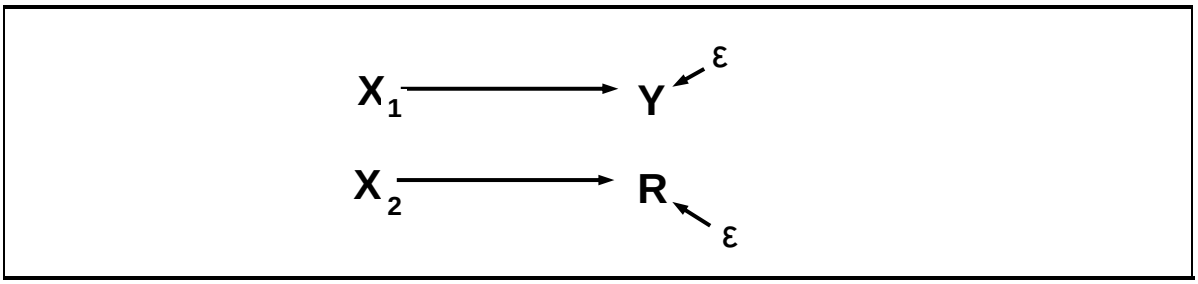
The coinage of the terms "ignorable" and "nonignorable" by Rubin (1987) focused ideas on two types of nonresponse error. Nonresponse, while presumably generative of nonresponse error can nonetheless be "ignored" to the extent that sample representativeness can be restored if corrective measures are applied to the truncated dataset. Nonignorable nonresponse is a particular form of nonresponse error whereby the values of the survey variables of interest (e.g. outcome or dependant variables) are determining, in some fashion, the decision to respond to a survey request. Following Groves (2006b)'s lead, ignorability and nonignorability can be represented and elaborated through a causal mechanism that models both the propensity to respond and the survey variable of interest. This section will discuss and specify which causal mechanism are implicated in the production of ignorable nonresponse and which are associated with nonignorability.

2.1.3.1 Ignorable Nonresponse: Missing Completely at Random

While it may appear counterintuitive, it is entirely conceivable that nonresponse may be completely ignorable. Missing surveys are ignorable if the causes of nonresponse are

unrelated to the values of the survey variables of interest, as shown in Figure 2-3. When there are no causes of nonresponse (X_2 is nonexistent), there is by definition no bias due to nonresponse. If, however, some population groups (X_2) are responding in greater numbers than others, some over or under representation in the respondent pool will occur, but again no bias will occur because these population groups are unrelated to the survey variables of interest (Y). This particular type of nonresponse is labelled missing completely at random (MCAR) to the extent that nonresponse will not influence the survey variables of interest (other than sampling error). The resultant dataset, despite being truncated (e.g. less than 100% response rate), can be assumed to be free of non-sampling nonresponse error[16].

Figure 2-3 Missing Completely at Random Causal Model



Adapted from (Groves, 2006b)

However, this situation is, in reality, highly suspect considering the literature on survey nonresponse which names specifically socio-economic status, education, gender, and age as known correlates of response behaviour (Groves, 1989: 201-206) and by all standards these same variables are sources of important sociological correlates as well. Indeed, it is

[16] The fact that responses are free of any systematic biases does not resolve the issue of sampling bias. Assuming that the decision to respond is truly random, the resultant response pool, although free of systematic bias, will be exhibit large sampling error. The lower the response rate, the lower the effective sampling size will be; thus, as a result, the higher the sampling error and the larger the confidence intervals around the variables of interest.

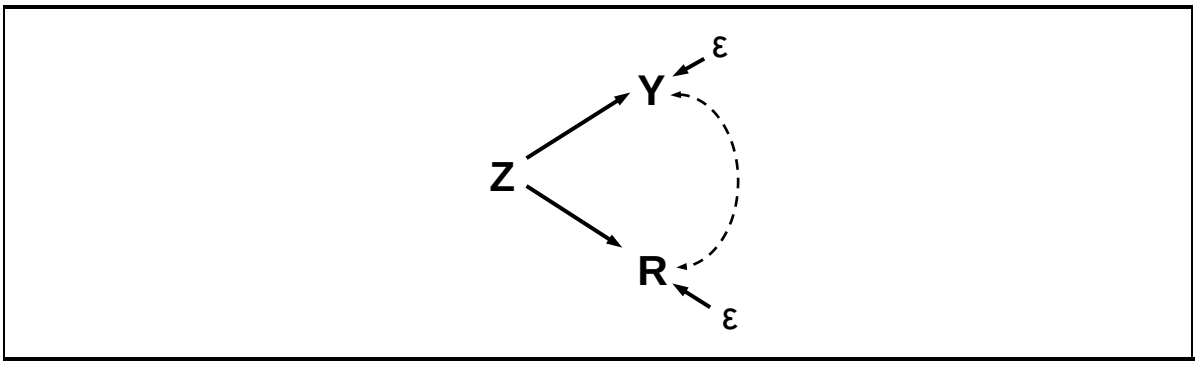
more likely that the decision to respond to a survey request and the variables of interest be linked in some fashion. This is what we are addressing next.

2.1.3.2 Ignorable Nonresponse: Missing at Random

If one acknowledges the possibility that variables causing nonresponse may also be cause to the survey variable of interest, nonresponse may still be considered ignorable but under strict conditions. This is because the presence of common-cause variables (Z) will show a spurious relationship between the survey variable (Y) and the probability of responding (R), (shown by the dotted line in Figure 2-4). If the covariation between the decision to cooperate (R) and the survey variables of interests (Y) can be fully elaborated by one, or a set of, common causes (Z), nonresponse can be ignored. The original (spurious) nonresponse error between the survey variables of interest (Y) and the response propensity (R) can be partialled out by the introduction of the common-cause variable (Z)[17]. This type is nonresponse is labelled missing at random (MAR). It reflects the fact that bias is caused by the overrepresentation of one category (Z) versus another. However, within a given category (Z), respondents and nonrespondents are essentially the same with regard to the values of the survey variable of interest.

[17] This presupposes that a common-cause variable (Z) is fully measured for all sampled cases including missing ones. If the variable (Z) is only partially measured, the nonignorable condition has not been surmounted. Nonignorability can take two additional forms other than the one mentioned so far (Baker, 2000: 48-49)

Figure 2-4 Missing at Random Causal Model



Adapted from (Groves, 2006b)

Here, unlike the missing completely at random (MCAR) condition, this second type of ignorable nonresponse does not signify complacence. Indeed, ignorability hinges on how the variables of interest will be analyzed. If descriptive statistics are required, such as the reporting of means and proportions, representational bias must be overcome by a reweighing of the respondent to match population ratios. Nonresponse is ignorable only to the extent that one has at one's disposal all the variables (Z) to elaborate fully the spurious relation between (Y) and (R). Should these variables be unavailable, representational bias cannot be corrected, and as a result, variables of interest (e.g. its marginals) will remain biased. In effect, missing at random condition is an issue of specification and not of response rates per say. Thus, in cases where descriptive statistics are required, ignorable nonresponse is highly suspect if common-cause variables are not commonly found in the sampling frame such as attitudinal ones.

Nevertheless, it is possible to uphold the ignorability status, and to circumvent to need for sampling frame variables -- even if a survey variable is suspected to be a common-cause variable. The key is to forgo descriptive statistics and to focus upon bivariate and

multivariate relationships with the variable of interest. As was mentioned earlier, by definition, missing at random stipulates that respondents and nonrespondents are essentially similar to one another within a given category of the common-cause variable. An increase in response rates will alter the representational bias, but will not alter the distribution of values on the variable of interest within a given category. In effect, changes in response rate will not alter the form of the relationship between the common-cause variable and the variable of interest. This is the basis of the form resistant correlation hypothesis (De Leeuw, Mellenbergh, & Hox, 1996; Goudy, 1976, 1978).

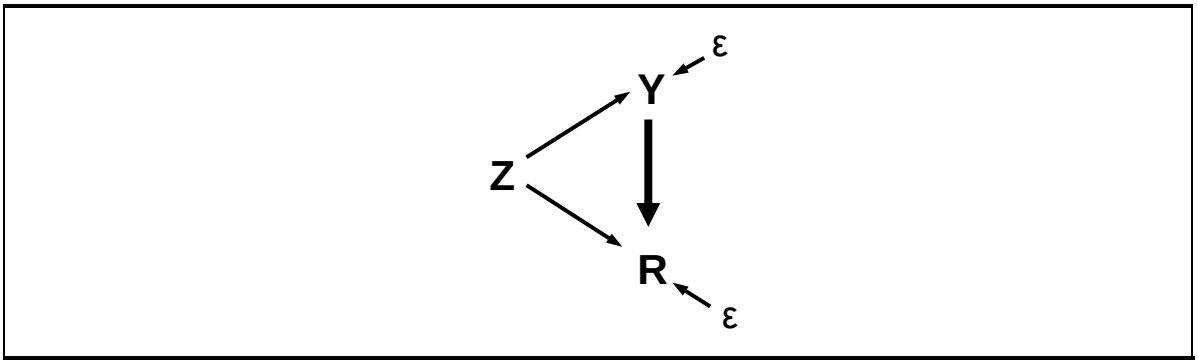
Of course, in the event that the decision to cooperate with a survey request is determine by the variable of interest itself. The form resistant hypothesis is nullified. Respondents and nonrespondents are no longer identical within a given category. This within-category bias may operate in one or all categories of the survey variable of interest. In cases, where surveys are highly salient to some, and not to others, it is likely that the form resistant hypothesis may prove to be false.

2.1.3.3 Nonignorable Nonresponse

When the decision to respond to a survey is based on the variable of interest, nonresponse is said to be nonignorable. Nonignorable nonresponse can be defined as "[…] response bias in the sense that a respondent and nonrespondent with exactly the same values of variables observed for both have systematically different values of variables missing for the nonrespondent" (Rubin, 1987: 202). Said differently, the values on the survey variables of interest will be systematically different for respondents and nonrespondents. Typically,

nonignorable nonresponse is represented by the causal arrow, shown in Figure 2-5, from the variables of interest (Y) towards the probability of being a respondent (R)[18]. It could be conceived, for example, that the decision to withhold participation may be based on the reluctance to reveal one's low income (Y). The resultant dataset would underestimate the range of incomes in the sampled population, showing a bias towards higher incomes.Under such a condition, sample representativeness is irretrievable without some knowledge of the mechanism that caused one to respond or not to respond to a survey request (Z).

Figure 2-5 Nonignorable Nonresponse Causal Model



Adapted from (Groves, 2006b)

One possible source of such mechanisms generative of nonignorable nonresponse can be conceptualized as stemming from the topic of the survey itself. Figure 2-6 elaborates the effects of survey topic on the relationship between types of nonresponse and the researcher's control over the probability that a sampled individual responds or not. The rows expand on Groves (1998)'s cooperation model illustrated in Figure 2-1. What are deemed out of researcher's control are personal and contextual characteristics of the sampled individuals that may have an impact on the decision to participate; under researcher's control are all

[18] See (Groves, 2006b: 650-652)

aspects of the survey design and administration of the survey that can be tailored to entice or negate the sample individuals' decision to participate. The columns expand on the three types of response hurdles one may encounter when sending a survey request; a survey request assumes contactibility of the sampled individuals, their cooperation, and the absence of other unknown barriers to participation such as health-related issues or illiteracy. Both "contact" and "other" are behavioural or physical conditions of the sampled individuals that will enhance or impede survey response. "Cooperation" is cognitively based. It involves a decision-making process which, depending on one's motivation, can range anywhere from the shallow heuristically-based to the more in-depth calculative modes.

Figure 2-6 Examples of Nonignorable Nonresponse Due to Survey Topic

	Types of Nonresponse		
	Contact	**Cooperation**	**Other**
Out of Researcher Control	Travel survey: frequent travels	Satisfaction survey: irate, bad experience	Health survey: illness, disease
Under Researcher Control	Travel survey: poor call-back design or survey period too short	Satisfaction survey: no incentives provided, tasks too demanding	Health survey: dexterity, arthritis and use of mail surveys

As can be seen from Figure 2-6, some survey topics are inherently prone to nonignorable nonresponse. Surveys interested in collecting information on travel behaviour , for example, are likely to introduce bias because segments of the sampled population who travel more frequently may not be easily reachable (Zimowski, Tourangeau, Ghadialy, & Pedlow, 1997); a situation which can be compounded by a poor call-back or follow-up survey design. Surveys interested in collecting information on special populations such as

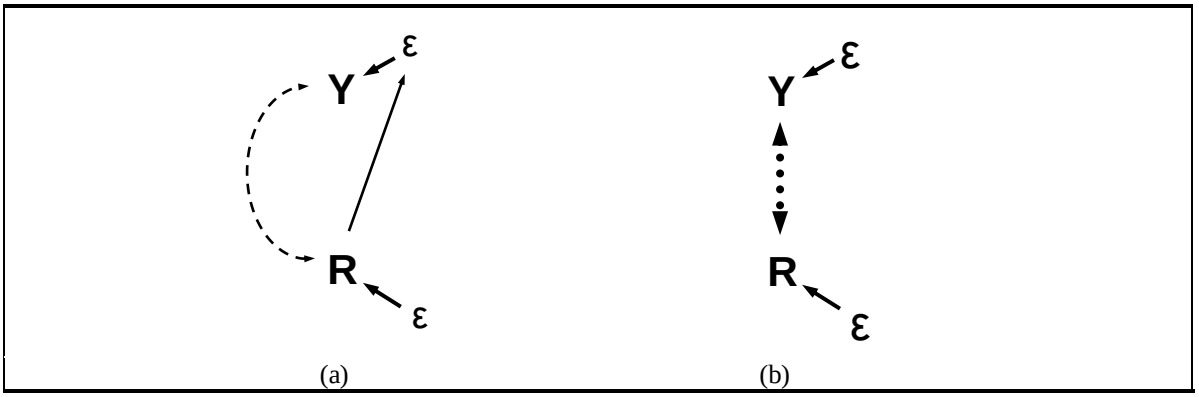
hospital patients or the elderly may also encounter difficulties in securing a response. Nonresponse studies on health-related surveys on the aged, for example, have shown underreporting of health conditions for the elderly 75 years of age and over (G. Cohen & Duffy, 2002). Both of these survey topics pertain to behavioural aspects of the sampled individual. Cooperation, because of its cognitive nature, will produce bias if the survey topic is judged as important or trivial by the sampled individual. Here interest in the survey topic, as opposed to strictly behavioural aspects, plays a role in the decision to participate in the survey. Satisfaction surveys, for example, may attract individuals who had a particularly good or particularly bad experience, misrepresenting, as a result, the sampled population's real satisfaction levels. A lack of incentives may not counteract the costs perceived in voicing one's satisfaction level, and as a result, attract those respondents who find the survey topic particularly interesting.

2.1.3.4 Nonignorability and Measurement Error

Nonresponse error is rendered increasingly more complex by the fact that non-sampling survey errors are interrelated with one another. Three different types of such errors may dissimulate or spuriously create a covariation between survey variables of interest (Y) and survey cooperation (R). The first of such errors, shown in Figure 2-7a, and illustrated by the causal arrow between survey cooperation (R) and measurement error on survey variables (ε on Y), creates a spurious covariation between survey cooperation (R) and survey variables of interest (Y). As the causal mechanism indicates, survey cooperation will affect the quality of responses on the variables (Y). It is conceivable, for example, that increased efforts at obtaining cooperation may yield answers on the survey variables (Y) which are less

reflective of the respondent's true opinion but more off-the cuff "non-attitudes" (Converse, 1970) such as "saying yes to anything" (Ray & Still, 1987: 572). Thus for nonresponse to remain ignorable, response biases such as acquiescence or social desirability must be eliminated.

Figure 2-7 Nonignorability and Measurement Error Model



Adapted from (Groves, 2006b)

The second and third types are illustrated by the measurement error (ε) on both the response (R) and survey variables (Y), here both shown in Figure 2-7b. In both cases, measurement error in and of itself will introduce variability in the range of possible values for each of these variables. Increased variability (noise) will drown the capacity to detect any real bias between the survey variables and survey cooperation. This may be the case when the survey variables (Y) are poorly measured and unreliable such as the case where respondents may not always understand the question, or may not be able to map their answers on the choices provided. The respondent may approximate an answer or choose to skip the question altogether. It may also be the case, as Stoop (2005) convincingly demonstrated, that survey cooperation may not always be recorded reliably by survey interviewers, or that one may not be able to distinguish between non-contact and refusals. In

either case, the net result is to conclude erroneously that nonresponse is ignorable when in fact it is potentially nonignorable. Thus, ignorability requires reliable measures in both the survey variables of interest and survey response.

2.1.3.5 Nonignorability and Population Coverage

Finally, there may be an insidious relationship between the homogeneity of the population or sampling frame and nonignorable nonresponse. Indeed, by definition, nonignorability presupposes differential values with respect to the variables of interest between respondents and nonrespondents. Should this differential be absent because of the homogeneity of the sample, nonignorability is rendered an impossibility. Sample homogeneity can be a fact of the population under study, or be the result of the design of the survey practionner. Indeed, the latter through the decision related to population coverage might render a sample more homogenous with respect to the values on the variables of interest. Particularly, population coverage error, that is to say excluding groups susceptible to having different values on variables (Y) and having a different response propensity (R), will render the sampling frame more homogenous and unwittingly remove sources of nonignorability.

This issue of who gets to be in the sampling frame and who does not strikes at the heart of the issues surrounding nonignorability. Nonignorability is a problem only to the extent that one is interested in representativeness. If the intent is gathering opinions and only dissatisfied ones for example, the best design strategy is precisely to leave a card on a hotel coffee table, or to install a message box. The design implicitly weeds out “unwanted” responses. If, however, it is the intention to get a sense of how satisfied customers or clients

are with one's organization then representativeness and the nonignorability issue must be dealt with. This is particularly problematic in the case of student satisfaction surveys, where withdrawal and failed to register students are routinely excluded from the sampling frame. These students may hold very different satisfaction scores regarding their university experience relative to university alumni, who ultimately received what they aimed for.

2.2 Topic Saliency and Satisfaction

In the previous section, it was suggested that topic saliency, as a source of nonresponse error, is a subset of a much larger collection of nonignorable nonresponse stemming from the survey topic. It was also suggested that nonresponse could be deemed ignorable by the retrieval of common-cause relationship between the survey variables of interest and the decision to respond to a survey request. The form resistant hypothesis would hold if the variable of interest were itself not cause for nonresponse. In most situations, this hypothesis might be sufficient to resist the knee-jerk reaction to increase response rates at all costs. However, in relation to ignorability, topic saliency generally, and satisfaction surveys in particular, may be a special case. There is some evidence in the literature to suggest that topic saliency may generate distinguishable respondent and nonrespondent attitudes – attitudes that stem from the variable of interest, unlikely to be available in the sampling frame, and thus by definition cannot be specified. This section will first define the term "topic saliency" and "satisfaction". It will be followed by an exploration of the complex linkages between satisfaction and saliency. The section concludes with an assessment of empirical findings of satisfaction bias.

2.2.1 Topic Saliency: Definitional Elements

A survey topic is salient to the sampled individual when it highlights events or situations that are "standing out from the rest; noticeable; conspicuous; prominent" (Dillman, 2000: 155). These can be understood in terms of a life changing event, "one that marks some sort of turning point in one's life, an event after which some portion of one's life is different from what it was before" (Groves, 1989: 430-431); or might be understood in terms of events (behaviours and beliefs) central to one's life. According to Heberlein & Baumgartner (1978), a survey topic can be considered "very salient" if it deals with "important behaviour or interests that were also current"; it is "possibly salient" if it pertains to "important issues or behaviours that were not necessarily current or timely"; and, it is non-salient if the survey topic "neither concerned important issues or behaviours nor [was] current" (Heberlein & Baumgartner, 1978: 449). Topic saliency is of particular importance to survey practionners due to its known positive impact on the decision to respond to a survey request (Goyder, 1987; Groves, Presser et al., 2004; Groves, Singer, & Corning, 2000; Heberlein & Baumgartner, 1978).

The literature reveals three dimensions to topic saliency[19]. First, the motivation to respond may come from topic interest. Individuals may be attracted to, or at the very least not repulsed by certain survey topics (McDaniel & Madden, 1987; Senf, 1987). This is equivalent to a "possibly salient" topic, and depending on the immediacy of the interest in question, a "very salient" topic. Second, the decision to respond may be based on topic

[19] The reader should understand these dimensions as tentative typologies and not as mutually exclusive categories. Although the literature does not specify in great detail how these typologies inter-relate, it is likely that three dimensions will shade into one another.

involvement. An individual involved in specific behaviours or practices may find survey topics addressing issues or questions related to his or her involvement salient and worthy of response. Again, the issue of immediacy made salient by on-going behavioural/practice involvement may underpin a "very salient" survey topic. Moreover, the notion of self-interest is key: individuals who were members of voluntary associations (Donald, 1960; Martin, 1994; Roose, Waege, & Agneessens, 2003), or were interested in recycling practices (Kojetin, Borgida, & Snyder, 1993), computing (Goyder, 1987), or a consumer product (Van Kenhove, 2002), were all likely to respond in greater numbers to a survey topic in their respective field of interest. Third, the impetus to respond may be driven by topic intensity. Individuals may hold strong opinions about some or all aspects covered by the survey topic, and choose, as a result, to respond to the survey request. They may wish to register their discontent, their support, or their general opinion on the topic at hand (Benson, 1946; Pearl & Fairley, 1985).

To posit that satisfaction surveys might be salient to a given sampled population says very little as to how satisfaction can be deemed salient and to which three dimensions it pertains. A closer look at the literature suggests that topic saliency as it relates to satisfaction surveys cannot be understood in abstraction of the manner in which satisfaction appears to the minds of sampled individuals. To get a clearer picture of the linkages between topic saliency and satisfaction, we must first delve into the nature of satisfaction, and secondly trace their origins within topic intensity.

2.2.2 The Nature of Satisfaction

The nature of satisfaction, as a state of being, has not yet achieved conceptual clarity. Babin (1998), basing upon his literature review on the nature of satisfaction, demonstrates how widely varied the concept of satisfaction seems to be: 1) "consumer satisfaction with a product refers to the favourableness of the individual's subjective evaluation of the various outcomes and experiences associated with buying it or using it" (Hunt, 1977); or 2) "satisfaction may best be understood as an evaluation of the surprise inherent in a product acquisition and/or consumption experience" (Oliver, 1981: 128). As a result, it is extremely difficult to distinguish satisfaction from other conceptual constructs. As Babin (1998) notes, one may easily conflate satisfaction with "perceived performances, disconfirmation, happiness or decision regret" (Babin & Griffin, 1998: 128). It is also clear that satisfaction is a complex mix of individual expectations, intents, experience, and personality. A formal definition is offered by Oliver (1997):

> Satisfaction is the consumer's fulfillment response. It is a judgement that a product or service feature, or the product or service itself, provided (or is providing) a pleasurable level of consumption-related fulfilment, including levels of under- or over-fulfillment (Oliver, 1997: 13)

Satisfaction, as the outcome of a cognitive process, as opposed to a state of being, has been conceptualized as the expectancy disconfirmation model[20], otherwise known as "gap

[20] The expectancy disconfirmation model, while still dominant, has come under greater scrutiny due to conceptualization and measurement problems associated with the inclusion of two dimensions (i.e. "expectancy" and "disconfirmation") within the same indicator (Cronin & Taylor, 1992). The service quality (SERVQUAL) instrument in particular was severely criticized by (Teas, 1993). This latter author proposed an alternative model based on affect. A firestorm of criticism of Teas' model and SERVQUAL ensued. See (Cronin & Taylor, 1994; Dyke, Kappelman, & Prybutok, 1997; Parasuraman, Zeithami, & Berry, 1994; Teas, 1994)

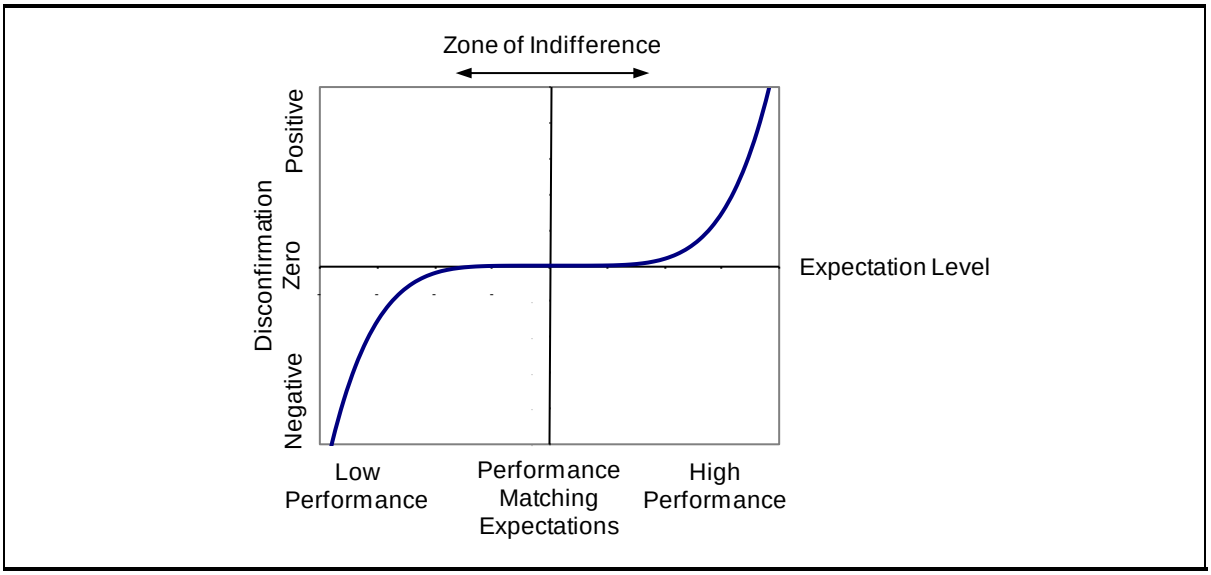
models" (Silva, 2000; Stallard, 1996). Satisfaction is the outcome of a comparison between one's prior expectations about a given service or product and one's experience. When experience exceeds expectations, the model predicts the expression of satisfaction; an experience which falls below expectations will translate into dissatisfaction. However, when expectations are met, no change in satisfaction or dissatisfaction will occur. That is to say that meeting expectations may not always lead to satisfaction. It is entirely conceivable, for example, that one may initiate a service request or a product purchase with very low expectations and receive exactly what was initially expected (Oliver, 1997). This interplay between expectations and service outcomes leads one to conclude, perhaps erroneously, that satisfaction is a conscious rational process, and by extension that satisfaction is reducible to saliency. The connection between these two concepts is explored next.

2.2.2.1 Satisfaction and Topic Intensity

One of the crucial assumptions regarding topic saliency is that a survey topic will bring forth images, experiences and emotions to such an extent that they will serve as a motivational force to cooperate with the survey request. Yet, cursory self-introspection will reveal that we rarely attribute satisfaction scores to every event we happen to be involved in. If this assumption is correct, it is reasonable to deduce that satisfaction does not necessarily lead to saliency and as a result would not immediately be of concern for nonignorability. As Mori (2002) is quick to point out, gap models leave the impression of a linear relationship between what one may have had and what one may have wanted. It could be construed, for example, that for expectations held constant, a change in one's experience with a given service will translate into a change in one's satisfaction level. Several works in this area have

shown this to be incorrect. Two cognitive processes will be reviewed: the zone of indifference and satisfaction as a dual construct.

Figure 2-8 Zone of Indifference

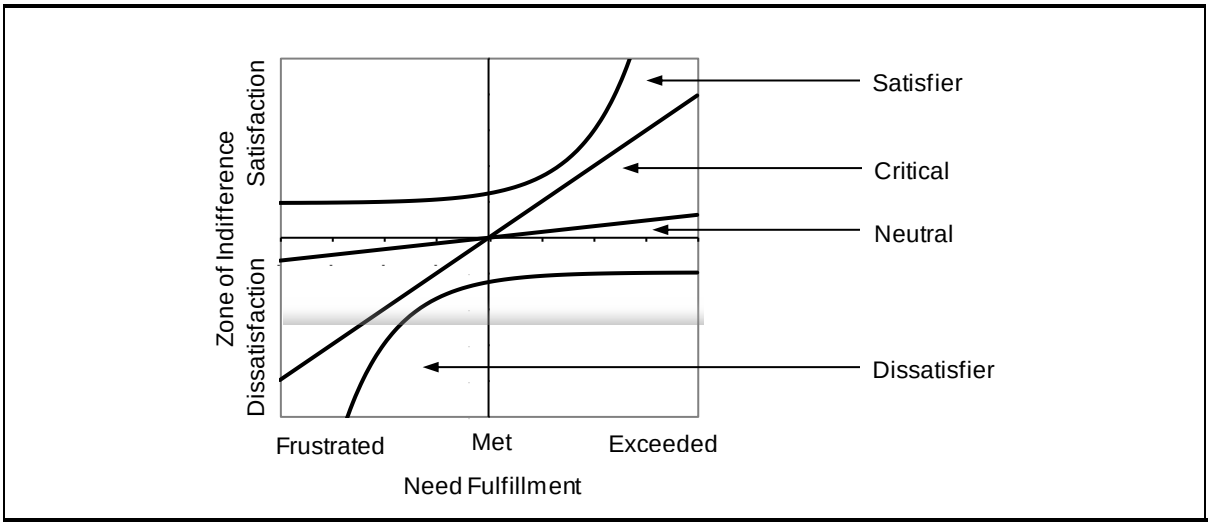


Source: (Oliver, 1997: 113)

Zone of Indifference. The zone of indifference can be understood as the outcome of a rational assessment to the extent that people will "absorb some positive or negative disconfirmation of expectations" (Mori, 2002: 25). It is, from a normative point of view, a pragmatically derived range of acceptability. Figure 2-8 illustrates the consequence of this assessment on perception. Satisfaction, the evaluative outcome of perception measured against expectation, will remain unaltered because performance variations are either not noticed or simply disregarded. Once the zone of indifference has been exceeded, satisfaction levels can change dramatically. The actual shape of this change can be linear, or most likely geometrical. We would postulate that outside the zone of indifference, patience has run its course and dissatisfaction will change suddenly; or, conversely, outside the zone of

indifference, unsolicited acts of kindness will produce changes in satisfaction levels. At the zone of indifference, the satisfaction may rest equally at the satisfied or no opinion level. At that specific point, perception will be a pragmatic meeting of expectations.

Figure 2-9 Satisfaction – Dissatisfaction Nonlinearity



Source: (Mori, 2002: 26; Oliver, 1997: 152)

Dual Construct Satisfaction. The concept of satisfaction as a dual attitudinal construct can be traced to the works of Herzberg (1966) on employee attitudes. In a study of accountants and engineers, Herzberg (1966) found that being satisfied was related to "motivational factors" such as self-actualization; their absences lead to no satisfaction – as opposed to being dissatisfied. Conversely, being dissatisfied was related to the lack of "hygiene factors" such as salary and good working conditions. Should the hygiene factors be present in abundance, participants expressed not being dissatisfied – as opposed to being satisfied. For Herzberg there are two constructs associated with satisfaction; there are "satisfiers" ranging from being satisfied to no opinion, and "dissatisfiers", ranging from being dissatisfied to no opinion. If one did receive proper amounts of contextual factors such

as good salaries, but didn't get motivational factors such as advancement or achievement, the individual would fall in a dead zone of satisfaction, neither satisfied nor dissatisfied.

This dual construct can be seen in Figure 2-9. No matter how frustrated one may be in regard to unmet motivational factors such as the provision of advancements, or increased role responsibility, satisfaction will remain at a no opinion level. It is only when these goods are provided that a perceived satisfaction will be registered, hence the name "satisfiers". The reverse is also true in the case of dissatisfiers: when good salaries are provided, perceptions of dissatisfaction will remain at a no opinion level; lowering salaries will undoubtedly register dissatisfaction. The same can be said in the educational sector for teachers' class preparedness as a dissatisfier. Sloppiness is likely to generate dissatisfaction. An increase in teaching performance, as in coming to class prepared, will not generate satisfaction, but simply no dissatisfaction. In this instance, teaching performance is detected by its absence. An overabundance of effort towards class preparedness, such as PowerPoint presentations drawn by Rembrandt, will generate neither dissatisfaction nor satisfaction. It is neither expected nor required.

This is not to say that linear relationships do not exist. There are one-dimensional satisfaction constructs, especially where time or responsiveness is involved (Mori, 2002: 24). One can, for certain issues, place satisfaction and dissatisfaction on the same continuum. In the case of education, returning the student's work on time is a clear one-dimensional construct. Should one complete a given task before expectations, it is likely to generate satisfaction, past the deadline, growing dissatisfaction (critical line). The critical, most salient features of undergraduate experience are reported as student-centeredness, campus climate,

instructional effectiveness (class preparedness and quality of instruction), and faculty-student interaction (Cotten & Wilson, 2006; Elliott & Healy, 2001; Krahn & Bowlby, 1997; Thomas & Galambos, 2004); for graduate students, quality of interpersonal relationship with supervisor, supervisor accessibility, financial support, and supportive climate (Dumaresq & Lambert-Maherly, 2002). That being said, not all student experiences cross the threshold set by the zone of indifference. Performance which is unrelated to prior expectations will likely not have any impact on one's satisfaction (neutral line).

To sum up, satisfaction is not just a linear function of whether events meet our expectations: the more distant events are from our expectations, the more satisfied/dissatisfied we will be. Some events are unrelated to our expectations; others are noticeable only by their absence; still others only by their presence. Moreover, there is an unknowable range of indifference, whereby people will simply not alter their satisfaction levels despite obvious changes in their experiences. In fact, for satisfaction to be truly salient, in the sense of being involved and interested, requires a definite breach of expectations, positive or negative. Satisfaction levels, which fall into the zone of indifference either, are neutral events, or are congruent with expectations. The implication for nonignorability is that satisfaction, unless generated by large disconfirmation, is not in and of itself conducive to topic saliency. As a result, satisfaction salience is a function of topic intensity, which can be brought about through topic interest and/or topic involvement.

2.2.3 Empirical Studies on Satisfaction Bias

This section summarizes the findings of empirical studies on the relationship between satisfaction and nonresponse error. Table 2-1 presents studies of satisfaction bias conducted within various organizations such as hotels (customer surveys), hospitals (patient surveys), employment (employee surveys), and education (student satisfaction). It lists the target population, the survey design, response rate obtained, and findings of bias. There are two types of studies excluded from this list. All studies on nonresponse bias that did not specifically pertain to satisfaction were systematically excluded. Moreover, excluded from this list are studies of nonresponse that did not explicitly test whether response differentials had an impact on satisfaction variables. The goal is to ascertain the presence of satisfaction bias, at what response rate, and its direction.

A glance at the findings of the studies reported in Table 2-1 reveals a chaotic pattern of nonresponse rates and satisfaction bias. Indeed, satisfaction, particularly intense emotions, was presupposed to be sufficiently salient to generate response patterns conducive to nonresponse error. Patient surveys, for example, were presumably susceptible to nonresponse error because responding patients were more likely to evaluate their care more favourably (I. Green, 1991). Yet, four studies reported non-significant differences between respondents and nonrespondents for variation in response rates ranging anywhere from 33% to 80%. Admittedly, findings of no bias are problematic because they assume the existence of a sound methodological design to prove the null hypothesis. In the case of two studies, that of Lasek et al. (1997) and Krushat and Molnar (1993), there is insufficient statistical power to detect

only but large (i.e. 10%)[21] deviation from the sampled population proportions. In the case of Moum (1994), the original sample of 85,100 was reduced to 74,977 when eligible sampled individuals were removed due to their frequent absence, illness or disability. While the sample size difference is small, those who may have had a stake in their quality of care have been systematically excluded rendering the sample more homogenous.

Perhaps more in line with expectations, the balance of satisfaction bias studies have reported the presence of bias despite variation in response rates from 1% to 81%. The recruitment of a wider satisfaction spectrum population, through minimal collection efforts, seems to reduce the initial dissatisfaction bias. Thus, Whipple and Muffo (1982) alumni survey showed a significant bias between early and late respondents: the latter were less critical about the program, material, quality of teaching, faculty contact, and overall satisfaction of program. Perneger, Charnot and Bovier (2005) wave analysis showed a positive bias on the patient experience scores where late respondents reported more problems with their perceptions of health care received. In the same vein, Allen (1998) wave analysis also showed a positive bias for consumer health plan experience assessments. Trice and Layman (1984) hotel-guest survey found that an increase in response rate from 1% to 20% produced a substantive increase in the proportion of satisfied hotel guests. In the case of a better quality survey, Hirdes et al. (1998) patient care study found that response was biased towards respondents who had the best experience and reached the highest level of self-

[21] In Krushat and Molnar (1993)'s study, for example, a post-hoc sample of 107 nonrespondents was utilized to assess the presence of bias against to initial set of respondents. At N=107, for an alpha level of 0.05, the expected precision levels are set at about ±10%.

functioning outcomes, as opposed to nonrespondents who were mostly too ill to respond and could have different ideas about patient care.

Other studies, using more innovative methodological techniques, have also reported significant satisfaction bias. Rogelberg et al. (2000) studies on employee attitudes towards organizational surveys showed that those who were less likely to report a desire to participate had greater intentions to quit their employment, were less satisfied with their work, and exhibited lower organizational commitment. Rogelberg et al. (2000) study operating on the distinction between passive nonrespondents (convertible refusals) and active nonrespondents (terminal) reported the former to be closer to respondents, and the latter showing significant differences on satisfaction with sponsoring agency; other satisfaction scores remained unbiased however. Barkley and Huxley (1992) study of hotel guest satisfaction revealed perplexing results. Using three quota sample groups, Hi-Q given a $10 discount, Med-Q given a drink coupon, and Lo-Q given no incentive (see Trice), the authors compared each group against a battery of satisfaction items. The Lo-Q group was found to be the most unstable and inaccurate, generally dissatisfied with their accommodations and room rate, compared to other quota groups. The authors conclude, for the Lo-Q group, that "without the benefit of an external motivation, guests who fill out surveys and questionnaires are more likely motivated by an internal force: disappointment, pleasure and boredom" (Barsky & Huxley, 1992: 24).

To summarize: It can be deduced from both hotel and alumni empirical studies that, at very low response rates (<14%), the modus operandi of response behaviour is dissatisfaction. If, as is the case for most hotel surveys, one's overall sampling strategy

constitutes a response card on the coffee table, survey results have shown a consistent pattern of dissatisfaction from customers. Minimal attempts to draw from a wider pool of hotel guests, such as a survey request during hotel checkout, have shown a dramatic increase in satisfaction levels. A sampling strategy based solely on self-selection is likely to yield a strong dissatisfaction bias.

A sampling strategy designed to gather responses from a wider pool of individuals is likely to attract the more involved, and the more satisfied first. This is corroborated by empirical studies on employee satisfaction surveys. This conclusion is also supported by health studies on patient satisfaction, which show a consistent positive bias amongst in-patients as opposed to outpatients. It appears that being part of an organization for an extended duration of time, either as an employee or as an in-patient, is a strong motivator to respond, and in particular for those who had a satisfactory experience.

It remains unclear, however, how the negative bias, as discussed in the hotel studies, and the positive bias, as discussed in the previous paragraph, relate with one another. Are these separate classes of surveys, one being a customer survey, the other an organizational survey, producing in each different response bias behaviours? Alternatively, are we to conclude that a very early negative bias (<14 % response rate) will be accompanied by a subsequent positive bias? What is clearer is that studies that have exhibited bias seemed to have diluted their initial satisfaction bias with the recruitment of a wider spectrum of satisfaction in the respondent pool. However, four studies did not find any significant satisfaction bias at all. It would seem that no decisive consensus could be reached as to nonresponse ignorability for satisfaction surveys.

Table 2-1 Summary of Satisfaction Bias Studies

SATISFACTION BIAS WORK SUMMARY RESULTS				
Author(s)	**Population**	**Method**	**Rate**	**Findings**
(Ley, Bradshaw, Kincey, & Atherton, 1976)	In-patients Non-random N= 158	Wave analysis From last survey in a multiple survey Cohort study	39.8%	No significant differences in reported satisfaction with communication patient-staff, between respondents and nonrespondents
(Whipple & Muffo, 1982)	Alumni Random N=1749	Wave analysis	14.5%	Significant bias found between early and late respondents. Late students were less critical about program, material, quality of teaching, faculty contact, and overall satisfaction of program.
(Trice & Layman, 1984)	Hotel guests Non-random	Survey design comparison	1% to 20%	Uncontrolled test, but substantive differences in the satisfaction ratings from no design to a minimal design.
(I. Green, 1991)	Patient Census N=864	Wave analysis	33.1%	Controlled for the treatment received, length of stay, number of therapy sessions, there was no significant difference between early-late
(Barsky & Huxley, 1992)	Hotel guests Systematic sampling N=100,100,100	Three sampled populations; first two were systematically polled at exit; third was retrieved from the hotel guest room, voluntary only	100% (quota)	Systematic differences between polled surveys with systematic requests at check-out. Average expectations met higher for hotel guest room surveys, more critical responses with polled surveys. Different customer base.
(Krushat & Molnar, 1993)	Social Security Random N=646	Nonrespondent follow-up	80.2%	Nonrespondents had lower levels of satisfaction but statistically non significant
(Grosset, 1994)	Alumni Random N=1936, 1740, 1547	Follow-up survey of original alumni survey	47%, 51.2%, 44.6%	Respondents scored higher on persistence and academic engagement indicators (GPA, credit earned, enrolled semesters)
(Moum, 1994)	Test-screenees Census N=74,977	Dual survey (wave analysis)	76%	Controlling for socio-demographic and impairment, there were no significant relationships between life satisfaction and response rate; no bias was found.

Table 2-1 Summary of Satisfaction Bias Studies (Continued)

SATISFACTION BIAS WORK SUMMARY RESULTS				
Author(s)	**Population**	**Method**	**Rate**	**Findings**
(Etter, Perneger, & Rougemont, 1996)	Patient Non-random N=395	Follow-up of outpatient survey; 2x2 experiment with different letterheads (sponsorship medical versus university)	80.5%	No statistically significant findings of bias for sponsorship effect (between factorial groups). Early respondents more satisfied, than late respondents (only on university sponsor).
(Lasek et al., 1997)	Patient Random N=16,267	Satisfaction survey, mail first follow-up; phone second follow-up Wave analysis	54%	Satisfaction higher among mail versus phone respondents Satisfaction differences between respondents and nonrespondents were small (5 points on 100 point scale)
(Allen, 1998)	Employee Random N=14587 (1993) N=9294 (1995)	Longitudinal panel study 1993-1995. Nonrespondent follow-up	51% (1993); 52% (1995)	Early respondents more satisfied with health plan performance than late respondents
(Rogelberg et al., 2000)	Employee Non-random N=153	Personal interview Profile pre-survey Job attitude survey	100%	Anticipated non-compliance showed less satisfaction with work and less commitment to organization
(Rogelberg et al., 2003)	Undergraduate Non-random N=405	Profile pre-survey Two mail survey: Pre-Post(Wave)	Pre-survey 100% Post survey 26%, 20%	Active nonrespondents are less likely to be satisfied with organization. Passive nonrespondents do not differ from respondents
(Perneger et al., 2005)	Patient Census of discharged patients N=2156	Wave analysis	70%	Early respondents more satisfied with care received than late respondents

2.3 Determinants of Topic Saliency

As we have seen in the previous section, there is ample evidence to suggest that topic saliency is a strong motivator to cooperate with a survey request. Judging by survey responses to satisfaction surveys, however, topically motivated respondents are not automatically different from nonrespondents. In an attempt to harness when topic saliency is ignorable and when it is not, this section will elaborate the relationship between satisfaction and survey response. To that end, common-cause determinants of response behaviour and of satisfaction will be explored. Where common-cause determinants are not be explicitly found, further determinants of satisfaction with the university graduate experience will be discussed as potential avenues for topic saliency bias. Following the survey cooperation conceptual framework elaborated in Figure 2-1, we will begin with the student and contextual determinants of topic saliency (e.g. "out of researcher control"), to be followed by survey design (e.g. "under researcher control"), and conclude with the cognitive determinants of topic saliency (e.g. "student-survey interaction").

2.3.1 Influences of Student Characteristics

Alumni surveys are often analyzed in terms of organizational statistics such as socio-demographic, enrolment, and departmental statistics. Such statistics are important because they form the primary language through which an organization will understand its population in terms of its own organizational goals. Basic issues such as student-faculty contact and research activity of students are essential in terms of allocation of resources for each degree program. Equity issues, whether one refers to the equitable treatment of minority groups or more specifically to the fairness of income distribution, are of primary concern for most

public institutions and universities. An accurate representation of its population is essential for the organization's understanding of its student population. However, since student experiences will be determined by the success or the failure of these organizational policies, survey responses may not be an accurate representation of the student population if the latter base their survey cooperation on these experiences. Survey responses may reflect an overly optimistic (or pessimistic) picture of student experiences. This section focuses on enrolment characteristics such as gender and visa student registration.

A review of alumni surveys provides some indication of bias amongst enrolment variables. Women are more likely to respond to an alumni survey (Dey, 1997; Lasek et al., 1997; Planning, 1996; Research, 2003; Stopher & Sheskin, 1981) and are also more likely to encounter gender inequality in terms of faculty representation and harassment (Jacobs, 1996), gender stereotypes (I. Lewis, 1984; Morrison, Bourke, & Kelley, 2005), streaming towards more feminine roles (Levy, 1982), chilly climate and socio-economic support (Leonard, 1997). In addition to gender, one's race also relates to survey response. Black students are less likely to respond to a university survey (Dey, 1997). They are more likely to express lower satisfaction with their degree program in terms of cultural climate (Ancis, Sedlacek, & Mohr, 2000) and campus diversity (Einarson & Matier, 2005). Other enrolment correlates of survey response such as high school grades, self-reported academic ability (Dey, 1997) and GPA (Porter & Whitcomb, 2005) will likely translate into bias; survey responses are likely to over represent the most academically successful students. Finally, students receiving financial aid will also be less likely to respond (Porter & Whitcomb, 2005), and possibly be more likely to experience financial stress (Bone, 2002). Findings from an early study of

student survey bias point to length of stay in college and rural upbringing/family background as two other determinants of cooperation (Reuss, 1943).

No differences were observed amongst several other important enrolment variables. Degree program has not been shown to be a predictor of survey cooperation. This is unexpected because program duration, tuition fees and employment prospects are known to play important roles in the experiences of Master's students (Bone, 2002), and could have been deemed salient since such topics are routinely broached in a student satisfaction survey. International student registration (visa students), does not relate to response behaviour. Visa students must pay higher tuition fees, experience difficulties in securing sources of income in terms of grants and campus employment, and are more likely to experience "transition issues" and "isolation feelings" (Andres & Carpenter, 1997). Yet, another study points out that visa students do not expect as many university resources such as teaching, technology, library services, as their local counterparts. These services are not perceived as necessary preconditions to expressing satisfaction with their degree program (Tsarenko & Mavondo, 2001). The literature would seem to suggest that degree program and visa status are not related to the decision to respond to a survey, and thus are ignorable (missing completely at random).

The question is whether the biases introduced by gender, race, academic success, financial aid, length of stay and rural background are ignorable as well. Certainly, as we have seen in the previous section, these enrolment variables are clearly not missing completely at random (MCAR); they are related to the propensity of responding to a survey request. The issue remains whether nonresponse satisfies the missing at random condition (MAR). If the

probability of obtaining satisfaction responses depends on enrolment variables, and thus the probability of responding may vary by race or gender but not on one's satisfaction level, then MAR condition is satisfied. In other words, the mean satisfaction levels within a given category are not expected to change for an increase in response rate. If these enrolment variables are available for the sampled population, then the bias can be corrected. If these variables are not available or if satisfaction levels do relate to the probability of responding, then the bias is nonignorable. One would expect that an increase in response rate would also be accompanied with a change in the mean satisfaction levels for a given category of enrolment variables. Thus for gender, race, academic success, and financial aid to be nonignorable means that female, black, academically weaker and financially stressed students are responding to the survey according to their satisfaction levels.

However, to conclude that these inequalities are topically salient, and that nonignorable nonresponse will ensue from population groups subjected to such inequalities, is controversial on two counts. First, it is not entirely clear that a student or alumnus would specifically cooperate with a survey request based on his or her minority group status. As Goyder (1987) would point out, "rarely do people reason out a decision about responding to a survey on socio-demographic grounds" (Goyder, 1987: 183). To say that survey responses are based on such considerations presumes a high level of socio-political consciousness[22]. A study of perceptions and experiences of gender inequality, for example, clearly shows that one's perception of gender stereotypes does not always translate into a recognition of gender

[22] Of course, it doesn't mean that such status would not predispose one to respond or not. It need not be a conscious factor in a decision one makes. However, topic saliency presupposes some elements of consciousness as to why one responds. Consequently, the decision to respond to a survey based on structural inequalities such as race or gender requires a modicum of political consciousness.

inequality (Morrison et al., 2005: 154-155). Second, it is presupposed that institutional inequalities are systematically operating across all educational institutions and operating on all minority students equally. The "chilly climate thesis", for example, has not been consistently supported (P. Grayson, 1999). For these inequalities to be topically salient seems to require clear and systematic existence of structural inequalities and presence of a socio-political consciousness. It is uncertain if and when these two factors will converge to influence the decision to respond to a survey request from that particular institution.

2.3.2 Influence of Organizational Characteristics

The analysis of alumni satisfaction surveys using socio-demographic and program-related statistics provides some understanding of graduate experiences in terms of organizational goals set by the university. In recognition of the hierarchical structure of the university, these analyses are not only conducted on the university population taken as a whole, but also, and often more importantly, at the faculty and departmental levels. The latter is particularly important since departments and faculties are the primary conduit through which most of the university's resources funnel to the student, and the area in which the student interacts with peers and faculty alike. It is reasonable to conclude that students who have experienced their graduate program within the same department may come to think in similar ways about their satisfaction with their degree program. It is equally conceivable that such similarity may translate into variations of response rates across departments. Consequently, alumni surveys may not provide an accurate portrait of student experiences within a given department.

Contextual analyses of survey nonresponse, albeit very limited, suggest some potential sources of bias. Following Groves and Couper (1998) analyses of survey nonresponse, there are four basic contextual effects applicable to alumni satisfaction surveys conducted within a given institution. Survey nonresponse may be susceptible to "decline in civic duty" (alienation), "over-surveying effect", population density (crowding) and social disorganization (low cohesion) (Groves & Couper, 1998: 155-170). These would translate in the realm of alumni surveys as disengagement with one's institution, survey fatigue, departmental crowding, and departmental cohesiveness. We will review in turn all four effects and their implication for ignorability. The first two are cultural in nature, the remaining two focuses on department. These will be complemented by a fifth contextual determinant of student satisfaction which may have nonresponse implications.

Survey fatigue. During their stay at a given institution, it is likely that students will be asked to participate in an array of satisfaction surveys with topics ranging from library services, to career and counselling services, to degree program evaluation. As Porter (2004) rightly pointed out, it is entirely conceivable that some of these survey requests may occur back-to-back or even simultaneously. These multiple requests have an impact on survey cooperation rates. In an experiment conducted upon undergraduate students at a liberal arts college, Porter (2004) observed that back-to-back requests have the strongest effect on nonresponse. Interestingly, topic saliency may have interacted in subtle ways. While it is clear that students may have been more interested in one survey topic over another, blurring some of the observed effects of survey fatigue, it is equally possible that multiple survey

requests enhanced topic saliency[23]. Further study on the effect of over-surveying produced no observable differences in terms of students' attitudes towards the survey and its sponsor - despite being contacted nine or more times (McCarthy & Beckler, 2000). From these results, one may hypothesize that multiple non-salient survey requests will result in greater nonresponse, but nonresponse will be ignorable. It is likely however, that one survey topic amongst all others requested will be salient. Consequently, it is likely that multiple simultaneous survey requests are nonignorable.

Disengagement. To the extent that survey cooperation is akin to cooperation to other "forms of participation in the polity", survey nonresponse may be linked to an increase in alienation and a general disengagement of the general public in "civic duty" (Groves & Couper, 1998: 170). In the realm of education, an analogous phenomenon has been observed as a "culture of disengagement" (Hu & Kuh, 2001). According to Astin (1999), a disengaged student is one who "neglects studies, spends little time on campus, abstains from extracurricular activities and has infrequent contact with faculty members or other students" (Astin, 1999: 518); presumably such a student would be less likely to reciprocate by cooperating to survey request. It should be no surprise that student engagement is strongly related to program satisfaction (Hu & Kuh, 2001; Pike, 1991) and that student engagement is related to the decision to cooperate to a survey request (Porter & Whitcomb, 2005). Contextual influences on student engagement, however, are mixed. Institutional emphasis on acquiring vocational skills had a positive relationship with engagement, but emphasis on

[23] Increased stimuli may simply force greater impetus to choose which survey request to honor. A fruitful avenue for the conceptualization of this relationship might be the analyses on information overload and boredom (Klapp, 1986).

scholarly activities and congenial relations did not (Hu & Kuh, 2001). Research-based schools (ratio of graduates to undergraduates) are less likely to be engaged and less likely to produce cooperation amongst first year students with a mail survey and senior students in a web survey (Porter & Umbach, 2006); the relation might be reversed for graduate students for whom research is at the center of their university experience. Public schools are less likely to be engaged than private ones, and less likely to respond to a mail survey; part-time students who are also (by definition) less engaged are marginally less likely to respond to a web survey (Porter & Umbach, 2006). The empirical evidence would at least suggest that student engagement has both an individual and contextual characteristic; together these characteristics have an impact on one's satisfaction. Whether engaged students located in engaging departments are more likely to participate remains an open question. Engaging departments, measured as frequency of interaction between student and faculty, will have mixed results on satisfaction: class size is unrelated to program satisfaction but related to personal and general development; the pool of full tenure professor is marginally related to program satisfaction (Umbach & Porter, 2002).

Departmental crowding. Groves (1998) defines crowding as an experience that may result from population density: "experience of excessive social encounters in high-density areas that lead to social overload" (Groves & Couper, 1998: 176). The author goes on to suggest that this density of interaction has the overall effect of curtailing social exchanges and restricting them to social relations between "friends and family" (Groves & Couper,

1998: 176)[24]. There is virtually no research on the impact of departmental indicators of crowding such as office space and enrolment size on students' satisfaction level. One study reports that students in densely packed schools, measured as the number of students per acre of campus, will be less likely to participate in web surveys; and students registered in urban schools, as opposed to rural areas, will be less likely to respond to a web survey (Porter & Umbach, 2006). Whether these indicators are precise measures of crowding, and whether crowding itself is a useful measures remains debatable, one could minimally make the contention that larger and ill-equipped schools and departments unable to accommodate all their students may be prone to lesser response rates and greater dissatisfaction.

Departmental cohesiveness. Cohesiveness (lack of) relates to the concept of social disorganization and lack of social control. Individuals are less integrated with one another, less willing to interact with one another, and less likely to recognize functional authority (Groves & Couper, 1998: 177). Indeed, a study conducted on the impact of work groups on employee satisfaction shows that cohesive groups are more likely to display organizational commitment, more courteous, more likely to create a favourable climate for exchange and helping behaviour (Kidwell, Mossholder, & Benett, 1997). In another study on substance use treatment, similar contextual results were obtained using measures of collective efficacy;

[24] There are undoubtedly intuitive links to be made between crowding and the concept of disengagement elaborated above. The effect of dense social relations can be construed as one of disengagement to the extent that students may opt out of social relations with their peers, their professors, and presumably with their studies as a whole. However, the conflation should be avoided with proper conceptual focus. Student disengagement refers to the lack of commitment and energy dedicated to one's studies, to enact and sustain graduate student roles. Crowding refers to the density of social relations within a given area; it results in less interest in the affairs of all (e.g. departmental or university-wide affairs) except those of close graduate peers. In that sense disengagement and crowding are separate entities. Crowding and disengagement may lead to the same outcome but the former does not preclude the possibility of the latter, that is to say, to uphold and maintain one's role as graduate student. Crowded spaces should not be conflated as causes for the lack of time dedicated to studying, reading, writing or researching.

communities exhibiting higher collective efficacy were more likely to influence an individual living in the given community to respond to a survey request, and to have an impact on the values obtained in the survey (Johnson, Cho, Campbell, & Holbrook, 2006). Within educational settings, the concept of peer influence (as opposed to peer pressure) might create a more cohesive environment to the extent that such influence may create a "frog pond" effect (Hox, 2002). One such measure, average GPA, seems to be ignorable: while average GPA (SAT scores) had a statistically significant but substantively small impact on survey response to mail surveys (school level) (Porter & Umbach, 2006), it had no significant impact to the evaluation of program satisfaction (department level) (Umbach & Porter, 2002). Furthermore, the concept of group diversity may play a role in how a given student assesses their degree program[25]. Diversity measure, such as proportion of female and non-white undergraduates showed marginal results: the former was marginally related to satisfaction with one's major (undergraduate), the latter remained nonsignificant (Umbach & Porter, 2002). In sum, the first set of studies, suggest that nonresponse will be nonignorable if students within cohesive departments are more willing to cooperate with a survey request (helping behaviour) and as a result of this cohesiveness have a satisfactory experience; the second set of studies, albeit with imperfect indicators, seem to suggest an ignorable outcome.

2.3.3 Influences of Survey Design on Topic Saliency

Thus far, we have summarized determinants of topic saliency, those which, according to Groves' conceptual model of survey response (see Figure 2-1), remain outside the

[25] See (Chang, 1996; Millem, 1998) for a discussion of the impact of diversity on college students, both cited in (Umbach & Porter, 2002)

researcher's control. What remain under the researcher's control are all aspects of the survey design such as the choice of survey mode, the establishment of a survey protocol, questionnaire design, and the administration of the survey. The question is whether the absence or presence of such design features, designed to maximize response rates, will also produce or inhibit topic saliency[26]. The literature suggests three general trends: 1.) survey designs can highlight what is most salient about the survey; 2.) survey designs can amplify or thwart saliency effects with an overly burdenful survey or with the usage of incentives; 3.) survey designs can trigger measurement errors, which prevent one from detecting nonignorability.

2.3.3.1 Survey Mode

Though the answer may seem obvious, it is worth asking the seemingly simple question: How does the sampled individual know what the topic of the survey actually is? Is it derived from the sum impression of questions in the questionnaire? Is it what is said during the introductory component of the survey request? Is it a combination of both? The literature on topic saliency is replete with assumptions about the saliency of the topic to the respondent and yet there is little research to provide clues as to how such a notion of topic is arrived at in the respondent's mind. To compound the problem, it is unclear whether the respondent is acting upon the "topic" as a whole, or a particular component of the said "topic".

[26] To remind the reader: Our use of topic saliency refers explicitly to an earlier discussion between topic saliency and satisfaction. It was ascertained that satisfaction is topically salient to the extent that a clear departure from the zone of tolerance has occurred. Topic saliency as it relates to satisfaction survey is essentially one of topic intensity. It is conceivable that a given survey design may appeal, foil or even produce topic saliency stemming from the recalled intensity of the graduate experience.

The survey mode might play a role in the determination of the survey topic. In the case of a mail survey, the introductory component of the survey request is put forward in the advance letter and the cover letter. It is the cover letter that will indicate to the respondent what the purpose of the survey is, who sponsors it, and what is the confidentiality and voluntary nature of the request (Dillman, 2000). This cover letter is mailed along with the questionnaire. Telephone and face-to-face surveys, however, will begin their introductory components dynamically by voicing to the respondent the nature of the request (Groves & Couper, 1998). Respondents are not given any questionnaires to peruse, and are completely dependent upon the interaction with the interviewer. It is precisely this distinction between survey modes that has propelled (Groves, 2006a; Groves & Couper, 1998) to hypothesize the possibility of nonignorable nonresponse to mail surveys. With the questionnaire in hand, the potential respondent may not be swayed by the cover letter's content and choose instead to scan the questionnaire for the "real topic". As a result, mail surveys are by design disadvantaged towards uncontrollable topic saliency effects. Said differently, the interviewer has lost the ability to tailor what the potential respondent may perceive as the survey topic. It could be speculated that repeated reminders may only compound the problem (Groves, 2006a: 734)

Empirical research on the motives of people who refuse to answer a mailed survey is understandably rare. More common are studies conducted during phone and personal interviews. A compilation of respondents' remarks shows "I'm too busy" is the likely reply; if any questions ensue they are likely to be about the purpose of the survey, the completion time, and how the respondent was chosen (Groves & Couper, 1998: 234). The question is

whether this pattern of behaviour is also valid for mail recipients. A qualitative study conducted on twenty-three sampled business students shows the same interaction patterns noticed in other survey modes. The recipient quickly classifies mail as to its type (junk, bills, etc), then assesses the appearance of the items in the envelope, and only then considers the topic. Helgeson (1994) is somewhat ambiguous on this point, but seems to suggest that the survey topic is gathered from both the cover letter and the questionnaire itself. Of interest is the cost-benefit analysis between survey topic and questionnaire length: "I'll usually take a moment to look how long it is and decide whether or not I'm going to look at the subject-matter, and if it's something that I have some interest in, I might fill it out, if it's not too long" (Helgeson, 1994: 341). The study, albeit not generalizable, does lend support to Groves' contention that mail survey cooperation may be topic sensitive. However, evidence on this particular interaction with a self-administered questionnaire is extremely difficult to ascertain since other factors (see below) cannot be controlled for.

Once a "topic" is acknowledged, however, the sampled individual may systematically react to the interviewer, to the sensitivity of the topic, or to the appeal of the method (Nachmias-Frankfort & Nachmias, 1996). In a study which examined the differential effects of phone, mail and face-to-face modes on a Medicaid survey, it was observed that respondents who answered by mail were more likely to be in good health, 35 years and older, and because of the dual-language format, to include more Spanish speakers. Follow-ups executed by phone and then by face-to-face modes were more likely to attract respondents who spoke English as their primary language. Significant differences on health ratings were found during the last phase of follow-ups indicating a significant positive bias (Gallagher,

Fowler, & Stringfellow, 2005). In another study of patient satisfaction, it was also observed that mail surveys seem to attract more positive responses than those obtained through face-to-face interviews (Larsson, 2000). These findings were not replicated in a controlled field experiment which tested the impact of phone, mail and face-to-face modes on two structural models (De Leeuw et al., 1996). One possible explanation is that mode effects are an artefact. It was not the mode in and of itself that generated bias. Face-to-face interviews, which are known to generate greater response rates, uncovered a positive bias in the response pool.

In sum the literature on survey modes points to inconclusive results: mode effects on topic saliency, whether generated by questionnaire accessibility or respondent preferences, remain ad hoc. Questionnaire accessibility might enhance or even create a "topic" which the sampled individual assesses when deciding whether to cooperate with the survey request. In that case, topic saliency would be a strong motivation to respond. However, the topic might be threatening or embarrassing and the subject may choose not to answer using the given survey mode and/or use another survey mode when provided. Yet, survey modes have not been conclusively shown to be the cause of systematic nonresponse error.

2.3.3.2 Survey Protocol

A corollary problem to the use of multiple survey modes is the establishment of a sound survey protocol, here understood as the use of incentives and reminders to produce better returns for a survey request. A review of the literature indicates two protocol elements that may bear upon topic saliency: the impact of reminders and concerns for privacy.

Sending reminders to nonrespondents not only "reminds" students to send in their surveys, it also gives the opportunity to convince recalcitrant nonrespondents to cooperate with the survey request (Dillman, 2000). A diligent follow-ups implementation has been shown to increase response rates by converting refusals (Dillman, 2000: 314-319; Porter, 2004). The increase of response rate has been shown to dilute observed initial topic saliency effect (Groves, 2006a; Groves, Presser et al., 2004). However, sending multiple follow-ups may also be generative of an acquiescence bias whereby sampled individuals will not only acquiesce to responding but also to what the survey sponsor wants to hear. The problem, as one study seems to indicate (Ray & Still, 1987), is that reminders do increase response rates, and will produce a representative demographic profile, and yet recalcitrant respondents will say "yes to anything" (Ray & Still, 1987: 572). This may help to explain some of the incongruent results on satisfaction bias noted earlier. Consequently, the discovery of no significant differences in satisfaction between early and late respondents (Gasquet, Falissard, & Ravaud, 2001), may (or may not) be the result of measurement error. The danger is to assert nonignorability when in fact respondents have concealed their "true" answers by acquiescing.

A second issue related to survey protocol is whether the survey should be deemed confidential or anonymous. A confidential survey will contain some form of identification that will permit one to connect answers to a particular student name; anonymity provides maximum separation between student identity and their answers on the questionnaire. Assurances of confidentiality showed small effects on response rates, and seem to have greater impact on topics dealing with extremely sensitive topics (Porter, 2004). As Pryor

(2004) correctly points out, sensitivity of topics is not necessarily related only to the content of the questions but also to how the data will be handled, reported and distributed. In fact, concerns about privacy seem to hinge on "trust in the integrity of the data-collection agency not the nature of the assurance given to respondents" (Singer, Mathiowetz, & Couper, 1993: 479). Consequently, the issue of anonymity versus confidentiality hinges on the trust placed by the disaffected student in the institution requesting cooperation with a satisfaction survey. There is a possibility that the most irate and disaffected students will not choose to answer a survey if there is a possibility that answers can be traced back to them. Nonresponse, in that specific case, is likely to be terminally nonignorable as these students are unlikely to respond.

2.3.3.3 Questionnaire Design

Speculative hypotheses were introduced earlier about questionnaire accessibility and topic saliency. It was suggested that mail surveys, because of the accessibility of the questionnaire, were susceptible to topic saliency bias because one may deduce a topic by perusing the content of the questionnaire. This presupposes that the questionnaire design was constructed in such a manner as to present and sustain topic relevance by carefully structuring and organizing survey questions. Foddy (1993) makes the point that the questionnaire must establish the relevance of the topic in a clear and focused manner by inserting filter questions and filter response options, and by "making intelligible requests" so that the meaning of words is understood, questions are properly introduced to foster recall, and questions are properly framed according to the level of generality asked. In essence, the

questionnaire is a communication device through which the saliency of the topic is ascertained. In order for the survey variables of interest to have a role to play in the decision to cooperate with a survey request, the topic of the survey needs to be clearly laid out.

One of the ways in which survey designers establish topic saliency is by placing the most important set of questions first and the least salient (socio-demographic questions) last (Dillman, 2000: 87). If this strategy is correct, we should see a saliency effect when survey sections are systematically changed. Such a study (Robertson & Sundstrom, 1990) was conducted using an employee attitude survey whereby six topic areas were combined into twelve different questionnaires (6X2 factorial experiment); five with random topic placements and one organized according to employee priorities. Questionnaires with topics organized according to employee priorities had significantly higher response rates and significantly different mean attitudinal responses than randomized topics. As a result, topic order is a contributor to topic saliency effects, and nonignorable nonresponse.

2.3.3.4 Survey Administration

Survey administration, here understood as the interactive process between interviewer and the sampled individual, may contribute both directly and indirectly to topic saliency bias. In both cases, the manner in which the survey request is communicated will weigh in the balance as to whether one should cooperate or not with the survey request.

Survey administration has an indirect effect on topic saliency to the extent that the administrator's experience, expectation, motivation and supervision will determine how the survey process is enacted throughout and how the survey request is presented to the

respondent. In fact, key interviewer characteristics and skills such as tailoring and self-confidence (De Leeuw, 1999; Groves & Couper, 1998), self-efficacy (Lemay & Durand, 2002), training and motivation (Durand, Gagnon, Doucet, & Lacourse, 2006) have been shown to be essential components to ensure cooperation of the respondent. Experienced interviewers are more likely to successfully counter reservations against the survey request and to sustain an interaction with the potential respondent; self-confidence and experience translate into a greater number of questions answered in the survey (Groves & Couper, 1998: 242-243). Admittedly, these characteristics are not fully translatable for mail surveys' "interactions" with the survey respondent, but they do play a role in the successful enactment of contact follow-ups and refusal conversion through contact strategies and reminder letters (Dillman, 2000). The net result of a well-administered survey is the ability to mitigate against the effect of topic saliency by increasing the saliency of other features of the survey. When contact is badly followed through, and when minimal refusal conversion strategies are enacted, topic saliency remains one of the few reasons to answer a survey.

A closer look at the manner in which the survey request is presented reveals a direct effect on topic saliency bias. The common-cause factor is personalization within the context of a bureaucracy. To personalize cover letters and reminders is to go well beyond a word processor's ability to merge cover letter salutations with name databases; it is "what one would do in a letter sent to a business acquaintance who is not well known to the sender. It provides the look and feel of being from a real person, rather than a carefully programmed computer" (Dillman, 2000: 152). In an experiment conducted on community college alumni, a significant increase in cooperation rate was observed with a "warm" approach (personal

salutations, hand-signed letters, no labels, commemorative stamps, hand-written reminders) as opposed to a "regular approach" (computerized salutation, Xeroxed signatures, printed address labels, bulk-mail) (K. Smith & Bers, 1987). Moreover, the pleasantness of the survey experience was also found to be related to the probability of cooperating with future survey requests (Nederhof, 1987).

The problem is that personalization (or lack thereof) and pleasantness are also salient to the alumni experience. A lack of personalization may very well bring forth in the minds of alumni what they consider one of the most infuriating aspects of their university/college experience: the bureaucracy (Brenders, Hope, & Ninnan, 1999). When presented with bland bureaucratically- inspired cover letters, those who have been stonewalled by the university bureaucracy may also be the ones who decline to answer a survey request. In an ironic twist, survey administrators seem to share this frustration precisely because of the reported difficulty of transcending "legalistic cover letters" imposed by the organization (J. Grayson & Myles, 2005).

2.3.4 Influences of Survey Interaction

Previous sections on survey designs have dealt with the influence of the mode, the protocol, the administration and the questionnaire on topic saliency. This section deals specifically with the interaction of the sampled individual with the questionnaire in terms of measurement error. As previously discussed in section 2.1.3, measurement error is related to the detection of nonignorability. Two of such errors are discussed in this section: response errors as a function of survey cooperation and response errors as dilution. The first imitates

the effects nonignorable nonresponse on the variable of interest by introducing response bias as a function of survey cooperation. The greater the response rate, the greater will be the biasing effect on the variable of interest. The bias is the expression of response bias that is to say when there is systematic over or underreporting of one's otherwise true answer. The second source of error dilutes the detection of nonignorable nonresponse by introducing large amounts of randomness in the variable of interest. The added variance is the result of response variance, that is to say when the respondent deviates randomly from his or her true answer. The implications for nonignorability are discussed.

2.3.4.1 Measurement Error: Response Bias

Response bias has been observed on Likert scale based questions in satisfaction surveys. During a think-aloud interview about a student satisfaction survey conducted on twenty-four medical students, two important findings were observed. First, students gravitated towards the second highest category (satisfied) when uncertain about the meaning of some questions or when uncertain as to how to map their evaluations onto a five-point scale. Moreover, students were generally biased towards the positive end of the scale; the highest level was reserved for "unequivocal" satisfaction, while the second highest was an amalgam of reasons "not all of which seemed to indicate a positive evaluation" (Billings-Gagliardi, Barrett, & Mazor, 2004: 38). Second, students edited their evaluations according to how their chosen value on the satisfaction scale could be evaluated by others. These findings indicate clearly that social desirability, in this case the desire not to appear negative, exerts a strong gravitational force towards the second highest point of the scale. In fact, being "satisfied" might be construed by these respondents as the middle point of the scale. The

consequence of this measurement error is to reduce the natural variance in satisfaction answers towards the satisfied end of the scale.

A second type of response bias can be caused by a heuristic cognitive process called anchoring. Anchoring generally occurs when one bases one's judgement on one's initial assessment of the situation. One such example is the halo effect. Not unlike the response bias on Likert scales, the halo effect reduces the natural variance of satisfaction scores, and assimilates various facets of satisfaction (Wirtz, 2003). This particular type of bias occurs mainly when the respondent is asked to perform specific assessments of satisfaction but recalls only summatives or generalized assessments, or tends to magnify one experiential aspects over all others, draws from contextual cues in the order of similar questions, or is unwilling or unable to discriminate between various aspects and makes use instead of inferred similarities between experiential aspects (Wirtz, 2003: 99).

Other anchoring errors, equally damaging to natural variance, are ceiling and floor effects and attribution errors. Ceiling and floor effects occur when satisfaction judgements are made under extremely high or extremely low expectations in relation to actual performance (Oliver, 1997: 110); disconfirmations of expectations, under the expectancy disconfirmation model, are rendered inoperative due to the high levels of expectations. Attribution errors occur during the assessment of performance of individuals, the consequence of which is to "underestimate the extent to which behaviour is shaped by the constraints of the situation and overestimate the extent to which it is shaped by people's underlying dispositions" (Kunda, 1999: 429). The attribution error may lead to a correspondence bias whereby underlying dispositions are drawn from the observation of a

small subset of behaviours; the net result is a failure "to appreciate the extent to which situational forces had contributed to that behaviour" (Kunda, 1999: 430).

These response biases are consequential for nonignorability only if the measurement error introduced is correlated with an increase in response rate. While such biases are to be expected from some of the respondents, it is likely to be more prominent amongst those who found the topic less salient. Low salience topics are likely to produce, on the part of respondents, heuristically-driven, short-hand answers that will either gravitate towards the satisfied end of the scale, or anchored on some specific aspect of the questionnaire or their alumni experience. Since low salience is inversely related to survey cooperation, the incorporation of low-salience recalcitrant alumni in the alumni pool will introduce more measurement error in the variable of interest. In particular, if these heuristically-driven answers do gravitate towards satisfaction, then an increase in response rate will bring into the response pool an overwhelmingly large number of satisfied individuals. Depending on the initial response marginals, the bias will be positive if the initial respondents were more negative, or the bias will be negative if initial respondents were mostly very satisfied. In either case, the appearance of a nonignorable relationship between response rate and satisfaction is not caused by satisfaction per se (e.g. being irate or elated), but rather from the introduction of measurement error as a function of response rate.

2.3.4.2 Measurement Error: Response Variance

Response variance pertains to all sampled individuals whether they deem the topic salient or not. The issue with response variance is the introduction of unreliability in the

measurement of satisfaction. This added noise drowns out any real relations between the values of the survey variables of interest and survey cooperation; and consequently, obscures the capacity of detecting nonignorability measured as a change of satisfaction scores for a change in response rate. This situation occurs, as we have already mentioned in an earlier section, on whether the questionnaire was able to formulate its request in an intelligible manner (Foddy, 1993). When skip patterns are difficult to interpret, questions are not given out-options such as "no opinion" or "don't know", and words are difficult to interpret (ex: bureaucratese), the respondent may opt for a variety of strategies to cope with the uncertainty. When forced to produce an opinion, they may fall prey to formulation of quick off-the cuff opinions or non-attitudes (Converse, 1970). They may use contextual clues such as questions already answered, questions above or below, or the answers provided by the question itself to formulate an answer (Foddy, 1993; Schuman & Presser, 1981; Tourangeau, 1999). They may also forgo devoting the cognitive energy required to understand the request, and opt for various "satisficing" heuristics such as selecting the first reasonable response answer, mental coin flipping, agreeing with assertions, or simply acquiesce with the status quo (Krosnick, 1991). All these strategies will result in increased levels of randomness in survey variables, and may drown, as a result, the capacity to detect any real patterns of nonignorability.

Response variance may also be introduced, not by convoluted questions per se, but by the generality of questions. The latter are typical questions on evaluative-type satisfaction surveys whereby the respondents are asked to assess, in general or in retrospect, their satisfaction with a particular aspect of their program. The problem with such questions is the

uncertainty as to which specific element the respondent will actually use to answer the question. Will they calculate a satisfaction score for each relevant aspect of their program, and sum together to arrive at a general assessment; or will they anchor their assessments on one particular aspect of the question. The issue is not simply one of recalling all aspects pertinent to the questions; it is also a matter of arriving at a congruent answer amidst conflicting satisfaction scores. It may not be possible, even for the most salient of topics, to arrive at a reliable, stable answer when the generality of the question conjures up opposing, conflicting and ambivalent satisfaction levels (Zaller & Feldman, 1992: 584-585). As result, general questions, if not properly worded, defined and contextualized are likely to generate a great deal of randomness.

2.4 Decision to Cooperate

Previous sections have enumerated various determinants of topic saliency that elaborates why one's graduate satisfaction impinges on the decision to cooperate with a survey request. These determinants, however useful, do not provide on their own a comprehensive understanding as to how students base their decision to cooperate with a survey request on their university experience. Without a theory that links these determinants together, survey cooperation could easily be construed as a behaviourist (stimulus-response) response to these determinants, or as a matter of pure choice on the part of the sampled individual. Goyder (1987) has criticized at length the not-so-subtle behaviouristic assumptions often presupposed by survey researchers in relation to survey cooperation. Indeed, setting aside both behaviourism and voluntarism, survey cooperation needs to be explained, according to the author, in part by determinants, and in part by the individual's

own reasoning. Social exchange was the outcome of this approach to survey cooperation. Since then, however, theoretical debates have gone beyond the dichotomy between behaviourism-voluntarism, and broached the place of rationality and cultural norms in the decision to participate in a survey. A new theory, leverage-saliency, is emerging out of this debate.

The debate is still active but may no longer reside, as some authors have concluded, between heuristics and rationality. It does not completely reside, as De Vaus (2002) suggests, between the survey characteristics approach and the societal approach. It would be correct to suggest, we would argue, that the debate is primarily about the range of validity of any given theory, and ultimately about the place of theory for survey practice and for social life more generally (Goyder et al., 2006). Unfortunately, topic saliency only serves to highlight this debate, particularly when the survey request operates within the confines of an on-going organizational relationship. We will begin by reviewing the contribution of leverage-saliency to topic saliency bias, to be followed by an expanded look at social exchange within organizational-client relationships.

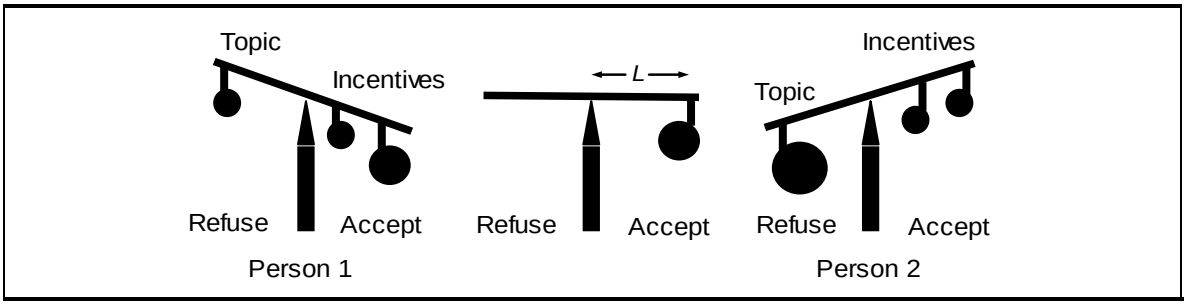
2.4.1 Leverage-Saliency Theory

Leverage-saliency theory is an on-going theoretical articulation that attempts to specify from the ground up how known determinants of survey cooperation (social, psychological and survey design) influence the decision to cooperate to a survey request[27].

[27] It is our understanding, from a review of Groves' works that leverage-saliency theory is the culmination of an on-going intellectual process about survey cooperation. It is easy to detect the frustration of the author and the anticipation to establish a "useful theory of survey participation" (Groves, 1991). The concepts developed in that conference paper were reiterated in the article "understanding the decision to participate in a survey"

Leverage-saliency conceives survey cooperation primarily in terms of heuristics. Because, as Groves and Couper correctly suggests, individuals "do not have a large personal interest in survey participation", sampled individuals are unlikely to devote the cognitive effort to the task; they are likely to use short hand, rule of thumb decision-making strategies (Groves & Couper, 1998). The decision to cooperate to a survey request may rest on the availability or anchoring heuristics particularly when individuals focus only specific aspects of the survey design; they may also base their decision on the representational heuristics where survey features are assessed according to their similarity to other surveys (e.g. mistaking governmental for commercial surveys)[28]. Thus the degree to which a given sampled individual devote the cognitive effort depends, in the final analysis, whether some survey features are noticed or not, and the intensity or importance one gives to that particular feature. It is these two elements, the salience and the leverage that explains the "threshold nature of the survey response propensity" (Groves et al., 2000: 300).

Figure 2-10 Leverage-Saliency Cooperation Model



Sources: (Groves, Fowler et al., 2004: 177; Groves et al., 2000: 300)

(Groves, Cialdini, & Couper, 1992), and further developed in (Groves & Couper, 1998). The singular focus on heuristics gave way to a more encompassing conceptual foundation based on probabilistic cognitive theory. Leverage-Saliency theory (Groves et al., 2000) is the product of this conceptual shift in emphasis. Further developments of the theory broached the concept of topic interest more fully (Groves, Presser et al., 2004) using the (Petty, Richard, & John T. Cacioppo, 1986; R. E. Petty, Cacioppo, & Schumann, 1983) dual track, central and peripheral routes to persuasion.

[28] For a discussion of various types of heuristics, see (Kunda, 1999).

The strength of leverage-saliency is to make sense, at the individual level, how a variety of survey design elements, social and cultural aspects comes to play a role in the decision to cooperate to a survey request. Figure 2-10 illustrates the basic mechanism. Dependent upon the interaction of the interviewer or the survey material with the sampled individual, some features of the survey will figure more prominently than others. The respondent may single out the topic itself and/or peripheral elements such as cash incentives. The more an individual zeroes-in on a particular aspect of the survey, the larger the size of the weight on the fulcrum. Saliency, in leverage-saliency, is analogous to a "top of mind phenomena" (Taylor & Fiske, 1978). The leverage, represented by the distance (L) from the fulcrum, reflects the importance of the salient object to the decision. The importance of a salient object is influenced by one's "background and experiences in ways unknown to the interviewer" (Groves et al., 2000: 301). Individuals, dependent on the saliency of a given survey feature, and dependent on the leverage given to each salient object, will accept or refuse to cooperate to a survey request. This decision-making process is probabilistic to the extent that no one survey request is exactly the same, or will be given the same cognitive attention.

In relation to topic saliency, the theory would postulate that 1.) interested individuals will answer in greater numbers, 2.) Overrepresentation is a function of the saliency of the topic relative to other salient features of the survey, 3.) topic saliency effects will dominate the decision to respond in the absence of other salient features. We would add, by extension, that topic saliency effects may prevail when other salient objects are given greater leverage towards the decision to refuse (i.e. cognitive burden). Three applications of the theory were

conducted to assess the relative impact of survey design attributes. As predicted by the theory, when community involvement is salient to the sampled individual, incentives, however salient, weights in less in the decision to cooperate, than for those individuals who did not find community involvement salient (Groves et al., 2000). In relation to topic interest, people who found the topic of interest to them did respond in greater numbers; a monetary incentive did increase cooperation amongst those who did not find the topic salient; however, in part due to methodological issues associated with the groupings of individuals with unknown topic leverages, only weak support could be extracted from the data regarding nonresponse bias (Groves, Presser et al., 2004). A subsequent experimental design succeeded in producing nonresponse bias due to topic saliency (birding experiment) (Groves, 2006a).

The major limitation of the current formulation of this theory is its preference for a strong methodological individualism[29]. The decision to cooperate with a survey request is essentially reduced to the cognitive processes of a given sampled individual. The criticism is not aimed to dismiss the notion that socio-demographics, social or cultural norms cannot be explained by the actions of individuals. We would agree in principle that such contextual factors are not directly causal to the participation decision. Instead, they produce a set of "psychological predispositions that affect the decision" (Groves & Couper, 1998: 32). However, the radical reduction of these factors to the cognitive property of individuals masks the overarching reach of these factors on the decision to cooperate with a survey request – factors, we must add, that were already well articulated in earlier models of survey

[29] Methodological individualism – for or against – is an intellectual pasttime that has enjoyed a rather lengthy pedigree (for reviews see (Hodgson, 1986; Laurent, 1994; Udehn, 2002)). Our intention is not to overturn its usage in the survey literature. Our issue is with the all-too-common tendency towards reductionism, still eloquently expressed in (Morgenbesser, 1964)

cooperation (e.g. (Groves & Couper, 1998)) but not found in leverage-saliency . It is likely under specific settings, such as organizational ones, that survey design features, their respective salience and leverages, might not be mutually independent predictors that sum into a positive or negative decision cooperate with of survey request. It is likely that these factors determine not just the leverage (e.g. their importance) of survey features but also which are to be considered salient, and which are to be ignored.

2.4.2 Social Exchange Perspectives

The social exchange perspective is a collection of theories of social action that attempts to bridge together economic, behaviourist, and cultural/anthropological conceptions of society and of human nature. Some authors will emphasize the rational and utilitarian strands of social exchange, others conceptualize human behaviour in terms of stimulus-response, and other theorists contextualize action within a cultural framework. All theorists would agree, however, that human behaviour within a given society is characterized by exchange. Much more than economic goods and services, exchange refers to the totality of resources, symbolic and material, between individuals. These reciprocal exchanges confer on recipients a set of obligations to reciprocate in kind, whether through rational calculus or cultural expectations. The strength of these obligations will depend partly on whether there is an on-going relationship reinforcing (making salient) reciprocity rules, and/or whether these normative rules are made clear at the level of strategies of exchange (Befu, 1980; Goyder et al., 2006).

Social exchange concepts, applied to survey cooperation, are distilled to three essential components. Following Dillman's approach, survey cooperation is mediated by assessments of costs, rewards and trust. Costs are all elements that the potential respondent must relinquish or endure in order to receive any promised rewards that may come out of the exchange; rewards are the gains, symbolic or material, one expects to receive during or as the outcome of the exchange; trust is the expectation that rewards or benefit accrued in completing the exchange will outweigh any immediate costs (Dillman, 2000: 14). In the vocabulary of survey design, ways of ensuring a successful exchange includes the minimization of costs such personal embarrassment, burdenful or intrusive questionnaires; the maximization of rewards by giving explicit rewards such as monetary incentives or assigning symbolic rewards such as social validation; the establishment of trust through reassurances that the information will be of use or through a legitimate authority (Dillman, 2000: 15-20). It is presumed that the transfer of rewards and promulgation of trust will not only outweigh the cost of the exchange, but will also create an obligation to reciprocate by exchanging personal information.

In relation to saliency, social exchange endeavours to create what Dillman calls "positive salience" (2000: 155). This strategy applies to topic salience as well as the more peripheral aspects of the survey. The design of each survey feature is carefully crafted to bring forth in the minds of sampled individuals maximum rewards, minimum costs, and a conviction of the trustworthiness of the requesting organization. Topic salience, another survey feature, must also be designed such that the topic be widely shared, be of interest, be easily comprehensible, and most important, be seen as an attempt to connect with sampled

individuals (Dillman, 2000: 92-93). In all cases, survey features are to be understood as "part of an overall implementation system" and not "self-standing" (Dillman, 2000: 156). Thus, the concept of saliency within social exchange is less about the objects of salience (e.g. survey features) or their actual leverage (e.g. relative importance); it is the basic thrust in an overarching strategy aimed at highlighting the exchange itself, to render salient the exchange between the requesting organization and sampled individuals.

The limitations of this particular application of social exchange to survey cooperation lie in its focus on the dyadic or restricted forms of exchange. Indeed, it is no surprise that De Vaus readily categorized Dillman and Groves under the same heading of "survey characteristics explanation" (De Vaus, 2002: 29). In focusing too much on the interaction between survey sponsor/researcher and the recipient of the survey, this interpretation of social exchange has relied too extensively upon the "strategies of exchange" at the expense of the cultural frame of reference that "sets the background which the norm, rules and strategies operate" (Befu, 1980). In limiting social exchange to the strategies of exchange, Dillman's imagery of the balance between "cost-benefit" analyses can easily fuse, in the minds of some survey practioners, into the fulcrum imagery of the probabilistic "leverage" theory. Much of the confusion surrounding the place of rationality (heuristics versus calculative), and of cultural norms, stem from this reductive understanding of social exchange[30].

[30] To be fair, it should be noted that Dillman's focus on the strategies of exchange is a reasonable approach when dealing with broad general population groups provided that the norm of reciprocity is still operative (e.g. salient) in the sampled individual's mind. To be operative, the norm of reciprocity should guide the saliency of "weights" and the assessment of "leverages", to use leverage-saliency theory terminology. If the norm is too

It is our contention that survey cooperation understood, as social exchange, must also consider the cultural frame of reference when assessing ignorability to organizational surveys. This framework is not part of the exchange per se, in that it "refers to factors exogenous to an exchange system, factors which are not recognized as variables affecting the outcome of exchange transactions in a given exchange system" (Befu, 1980: 205). One of such frameworks can be found within our relationships with organizations; these relationships alter the very nature of the exchange (Cropanzano & Mitchell, 2005). Thus, what we deem salient not only in terms of the survey topic but also in terms of specific aspects of survey designs, is based on our on-going relationship with the requesting organization. Three theories will clarify how organizational experience relates to survey cooperation: organizational citizenship behaviour, trust and consumer complaint.

2.4.2.1 Organizational Citizenship Behaviour

Organizational Citizenship Behaviour, or OCB, is a concept initially developed to describe a particular type of employee-employer organizational relationship. It is defined as the "behaviour that is discretionary, not directly or explicitly recognized by the formal reward system, and in the aggregate promotes the effective functioning of the organization … the behaviour is not an enforceable requirement of the role or the job description… the behaviour is a matter of personal choice" (Ogdan in Coyle-Shapiro, Kessler, & Purcell, 2004). Such behaviours exhibit traits of helping, sportsmanship, loyalty, compliance, individual initiative, civic virtues, and self-development (Podsakoff, Mackenzie, Paine, &

---

weak, thus inoperative, we would contend that dyadic social exchange can be absorbed in the leverage-saliency theory framework.

Bachrach, 2000). The outcome of OCB is an expansion of normal role expectancies to include "extra-role" behaviours. Moreover, there is a transformation of "in-role" behaviours to the extent the individual now integrates proactive values and norms of the organization as his/her own; it becomes part of one's job. Both of these behaviours stem from a desire to reciprocate for what is perceived to be "fair and good treatment" (Coyle-Shapiro et al., 2004). In essence, OCB is a transmutation of a restricted exchange between employer-employee towards a more generalized exchange between employee and organization – generalized exchange here understood as indirect exchange with other members of the organization and performed without expectations of immediate or direct reciprocation (Takahashi, 2000)

The concept of OCB was not lost on survey researchers in their attempt to understand cooperation with organizational surveys. In a study of 194 employees non-randomly sampled from different organizations, organizational citizenship variables such as organizational satisfaction and commitment were tested against intentions to participate in a survey about their organization. Anticipated respondents differed from anticipated nonrespondents in relation to their attitude towards their organization commitment, satisfaction with supervisor and with their job as a whole, and intention to resign. Overall, a positive bias was observed in the response pool on variables pertaining to job, management and the organization (Rogelberg et al., 2000).

The problem with studies that utilize OCB as the basis to understand survey cooperation lies in the fallacious characterization of all non-respondents as lacking in OCB. In other words, OCB may be an impetus to respond; those lacking in OCB may be less likely

to respond, but nonrespondents are not necessary lacking in OCB. Thus, studies such as (Spitzmuller, Glenn, Barr, Rogelberg, & Daniel, 2006), who conclude that those lacking in OCB were more likely to report unfair resource allocation, insufficient organizational support, and generally lower social exchange with their organization, should be understood in relation to the respondents who exhibited OCB characteristics, and not as a description of non-respondents per say[31]. As we shall see in the next section, nonresponse need not be construed in opposition to OCB organizational traits.

2.4.2.2 Trust and Functional Authority

Organizational relationships do not always transmute themselves into citizenship behaviour. In fact, very few do. This is not to say that successful exchange does not take place. Most of our social exchanges with collective organizations are largely satisfactory, and occasionally, they can be downright frustrating. The contribution of trust theories to our understanding of organizational relationships lies in the functional outcomes of trust in collective organizations. As users of organizational services, long term or short term, we are rarely invited, nor are we likely, to participate in the day-to-day decision-making of an organization. In the language of trust theories, we are relinquishing our judgement to public administrators. Public officials are entrusted to administer institutions effectively. They are given "authority of function". As Warren explains:

> We trust airline controllers, food inspectors, and the judicial system, and we do not feel any particular compulsion to

[31] It is important to note that respondents and nonrespondents may each have their own rationale for not cooperation with a survey request. To suggest that nonrespondent lack OCB characteristics should not be seen as *the* rationale for not cooperating. Nonrespondents may see their interests lying elsewhere than that of cooperating with managerial imperatives for example.

> substitute our judgement for theirs simply because those who guide planes, inspect food, or handle criminals and conflicts do so according to appropriate standards of expertise and appropriate procedures (Warren, 1996: 49)

It is in this sense of functional trust that one has to understand the exchange between clients, citizens or in our case students, and a given organization. Trust can be defined as "a state of favourable expectation regarding other people's actions and intentions" (Mollering, 2001). It is different from confidence in that one maintains expectations of satisfactory outcomes despite the awareness of potential betrayal. These favourable expectations are anticipatory in that they reduce the complexity of our dealings with others (Barber, 1983; D. Lewis & Weigert, 1985; Luhmann, 1979). In an organizational setting, trust operates as a form of social capital which facilitates cooperation between strangers, increases the likelihood of organizational citizenship behaviour and increases compliance to hierarchical authority (Kramer, 1999). Trust, in effect, reduces the salience of organizational relationships. Exchange, under such conditions, assumes that reciprocity will be returned, and that reciprocity will follow certain rules of conduct (Blau, 1967).

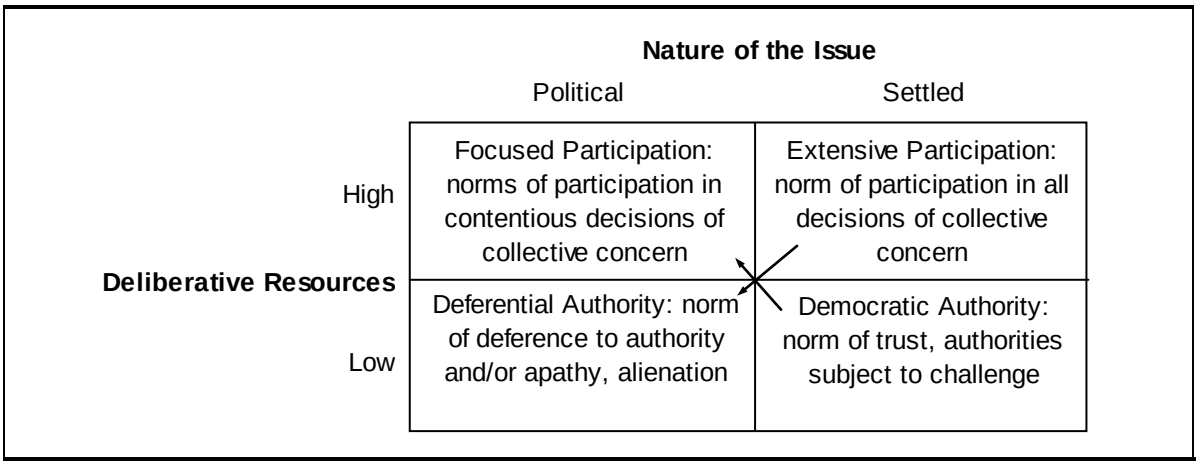
However, when procedures are followed arbitrarily or when individuals are treated badly, organizational relationships or the quality of services provided becomes highly salient. Distrust, here defined as a "lack of confidence in the other, a concern that the other may act so as to harm one, that he does not care about one's welfare or intends to act harmfully, or is hostile" (Kramer, 1999: 587), elicits a heightened sense of alertness. Intensive cognitive processing is spent uncovering hidden motives, to "actively entertain multiple, possibly rival, hypotheses about the motives or genuineness of a person's behaviour" (Kramer, 1999: 587).

The impetus is to reassert one's capacity for judgement, to question the authority of function. In a sense, the desire to get involved with the internal workings of an organization operates at the "margins of trust" (Warren, 1996: 49):

> While we may not wish to participate most of the time, we want procedures that allow us to do so when authority becomes questionable, and this occurs when authorities make decisions no longer functionally specific to the goods they serve. When questions of profitability compromise the safety of food or airline travel, when priests use their power over salvation to gain sexual access to bodies, when office managers use their power of organization to gain personal favors or excessive rewards, when elected officials sell their influence for money, at points such as these "democracy" is important and often demanded. (Warren, 1996: 49)

There is no claim made that survey respondents are seeing satisfaction surveys as a deliberative, political or even a democratic act. Student satisfaction surveys are certainly not as lofty a practice as deliberative democracy. They are after all a limited, passive and altogether benign form of student engagement in the decisions of their institutions or organizations; it is a way to voice one's opinions or concerns in the hopes of a more responsive organization; no more, no less. Moreover, the cognitive demands placed on the respondent are nothing compared to full-fledged democratic practices. Nevertheless, there are strong assumptions shared by satisfaction surveys and deliberative democrats on the subject of functional authority. By eliciting client/customer/citizen voices in management decisions, organizations are in effect violating the functional aspect of authority. To ask "how are we doing today?" or "how was it for you?" is simply not salient unless one's trust has been misplaced.

Figure 2-11 Deliberative and Authoritative Decision-Making



Source: (Warren, 1996)

An adaptation of Warren's (1996) model on functional trust within a democracy, illustrated in Figure 2-11 above, provides some clues as to how survey cooperation might be understood in an organizational setting. The rows indicate the amount of resources available to a given individual such as "time", "expertise", "attentiveness" and "psychological security" listed as essential elements needed for participation (Warren, 1996: 57). They easily apply to what is known about survey participation in terms of student and contextual characteristics. The columns reflect the nature of the issue in which one is asked to participate. Political issues are "relatively contentious" where there is "sharp disagreement about public concerns"; settled issues are "non-controversial" and usually "draw few challengers" (Warren, 1996: 57). Offering high levels of deliberative resources and ready to voice opinions on settled issues are the organizationally committed individuals. Organizational trust is an outcome of the authority of function; these individuals are not motivated to expend deliberative resources on settled issues. However, if issues become political, organizational trust gives way to focused participation. More problematic, from an

organization point of view, are those unwilling to expend deliberative resources despite recognized political issues. A part of this group is the once OCB motivated individuals, now highly disaffected. Such a group are likely to have experienced violations of role-expectations in terms of expected obligations and competence associated with a given role, and/or violations of rule-expectations in terms in a shared adherence to a normative system within the organization (Kramer, 1999: 578-579).

The value of this model for nonignorable nonresponse is to show that satisfaction does not always lead to organizational citizenship behaviour, and by extension reciprocation with the organization. In fact, a satisfactory exchange will sustain trust relationships with public administrators, and with the manner in which services are dispensed; exchanges will not transmute themselves into a more generalized form of exchange. Only when the exchange has been less than satisfactory, will there be some reciprocation in order to, presumably, re-establish trust relationships. Thus, functional trust theory specifies when and why satisfaction levels are cause for reciprocity. The saliency of the topic is not due to satisfaction per se, but to the organizational relationship now seen to be in jeopardy.

2.4.2.3 Consumer Complaint Behaviour and Loyalty

From a social exchange perspective, one could reasonably deduce that being dissatisfied with one's organizational experience would provide little impetus for reciprocation and would weigh heavily on the costs side of the exchange assessment. In fact, the disclosure of one's dissatisfaction would likely require some sort of "payment" in return to offset that cost. However, when survey response conceived as exchange is influenced by

the on-going relationship with the organization, this deduction is not entirely complete. Indeed, when the relationship is minimal, exchange is likely to be based less on cultural orientation and more on one's own assessment of the situation. The literature on consumer complaint behaviour conceptualizes this exchange in terms of maximizing rational calculus but with a focus on redress success. When the on-going relationship is strong, akin to OCB, dissatisfaction is an expression of loyalty, that is to say a desire to maintain one's relationship with a faulty organization; exchange requires no incentive or intense calculus. We will begin with consumer complaint studies, to be followed by loyalty.

Consumer complaint studies focus on the complaint behaviour associated with a negative service experience and/or a defective product. A common trend among these studies is the realization that dissatisfaction does not lead to complaint behaviour, but rather to avoidance expressed through "private complaining" (Gwinner & Stephens, 1998; Oliver, 1997; Singh, 1988, 1990; Volkov, Harker, & Harker, 2002). (Singh, 1990) found that 51% of irate customers would not complain to the manufacturer; instead, they are more likely to engage in word-of-mouth behaviour; the injured party will convey his or her displeasure regarding the offending organization to peers and family. Moreover, (Nyer & Gopinath, 2005) found that those who choose to engage in word of mouth behaviour (gossip) rather than to complain to the offending organization, will remain committed to their dissatisfaction level. Word of mouth behaviour is perceived as the least costly option: "they simply never return to the firm with whom they are dissatisfied […] they typically feel that complaining is 'not worth their efforts'" (Stephens & Gwinner, 1998: 184).

Oliver (1997) cites two explanations of complaint behaviour. The first highlights the cultural environment and the psychological resources presupposed before engaging in complaint behaviour. The act of complaining presupposes an organizational culture that either invites or at least shows a willingness to receive such complaints. The motivation to register a complaint quickly vanishes if it is belittled, ridiculed, perceived as an act of lèse majesté, or worse labelled as originating from a "chronic complainer" (Oliver, 1997: 363). Moreover, individuals must have at their disposal psychological resources that will allow them to confront the cultural or social undesirability of complaining such as self-esteem, communication skills and prior knowledge of complaint mechanisms (Oliver, 1997: 363).

The second, closer to the rational variant of the social exchange perspective, cites the costs and benefits associated with the decision to complain to the offending organization (see Table 2-2 below). Costs such as monetary loss, time and effort required to build a complaint argument are evaluated against the potential benefits of having one's product reimbursed or receiving an apology for surly service. The moderating factor is the probability of succeeding in acquiring such benefits; probabilities are measured against previous complaint experiences, the potential pressure that can be applied by the consumer and the organization's reputation in handling complaints. The probability of success is projected to be low when the consumer expects a fight against the "complaint escalation pyramid" (Heskett, Sasser, & Schlesinger, 1997); customers, willingly or not, will find themselves railing against escalating levels of management starting with the sales clerk, then the middle manager, right up to the vice president. At each stage, the cost of sustaining one's complaint

behaviour is ever more onerous. Understandably, under these conditions, fewer and fewer customers bother registering their complaints beyond sales clerks (Heskett et al., 1997).

Table 2-2 Cost-Benefit Decision to Complain

Costs	Benefits	Probability of Success
Monetary loss	Reimbursement	Response reputation
Ancillary loss	Replacement	Threat to business
Time required to complain	Extra compensation	Previous experience/efficacy
Effort required to complain	Correction of problem	
Product importance	Apology	
	Vent frustration	

Source: (Oliver, 1997: 361-362)

Customers, however, do complain. At times, their motivation for complaining seems to run counter to their economic self-interest. Clearly, if the product is faulty or the service is unsatisfactory, there is no reason to pursue a relationship with that organization; the rational consumer might seek a better organization/company instead of shoring up a faulty one. In other words, they should avail themselves of the "exit option" instead of seeking to "voice"[32] their disagreement (Hirschman, 1970). Underlying the act of registering one's complaint is a desire to pursue or to maintain an on-going relationship with the given organization. Complaint behaviour is related to the concept of loyalty. Following Hirschman (1970), loyalty has two fundamental characteristics: firstly, the extent to which customer-members are willing to trade off the certainty of exit against the uncertainties of an improvement in the

[32] Hirschman defines voice as "any attempt at all to change, rather than to escape from, an objectionable state of affairs, whether through individual or collective petition to the management directly in charge, through appeal to a higher authority with the intention of forcing a change in management, or through various types of actions and protests, including those that are meant to mobilize public opinion" (Hirschman, 1970: 30)

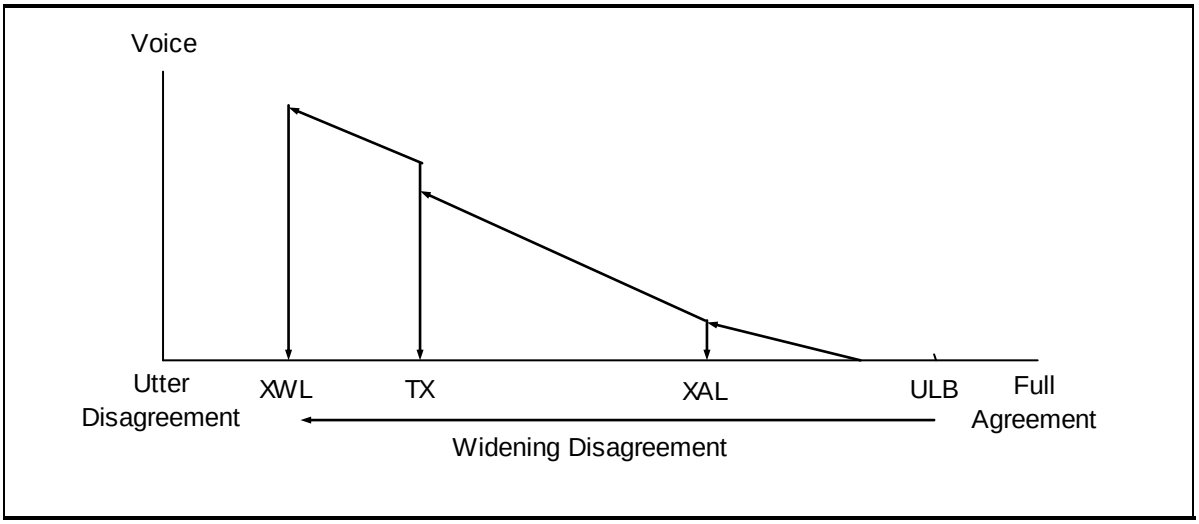
deteriorated product; and secondly, the estimate customer-members have of their ability to influence the organization (Hirschman, 1970: 77).

The second characteristic is an extension of the probability of success already discussed above, and tabulated in Table 2-2; the consumer evaluates success in terms of influence on the organization. The first element, however, implies a leap of faith; when the organization provides lesser products or services, a loyal customer will not only ride out the shortcomings (instead of seeking immediate exit) but will also find a way to be influential in order to get the organization or company "back on the track" (Hirschman, 1970: 78). A loyal customer will likely take the opportunity to voice an opinion. Consequently, complaint behaviour should not immediately be confounded with an overly critical chronic complainer, but rather it should be seen as a particular kind of customer, very different from the ones who have chosen the exit option.

Figure 2-12 illustrates how one's disagreement with organizational policies impacts the decision to formally complain to the organization (voice) or to engage in avoidance behaviour (exit). Changes in policies or organizational performance may not be immediately perceived by the customer; he or she might still be within a zone of indifference (ULB). Disagreement, now fully cognizant, will lead to the probability of voicing one's concern but will very quickly lead to exit behaviour for individuals not loyal to the organization (XAL). Continued disagreement will lead to threats of exit (TX) from loyal customers if the situation is not corrected. A much higher (and louder) probability of voicing one's concerns ensues when even higher disagreements persist -- at which point, even the most loyal of customers will also defect from the organization (XWL). Loyalty, in the context of cost-benefit analysis

is “paradoxical” to the extent that it promotes exchange with the offending organization when clearly it is not deserved; an exiting and engaging in word of mouth behaviour strategy would have been expected (Hirschman, 1970: 81).

Figure 2-12 Hirshman’s Loyalty Theory



Source: (Hirschman, 1970: 87)

The contribution of consumer complaint studies to nonignorable nonresponse to student satisfaction surveys is threefold. First, it seems likely that a large percentage of dissatisfied students will not voice their concerns directly through a student satisfaction survey. The likelihood is avoidance and word of mouth behaviour. Second, those who do complain will be those who are, or were, loyal to the organization. During the course of their degree program, loyal students, who for the most part have been central to the norms of the organization, may have become more and more disaffected with some aspects of their degree program. These students are likely to voice their concern unless the disagreement has reached a breaking point. Third, as an extension to both points above, respondents (satisfied and dissatisfied) are likely to be quite different from nonrespondents; They will be more loyal,

emotionally attached and were well integrated with the university system. Some support for this hypothesis is provided by (Thorsten, Langer, & Hansen, 2001)[33]

2.5 Conceptual Hypotheses

In this section, we will begin by summarizing the theoretical propositions and empirical evidence established thus far. The goal is to deduce the relationship between nonresponse and nonresponse error as it pertains to satisfaction-based cooperation behaviour. This will be followed by a conceptual model that will attempt to elaborate the hypothesized nonignorable nonresponse.

2.5.1 Nonresponse and Nonresponse Error

Following organizational citizenship behaviour and loyalty theories, the propensity to respond to a satisfaction survey will be higher among those most involved in the organization. They are more likely to express both their satisfaction (OCB) and dissatisfaction (loyalty) with the organization. It is hypothesized, in light of their sustained involvement, that they are more likely to produce a satisfaction bias. This is corroborated by Rogelberg et al (2000) empirical study of employee satisfaction surveys. This evidence is also supported by health studies on patient satisfaction, which show a consistent positive bias amongst in-patients as opposed to outpatients. It appears that being part of an organization for an extended duration of time, either as an employee or as an in-patient, is a strong motivator to respond, and in particular for those who had a satisfactory experience.

[33] Further support is also derived from Etzioni's model of organizational involvement within the context of employee loyalty. See (Jernigan & Beggs, 2005; Penley & Gould, 1988)

However, dissatisfied individuals will not only be composed of loyal involved students, but also those who experienced a discrepancy in their program. Following Warren's theory, the silent otherwise satisfied majority would give way to cooperation behaviour if a discrepancy were salient enough to be expressed. This compounded dissatisfied group may tip the balance towards a dissatisfied bias at lower response rates. An increase in response rate would eventually incorporate the silent satisfied majority (zone of indifference) into the response pool, and, as a result, bring the distribution of responses closer to the sampled population. It is important to note that the initial distribution of satisfaction scores will be truncated at the satisfied level giving way to a flatter distribution or even a bi-modal one as shown in Figure 2-13a. As the response rate increases, the distribution should peak at the satisfaction score (see Figure 2-13b). It is hypothesized that the net effect of discrepancies and topic involvement on satisfaction scores should be smaller as the response rate increases.

Figure 2-13 Hypothesized Effect of Response Rate on Satisfaction Scores

However, the existence of a core group of dissatisfied individuals who are much less likely to cooperate than any other experiential group may bias the effect of discrepancies on satisfaction scores yet again (see Figure 2-13c). The consumer complaint literature makes it

clear that dissatisfaction does not always lead to participation in a complaint mechanism, whether it is through a survey or through formal bureaucratic procedures. These individuals are overwhelmingly dissatisfied with their experience and their refusal to participate in a satisfaction survey would stem from the perceived futility of seeking redress. Following Warren's trust theory[34], such individuals are more likely to have experienced severe breaches in role-based trust and/or rule-based trust (Kramer, 1999); they are likely to feel alienated and/or be ostracized from the organization. In other words, it is hypothesized that their dissatisfaction, and their muted cooperation with a satisfaction survey, both stem from the evaporation of trust in the functioning of the organization itself. It is unclear how large such a group may be in relation to the sampled population. Unless we can infer that dissatisfied respondents are similar to dissatisfied nonrespondents, the size of the satisfaction bias remains unknown. Depending upon the effectiveness of bureaucratic policies to prevent these breakdowns, this experiential group should be relatively small. Nevertheless, there will always remain in the response pool a small absolute bias towards satisfaction until all sampled individuals are reached, and have cooperated.

2.5.2 Satisfaction-Based Survey Cooperation Model

In order to assess whether nonresponse error is ignorable or nonignorable, a conceptual model based on Michalos' Multiple Discrepancy Theory (see Figure 2-14 below), herein named MDT, will be used to specify the relationship between satisfaction and survey cooperation (action). The model is particularly useful in assembling the determinants of topic saliency enumerated above, but its main contribution is to specify classes of expectations in

[34] See Figure 2-11 lower left quadrant on disaffection and alienation.

the expectancy-disconfirmation model of satisfaction. Instead of relying exclusively on individually defined, highly subjective, and essentially unknowable expectations, MDT assembles various theories of motivation with the goal of clarifying how discrepancies are arrived at (see Table 2-3). Perceived discrepancies are based on social comparison with one's peers (had-others have); they are based on the fit between person-environment in terms of needs (had-need to have); they are based on equity assessments (had-deserve); they are based on developmental theory evaluating past (had-past best), present (had-expect by now) and optimism for the future (had-expect in future). A final discrepancy, utilitarian-based aspiration assessments (had-want), is to be explained by the previous six discrepancies. MDT has shown that aspiration theory, social comparison theory and equity theory had the highest influence on satisfaction scores (Michalos, 1985: 381-382, 390). MDT explained, on average, fifty-three percent of the variance in global satisfaction, indicating a reasonably good model on which to base our analyses of nonresponse error.

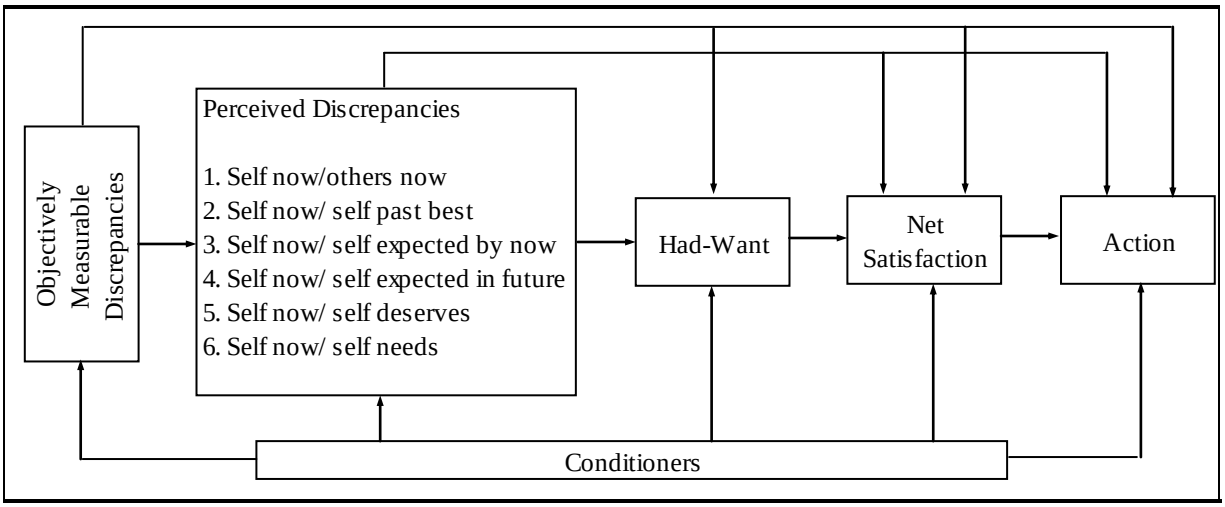
Table 2-3 Theoretical Expectations of Subjective Discrepancies

Theoretical Expectation	Discrepancies
Aspiration theory	What one has - what one wants
Social comparison theory	What one has - what relevant others have
Developmental model	What one has - the best one ever had in the past
Dissonance theory	What one has – what one expected to have
Optimistic theory	What one has - what one expect to have in the future
Person-environment theory	What one has – what one needs
Equity theory	What one has - what one deserves

Using MDT however, poses some conceptual difficulties that must be tackled in order to render the analysis intelligible. First, the model is over-specified in several aspects of the

student experience and under-specified in another. It is over-specified because the number of explanatory variables in the MDT model greatly exceeds those retrieved in the literature in terms of determinants of topic saliency. Moreover, the determinants are primarily objective ones and/or conditioners of topic saliency, not on perceived ones per se. Social comparison theory, for example, which was one of the most influential aspects in the expectancy-disconfirmation model, is not part of the literature on student satisfaction. These additional variables will be used, methodology permitting, in our analyses. The model is under-specified because determinants of satisfaction are conceived at the student level only. No contextual effects have been included in the model. Those effects, such as departmental crowding and involvement, will be included here.

Figure 2-14 Multiple Discrepancy Theory



Source: (Michalos, 1985: 354, 357)

Second, as Schultz (1995) is at pains to point out, there are conceptual ambiguities between what is to be considered a "conditioner" and what is an "objective discrepancy" on the one hand, and what is to be considered "objective" and what is "subjective" on the other.

An objective variable such as income can be just as easily conceptualized as a conditioner or an objective discrepancy; moreover, it is not entirely clear how income, in fact, translates into what students perceived they had in order to perform their subjective expectation-disconfirmation. Distributive justice, under equity theory, for example, certainly has an important subjective quality that is, hopefully perhaps, not entirely disconnected from an observable objective inequity. For the purpose of conceptual clarity, conditioners are defined as moderating variables that specify the relationship between satisfaction and action. Objective discrepancies refer to the quantifiable, physical aspects of the discrepancy which, following a realist epistemology, should exert some influence on how one perceives what one had. Despite these important conceptual considerations, socio-demographic variables are not expected to be important predictors of topic saliency bias. In fact, in a review of social indicators and quality of life (Stassen & Staats, 1988) found that such predictors explained no more than 10% of the variance on quality of life national surveys. Unless there is clear evidence of experiential disparity at the objective level such as blatant discriminatory practices operating at the university level, it is not expected that any such predictors will explain any discernable bias.

Third, the three dimensions of topic saliency need to be translated into the MDT model. Topic intensity maps directly into attitudinal valence, or the range of satisfaction scores. Topic interest is defined in terms of discrepancies; one's propensity to cooperate with a satisfaction survey request will be based on whether or not one experienced or perceived a discrepancy for one or several of the types of discrepancies enumerated above. Topic involvement, not listed in the model, is to be conceptualized as "the amount of physical and

psychological energy that the student devotes to the academic experience" (Astin, 1999: 58); it will serve as a conditioner variable specifying discrepancies and satisfaction. Following MDT, topic interest and topic involvement will influence topic intensity. It is precisely when discrepancies reach a significant level, above or below the zone of indifference that topic interest and topic involvement will translate into topic intensity. The latter, here understood as net satisfaction scores, is perceived as the main cause of cooperation behaviour. Thus, topic interest (discrepancies) and topic involvement are the explanatory variables of nonignorability.

All this begs the question as to what happens during the interaction between the student and the survey material. Unless there is evidence of "attitudinal crystallization" (Schuman & Presser, 1981) on the part of students about their educational experience prior to a survey request is made, it is likely that topic saliency, if present at all, will be constructed out of the interaction with the survey material. Here introductory statements such as those found in cover letters, topic order of the questionnaire, and question relevance will all play a role in forming topic saliency. Moreover, as we have seen in previous sections, nonignorability can just as easily be rendered undetectable because of unreliable satisfaction measures, and/or just as easily be an artefact of response bias. It is imperative that any analysis performed on the data itself be accompanied by some controls over measurement error and the formation of topic salience generally.

2.6 Conclusion

There are good reasons to believe that satisfaction surveys are particularly prone to nonignorable nonresponse. Empirical cases have shown that bias can occur in all directions. For some cases satisfaction bias was found, others found dissatisfaction bias, and still others found no bias at all. A review of satisfaction-based theories of behaviour across several fields, including political participation and consumer research, has also shown a disparate array of causes. Amidst all these potential causes for nonignorability, we have synthesized rudimentary hypotheses on the relationship between nonresponse to satisfaction surveys and nonresponse error. We also have developed a conceptual model to specify this relationship. The key questions are whether student satisfaction surveys are topically salient enough to cause nonresponse bias, and whether any of these hypothesized behaviours are indeed capable of explaining and specifying nonresponse error.

Chapter 3
Research Methodology

3.1 Introduction

The central aim of the research is to uncover whether and how nonresponse is ignorable or nonignorable with respect to key survey variables salient to the graduate experience. To that end, the research method deployed in nonignorability studies differs from the more traditional research structure. One will not find, for example, the typical dependant variable to be explained by predictor and control variables. Instead, one will encounter a pre-defined empirical model within which, for a range of response rates, multivariate coefficient variations need to be explained, specified, or interpreted. In essence, response rate is the independent variable and the empirical model the outcome. At times, the response rate variable is entered directly in the empirical model as a contextual variable; at times, a comparison of substantive models under different response rates is performed. As a result of this peculiar data analysis strategy, predictors embedded in the empirical model become at once the search for, and the expression of, nonresponse bias. The search for such variables is based as much on theoretical grounds as it is based on an empirical exploratory search. This mixture of experimental and anthropological research styles is a reflection of the difficulty of acquiring information on a population that is not responding, and perhaps not responding because, in this case, they are satisfied or dissatisfied with their experience at Waterloo.

This chapter outlines the methodology followed in this study. First, the section on research design will present the strategy that will permit the juxtaposition of all pieces of evidence found. This is followed by a critical assessment of the various quantitative data sources, and measurement models used in this thesis. Lastly, fieldwork interview strategies will be discussed.

3.2 Research Design

In determining the causes and consequences of nonignorable nonresponse to satisfaction surveys, we need to explore a wide range of variables that will help us specify, explain or interpret how, why and for whom satisfaction comes to play a role in the decision to participate. Survey cooperation is conditioned by the socio-cultural context, the survey design, the sponsoring agency's expectation and competence, and the sampled person's characteristics (Brehm, 1993; Goyder, 1987; Groves & Couper, 1998; Lessler & Kalsbeek, 1992). Several nonignorability studies have focused on the objective characteristics that differentiate respondents from nonrespondents (see chapter 2). What is omitted in such studies is an analysis of survey cooperation conducted at the subjective level. There remains much debate and uncertainty as to how cognitive determinants such as saliency and heuristics play a role in nonignorable response behaviour. Despite a tacit consensus on the methodological individualist premises of survey cooperation, little research seems to connect the objective socio-demographic analyses to cognitive ones.

The first component of this research is an attempt to expand the breadth and scope of nonignorable studies by the incorporation of cognitive interviews. Cognitive interviews[35] are analogous to a face-to-face interview but are structured in such a way as to tease out the thought process of the participant as he or she performs a given task. Through techniques of think-aloud, concurrent probes and retrospective probes, we are given a glimpse of what the participant understands, recalls, evaluates, and subsequently decides. Such interviews are commonplace in testing surveys for ambiguous wording, accurate recalls of past events, intelligibility of skip patterns, cognitive burden and question threats. Our use of cognitive interviews, however, is not intended to ascertain the quality of the questionnaire per se. In the attempt to understand the cognitive dimensions of nonignorability, the aim is, on the one hand, to uncover whether the survey is salient, a burden, psychologically threatening, and to assess the impact of measurement error, both bias and variance, on the detection of nonresponse error.

The second component of this research is an expanded quantitative analysis of nonignorability on two simultaneous fronts. First, empirical studies of saliency and satisfaction bias have provided only a scant quantity of predictors of satisfaction bias. In chapter 2, several potential determinants of topic saliency were uncovered; they will offer more vocabulary as to why alumni would refuse, based on their satisfaction, to answer a satisfaction survey. Second, empirical studies on nonignorability have been predominantly

[35] For a review of the literature and exemplars on cognitive interviews, see (Billings-Gagliardi et al., 2004; Campanelli, Martin, & Rothgeb, 1991);(Collins, 2003; Conrad & Blair, 2001; Conrad, Blair, & Tracy, 1999; Ericson & Simon, 1980; Fathi, Schooler, & Loftus, 1984; Jobe, 2003; Presser et al., 2004; Redline, Smiley, Lee, & DeMaio, 1998; Sudman, Bradburn, & Schwarz, 1996; Willis, 1999; Willis, DeMaio, & Harris-Kojetin, 1999; Willis, Schechter, & Whitaker, 1999)

confined to tabular results. The problem with such studies is that it is not always possible to control for confounding variables, and worse, to discern between a missing at random condition (ignorable) and a not missing at random condition (nonignorable). To further our analysis, we will test the effect of response rate variations[36] on a satisfaction regression model constructed from predictors of topic saliency. Taken together, these two approaches will permit this study to undertake a more pertinent and fundamental research question: does a change in response rate alter our understanding of student satisfaction? [37]

To that end, this research will follow a sequential mixed method design[38] (Tashakkori & Teddlie, 1998: 46-47). Under such a design, results from cognitive interviews and from the analysis of survey response effects on statistical models will be used in a complementary manner. Both are aimed to ascertain the degree to which topic saliency, understood here as student satisfaction, is the basic motive for answering or declining to answer the alumni survey. Looking at the same facet of survey response behaviour, they will elucidate its objective and subjective dimensions. Both methods will also expand the breadth and scope

[36] Techniques utilized in bias analyses are diverse. Commonly, researchers will utilize late respondents (e.g. converted refusals, late contacts) either to compare with early respondents, or to simply swell the response pool; other methods may include varying the survey design, conducting a study on nonrespondents, or comparative datasets such as population frame data, or an external surveys (Groves, 2006b; Porter & Whitcomb, 2005). The precise technique ultimately depends upon whether the researcher is conducting primary or secondary analysis. The former allows greater freedom to choose which technique is most appropriate; the latter, like all secondary analyses, limits the researcher's hand. Bias analysis on secondary sources depends on the quality of the dataset, the availability of paradata that accompanies the dataset, and the type of analyses to be conducted. For these reasons, the general techniques used to test our empirical model will be extracted once a dataset assessment is performed. Precise techniques will be elaborated at each analytical chapter.

[37] For similar studies that address ignorability in this manner see (De Leeuw et al., 1996; Goudy, 1976, 1978)

[38] In borrowing from the multi-method approach, we are not alluding to the "incompatibility thesis" (Howe, 2006), or the epistemological debate positivism-interpretivism (Blaikie, 1991). The juxtaposition of cognitive interviews with quantitative analyses does not take this research outside the basic "positivist" (Polkinghorne, 1983) epistemology. Our usage of the multi-method approach is to provide the analytical procedures to assemble several pieces of evidence drawn from several data collection techniques, which are all essentially housed in the same epistemological tradition.

(J. Green, Caracelli, & Graham, 1989) of our knowledge of the meaning of topic saliency itself. On the one hand, quantitative analyses will provide the added vocabulary to ascertain which aspects of the alumni's experience seem to be operating in determining response behaviour; on the other hand, qualitative analyses will provide much needed information on the alumni-survey material interaction, and how measurement error relates to nonignorability.

3.3 Quantitative Approach

The mixed mode research design begins with the quantitative approach to the study of nonignorable nonresponse. Its aim is to test for key salient experiential elements in the alumni graduate experience, as well as the effect of survey design, which may degrade or enhance topic saliency. To that end, this section will begin by assembling and assessing databases, to be followed by the operationalization of concepts elaborated in the previous chapter, and the general data analysis strategies used to assess nonresponse error.

3.3.1 Quantitative Data Sources

The assessment of bias in the reporting of satisfaction in a satisfaction survey will draw from four data sources: the University of Waterloo's Graduate Student Exit Survey 2000-2002, GSES 2000-02 population frame data obtained from student records, GSES 2002 factorial experiment on survey implementation, and Statistics Canada's National Survey of Graduates 1995. This section is composed of four subsections. The first two subsections present datasets from which nonresponse error is to be estimated; the remainder subsections present external datasets to be used as comparative measures. Each subsection describes the

data source in question, as well as its relative strengths and limitations associated with the goals of this study.

3.3.1.1 Graduate Student Exit Survey

The Graduate Student Exit Survey (GSES) constitutes the main source of data used to ascertain the presence of nonresponse error. This section begins with a description of the GSES survey, to be followed by a critical assessment of its suitability to our research goals.

Survey Description. For three consecutive years, at each convocation period (June and October), the University of Waterloo's Graduate Studies Office (GSO) mailed an exit survey to all graduate students terminating or graduating from their respective graduate program. The present author was responsible, under the supervision of the Dean of Graduate Studies, for the administration of this survey. Graduate students' opinions were sought on their overall progress at Waterloo, on the availability and quality of resources, and on the relationship between various facets of their degree program and the University's faculty members, staff and services. Figure 3-1 below summarizes the population frame.

Table 3-1 Population Frame Breakdown

Population Included	2000	2001	2002
Alumni	656	717	734
Withdrawal, Failed to register	59	90	0
Total	715	807	734

Only graduating students were surveyed during the convocation of 2002

The Graduate Student Exit Survey contains twelve sections (see appendix A). Each section targets one specific area of students' experience at Waterloo. It includes sections on program duration, research supervisor experience, comprehensive examinations (PhD only),

conference attendance and funding, financial support, teaching assistantships, department, Graduate Studies Office, University services, and Graduate Student Association. It also includes sections on the students' experience after Waterloo: intentions regarding future education, and their professional employment. Fifty questions (12 pages) were administered.

The survey, complete with cover letter and self-stamped return envelope, was mailed two weeks after convocation. No names or identifying codes were attached to the questionnaire or on the self-stamped return envelope. Anonymity, as opposed to confidentiality, was insisted on by the GSO, being deemed necessary to instil complete confidence of the alumni to voice their opinions without fear of reprisals. No specific reprisals were noted in the GSO's submission for ethical approval, other than the possibility of the alumni being identified. However, it is clear that the fear of being identified, considering that reference letters are required for future employment, would be a strong deterrent against the voicing of one's opinion. Even with anonymity, the prospect of being identified remains, as some departments will only graduate a few students each year.

A return postcard, to be mailed back separately, was included with the initial survey package, allowing the removal of the respondent from the mailing list without compromising his or her anonymity. A postcard reminder was sent to every student two weeks after the initial mailing. The remaining nonrespondents (living in Canada and the U.S.) were randomly selected for phone/answering machine reminders three weeks after the postcard reminder. This protocol was followed for the first two years. On the third year, spring convocation of 2002, a factorial experiment (see below) was conducted to assess the differential impact of changing the mailing dates and the final reminder type. During the fall

convocation of 2002, the survey protocol was modified to include a dean's letter as a final reminder instead of a phone message reminder.

Response rates for each of the three convocation years are typical of other alumni exit surveys but disconcertingly low nevertheless. Despite several modifications of the questionnaire design[39] and implementation, discussed above, the response rate[40] changed very little over three years. Convocation 2000 had a response rate of 35.9%, convocation 2001 had 33.0% and convocation 2002 had 39.4%. All three survey-years will be merged for most of the analysis herein, giving a total of 812 cases available for analysis.

Database Assessment. Satisfaction surveys, whether commercial or institutional, frequently contain sensitive information on client/customer attitudes towards products and services of the sponsoring organization. These surveys are rarely distributed publicly. In that regard, the Graduate Student Exit Survey is no different. However, my status as a UW student and my previous work experience with the GSES has provided the unique opportunity to access and utilize the data set.

Beyond such opportunistic reasons, the data set is well suited to the task. First, the survey contains a myriad of satisfaction question items, many of which are compatible with the literature on student satisfaction. Second, university student surveys can tap into a rich source of frame variables to better assess nonresponse bias. While they do not contain

[39] Changes in the layout of the questionnaire were performed during the convocations of 2001 and 2002. See appendix A for details.

[40] Undeliverable surveys returned to the GSO were re-sent if a new address could be found; those remaining were classified as non-contact. All students found on the convocation list were deemed eligible to receive a survey. Consequently, response rate calculations were performed using the number of questionnaires returned divided by questionnaires sent. No distinctions between contact rates and cooperation rates were made.

student attitudes, expectations or satisfaction levels, frame variables can be used to assess the extent of the differences between respondents and non-respondents. Third, a random factorial experiment was conducted to ascertain the impact of varying final reminder formats. Such an experiment provides the unique opportunity to test varying response rates ranging from 22% to 41%. Finally, a few select questionnaire items mesh well with Statistics Canada's own National Survey of Graduates (see next section). Response rate variations, from 36% to 77.3%, can thus be potentially analyzed.

Nevertheless, the Graduate Student Exit Survey's implementation strategies severely hamper our study of nonignorability. The GSES, like most mail surveys, does not provide the ability to assess the relative proportions of non-contact and refusals. First, "return to sender" returned mail may not accurately reflect the true non-contact population (Moore & Tarnai, 2002). Second, the GSES might be prone to a high level of "third-party delivery system". Prior to leaving Waterloo, students were asked to produce three addresses where they could be contacted. Despite the fact that phone reminders confirmed reception of most of the surveys at students' home addresses, it may be better to speak of remote-contact response. Since many students have their mail forwarded directly or indirectly to a new address, they may receive the initial mailing (cover letter and questionnaire) only, or receive the reminder on the same date. There may be an interaction effect between remote-contact and refusal. There is the possibility that alumni operate on an implicit triage of letters based on third-party interpretation of received mail from the University of Waterloo. Much of the contact depends on what (presumably) parents will do with the survey envelope and how they will interpret its content if asked to open it. It is likely that mail from the University of Waterloo

is stacked and kept until return of the alumni. In all cases, refusal because it's too late to send the survey, or because it's not deemed important or salient are likely reasons to refuse to respond.

Furthermore, the response outcome variable does not discern between early and late response. The comparison of early to late respondents requires direct access to the mail delivery system. Since questionnaires were received at the Graduate Studies Office and only released to the author at a later date, a reliable time line could not be constructed. In hindsight, even if such a strategy were executed, it would not have been meaningful for this study. To be meaningful, time-based comparisons, which presuppose that late respondents are akin to nonrespondents, require a reasonable response rate. One cannot intelligently compare, for example, 20% late respondents with 65% non-respondents, and simply assume that the former is similar to the latter. The assumption of a continuum of resistance, whereby late responders are akin to nonrespondents, has not only been disputed empirically (Lin & Schaeffer, 1995), but presumes precisely what it is that we are trying to research: the specificity of nonrespondents in relation to respondents. Finally, there are doubts whether a single motivational variable – that of being late – can serve as a proxy variable for non-respondents. Being late may simply be a consequence of remote-mail contact as mentioned earlier.

Finally, since the survey was designed to be completely anonymous, it is impossible to conduct a study on nonrespondents. The mailing of a second questionnaire to a random sampling of nonrespondents exposes survey results to duplication since no identification codes were attached to the original questionnaire. Only a phone interview, in which the

student is explicitly asked whether he or she has answered the survey, may alleviate duplication error. However, previous attempts to contact students during the final reminder phase of the survey have shown how extremely difficult it is to reach them by phone; parents were the most likely contact. Consequently, a study of nonrespondents may not yield a high response rate. Contact difficulties aside, one must have some idea of the nonresponse bias on the survey variables of interest in order to construct a sample size that would detect the presence or absence of bias (Groves & Couper, 2002: 20-21). Assuming simple random sampling and a population satisfaction proportion of 0.5 and a probability of error (alpha) set at 5%, a study of nonrespondents would require a sample size of 1067 for an estimated bias of ±3%, or 11 for an estimated bias of ±30%[41]. Considering all the above, notwithstanding the fact that the determination of the magnitude of the bias is a desired outcome of this study, a follow-up of nonrespondents is not a useful research strategy.

3.3.1.2 GSES Factorial Experiment on Survey Implementation

In addition to obtaining socio-demographic and program-related variables associated with the alumni population frame, the Graduate Student Exit Survey also provides the opportunity to test the effect of implementation strategies (survey design) on key survey variables. During the spring 2002 convocation, a 2X2 factorial experiment was conducted in the hope that a change in the implementation strategy, namely a change in the final reminder format and a change in the mailing dates, would be accompanied by a better response rate. It

[41] Sample size calculations for mean proportions given a simple random sample were performed using confidence interval formulas (Healey, 1996: 165-166):

$$Bias = P_s \pm (Z)\sqrt{\frac{P_u(1-P_u)}{N}}$$

was felt that the timing of the mailing period, two weeks after convocation, although salient, was a rather busy time for recently graduated alumni. A mailing date of September was tested in the hope that any uncertainties about life arrangements in terms of employment and housing would be settled by September, and they would thus be more receptive to answering the survey. In other words, the perceived costs associated with answering the survey such as time constraints or geographical dislocation would be diminished at a later date. Moreover, it was also felt that a telephone reminder was costly in relation to resources and ultimately did not reach the alumni themselves. The answering machine and occasionally the parents were the likely contact. A final reminder, signed by the dean of Graduate Studies, might invoke the authority of the sponsor (Groves & Couper, 1998), and induce alumni to return their surveys.

Table 3-2 2X2 Factorial Group Experiment

	Treatment	**Post test**
R	X_1 July mailing date, phone message second reminder (control group)	O_1
R	X_2 July mailing date, letter second reminder	O_2
R	X_3 September mailing date, phone message second reminder	O_3
R	X_4 September mailing date, letter second reminder	O_4

In the spring convocation of 2002, a list of alumni was acquired from the graduate studies office (N=331). The list was subdivided into four equal random groups (see table above). The first group, the control group, received the usual survey protocol. They were phoned five weeks after the initial mailing, two weeks after convocation. The second group had the same mailing date, but received a dean's letter five weeks after the initial mailing date. Groups three and four had their surveys mailed in the first week of September; the

former received a phone message, the latter were mailed a dean's letter. Since GSES is an anonymous survey, the groups were tagged on the return envelope. A hand-written room number was placed on the return envelope – a different color pen for each group. Upon reception, each questionnaire was labelled according to its respective factorial grouping.

Knowledge of which survey respondents are associated with each grouping will be used to test whether a change in a given implementation strategy was accompanied by a change in the values of key survey variables. We can also ascertain survey respondents' representativeness by comparing them to population data. The number of cases per grouping is quite low: 83 cases sampled per alumni group, and response rates for each group range from 22% to 41%. The analysis would not have enough cases to proceed with the full logistic model, as developed in section 3.3.2. Moreover, the statistical power is quite low. At maximum cooperation (100%), given a probability of error of 5%, and a desired statistical power of 0.80 (80% chances that the statistical tests will be significant, see (J. Cohen & Cohen, 1983: 60-61, 529)), regression models could only detect moderate correlations ($r \geq 0.3$). Since the goal of the research is to assess biases stemming from cooperation of less than 100%, the resultant statistical power and the ability to detect significant changes in regression models will also be greatly diminished. As a result, we may not always be in a position to adjudicate whether ignorability was due to a lack of bias or due to low statistical power.

3.3.1.3 Comparative Dataset: GSES Graduate Student Population Records

One of the main advantages of using student-based satisfaction surveys is the rich store of information on recent alumni contained in the university's student records. These records contain several socio-demographic and program-related variables that map directly to questions already asked in the GSES. These variables will provide the necessary information to assess the degree of sample representativeness as well as any distributional biases that may have occurred on the survey variables of interest. Several such variables were requested from the Graduate Studies Office. Unfortunately, not all requested variables were deliverable; some variables were omitted due to unavailability and time constraints, others due to database upgrades. Table 3-3 shows a list of the requested variables and those obtained from the GSO:

The availability of frame variables has also opened up the possibility of matching each respondent with the alumni population data. Merging both databases using matching variables available would provide the added benefit of deducing who answered the survey and who did not. Response propensity models could then be constructed and compared with key survey variables of interest. Unfortunately, the requirement of anonymity foreclosed the possibility of matching each respondent with the alumni population data. The extent of socio-demographic variables such as convocation year, department, degree type, citizenship status, and degree program start dates proved to be insufficient to consistently match respondents to their student records, or to ascertain whether any matches found were indeed the correct ones. For a number of departments, these merging variables taken together could still point to a dozen different alumni. More importantly, and disturbingly so, a deeper comparison

between frame and respondents variables revealed that, for some alumni, degree program start dates did not always match. Recall questions, particularly those involving frequencies and dates, are prone to response errors such as a "telescoping effect" (Rossi, Wright, & Anderson, 1983). The presence of such errors only compounded the problem and any attempts at merging data sets were abandoned. The frame data set will be used in an aggregated form only.

Table 3-3 Frame Variables Request

Variables Requested	Received
Sex	✓
Visa Status	✓
Enrolled degree program	✓
Age (not in survey)	✓
GPA (not in survey)	✓
Enrolled department	✓
Previous degree	
Supervisor assignment term	
Comprehensive examination completion term	
Thesis proposal approval term	
Number of teaching assistantships received	
Time to completion (program start/end date)	✓
Total financial support	
Mailing city, province, country	✓

Another useful source of data, beyond student records, is the report produced by the University of Waterloo's Institutional Analysis and Planning (IAP) called the *University of Waterloo Senate Finance Committee Operating Budget Supplementary Data*. It is freely available on the IAP's website. It is a summary of graduate and undergraduate enrolment, faculty and staff complement, and teaching unit budgets. This data is extremely useful in complementing the aggregated alumni population data from the GSO. Of interest is the list of

monies, faculty and staff complement for each department. They will provide a means to test whether faculty to alumni ratio in each department was partly responsible for the alumnus' satisfaction scores, and by extension, his or her decision to participate in the GSES survey.

3.3.1.4 Comparative Dataset: National Graduate Survey

The National Graduate Survey (NGS) is conducted periodically by Statistics Canada. The survey is administered to a random sample of alumni students, living in Canada or the U.S, two years after graduation. The data set contains eighteen sections mostly related to their post-graduation labour force experience, but also contains questions on degree program experience. The stratified sample draws from a population of Canadian universities, community colleges and CEGEPs. Each student was contacted by phone during the three summer months. The NGS public use data set is of value to this research to the extent that it duplicates in section K (program evaluation), a number of our own questionnaire items (see Table 3-4). It also surveys the same population group: recently graduated alumni. Finally, NGS has achieved a greater response rate than our own in-house graduate exit survey (GSES 36%; NGS 77.3%[42]). These two data sets permit one to parallel (Keeter et al., 2000)'s study noted in chapter two.

[42] Final response rate calculation was 77.3%. The total count of respondents, including refusal conversions is 43,040. The total count of eligible alumni in the sample was 55,649 (original sample of 61,759 – ineligible (2,298 + 3,093), duplicates (647), and death (72)). It should be noted that untraceable, yet presumably eligible alumni represented 11.3% of the original sample of 61, 759. See (Statistics Canada, 1995b: 20). If these latter alumni are excluded from the original sample, the response rate climbs to 89%.

Table 3-4 GSES and NGS 1995 Common Variable Set

Question Category	Question Item
Demographic Information	Sex, Marital Status, Dependents, Employment after graduation
Program Information	USIS, Degree program, Time to completion, Part-time period, Inactive period
Income	Student loans
Satisfaction	Quality of teaching, Computing facilities
Net Satisfaction	Would start program again, Would return to Waterloo

Unfortunately, due to unforeseen difficulties in accessing the NGS 1995 and 2000 master files[43], this research could not proceed with the original intention to extract, from the NGS dataset, alumni responses from the University of Waterloo. The consequence of using only the public version of the NGS 1995 data set was to deal with the problem of coverage and the costs of greatly diminishing validity. Two options were available. First, comparisons could be made at the provincial level. Using the study location variable, we could isolate Ontario alumni, and proceed with comparisons with alumni from the University of Waterloo. This assumes, of course, that University of Waterloo graduates are not systematically different from other graduates within the province of Ontario – an assumption immediately suspect if one glances at Waterloo's performance indicators (employment rates, OSAP

[43] The NGS 1995 and 2000 master data sets were inaccessible. Research Data Centres (RDC) operating within several University Campuses provide the opportunity to access Statistics Canada master data sets otherwise not available to the public. Unfortunately, and previously unbeknownst to this researcher, it appears that the 1995 data dictionaries did not accurately reflect the content of the NGS databases located in RDC offices. Variables that would have identified each University, -- ERLID (institution identification number) for instance --, have been systematically deleted from the 1995 and 2000 data sets located in RDC. We can only surmise that RDC was never supplied the true master file. Our request for RDC entry was refused on the basis that ERLID did not exist. Since no appeal was possible, and our original proposal took four months to be processed, this research confined itself to the publicly released 1995 data set. As a result, what could have been an excellent opportunity to compare 2000 Waterloo graduates with our own survey was lost.

default rates, and completion rates) at the undergraduate level. This option was abandoned upon the discovery that doctoral degree program identification was systematically deleted for some provinces, including Ontario. This would have reduced the analysis to Ontario master's alumni only. The second option was to use the entire national dataset in the hopes that, on average, when large and small universities are combined together, a better fit with the University of Waterloo will emerge. Opting for this "solution" does not eliminate the confounding bias introduced by coverage mismatch. It does offer, however, an empirical estimate from which a guarded comparative analysis can be performed on all alumni, including doctoral ones. This is ultimately the compromise to be made when secondary datasets are used.

At a more general level, there are four other important limitations in using the NGS data set. First, the two surveys were not implemented within the same context. GSES results may overestimate satisfaction if an authority relationship elicits a social desirability effect. While Statistics Canada could also be construed as an authority, it does offer some neutrality to the respondent. Second, the surveys used different modes of data collection. The GSES self-administered survey may yield more forthcoming and less acquiescent answers by virtue of being anonymous as opposed to the NGS phone survey. Third, the survey request was placed at convocation time for GSES and two years after graduation for NGS. A dissatisfaction bias may be introduced due to recall bias and use of information not currently available at graduation (Carr-Hill, 1992). Fourth, and perhaps least important, the surveys were not conducted on the same year. Any difference in the estimate may prove to be a true population change that occurred between 1995 and 2000. Our cross-sectional logistic design

cannot compensate for history effects. Despite these confounding errors, large estimate variations would indicate a potential problem of nonignorability.

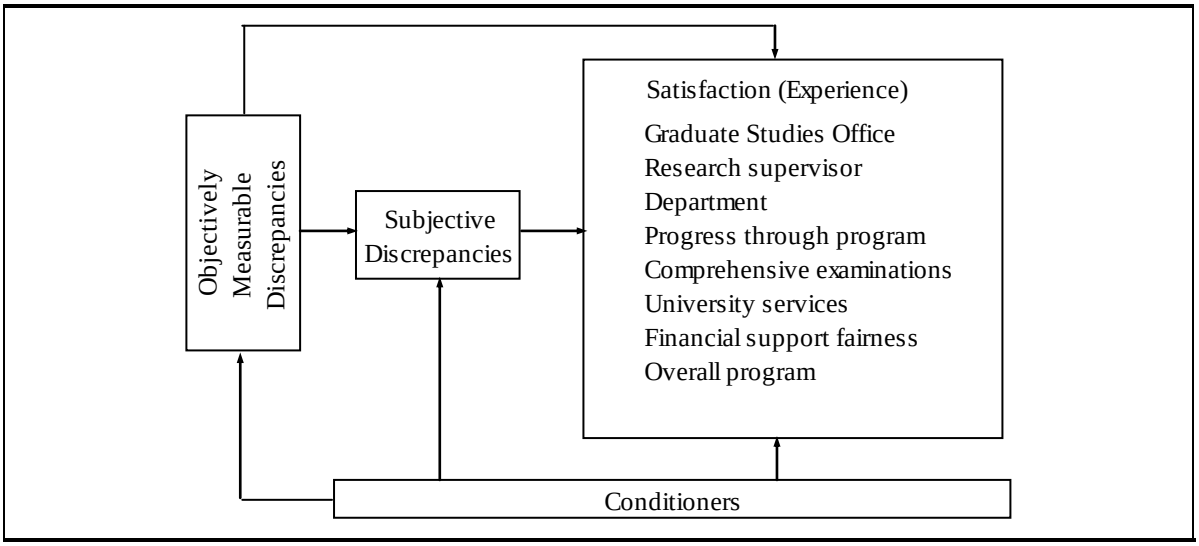
3.3.2 Graduate Student Satisfaction Measurement Model

As we have already alluded in the introduction, research on nonignorable nonresponse does not follow the classic methodology generally associated with quantitative analysis. While both may devise empirical models that would explain outcome variables of interest, the former is less interested in the explanation of variance in outcome variable than the latter; instead, the former makes use of an empirical model to understand the effect of nonresponse on the bivariate and multivariate relationships. As we have seen in chapter 2, the elaboration of nonignorable nonresponse requires predictors that are both strongly related to the decision to cooperate with a survey request, and with the variable of interest causing nonignorable nonresponse. As a consequence, the use of any empirical model must also contain variables that are conducive to explaining nonignorable behaviour on the part of the sampled individual, and must exhibit strong relationships with the variable of interest causing nonignorable nonresponse. Michalos' multiple discrepancy theory (MDT) was developed, in Chapter 2, to provide a coherent conceptual framework for grouping predictors of graduate students' satisfaction most likely to be topically salient, and as a result most likely to explain nonignorable nonresponse. Equally important, MDT explained 53% of the variance (Michalos, 1985) in student global life satisfaction – an outcome variable not too dissimilar from ours. MDT is a good empirical model to observe nonignorable nonresponse.

However, the operationalization of Michalos' MDT conceptual model into an empirical model was restricted by the number and type of variables available in our datasets. Two factors limit our ability to test the MDT conceptual model fully. First, satisfaction, in MDT, was conceptualized as the discrepancy between what one had and what one wanted, deserved, needed, saw others have, had in the past, hoped for the future, and expected to have by now. These expectation-discrepancy measures presuppose a comparison judgment mechanism, usually provided in the formulation and/or design of the question itself. The questionnaire items found on the GSES pertain mainly to attitudes measured as a four point Likert scale ranging from "very satisfied" to "very dissatisfied". These questions do not, on face validity, map well to a conception of satisfaction as expectation-discrepancy. They may presuppose, if Michalos' model is correct, a comparison-based judgment but these expectations remain unknowable when asked as an absolute attitude question as opposed to a relational attitude question (e.g. in relation to expectancy, what they had). The bulk of these Likert-based satisfaction questions will be relegated as outcome variables, to be submitted to factorial analysis; some, for which clear referents of discrepancies can be ascertained, will be placed under the subjective discrepancies variable set. Second, following Schultz (1995)'s criticism, MDT model does not clearly differentiate conditioners from objective discrepancies. The latter will be explicitly operationalized using indicator that objectively represent one of the seven discrepancies conceptualized in the MDT model; conditioners are operationalized using indicators that situate alumni in terms of their socio-demographic status, their particular educational program or the departmental context into which they were enrolled into.

Figure 3-1 illustrates the empirical model developed from the available variables present in the GSES data set. Satisfaction is the outcome variable of interest which, presumably, is cause of nonignorable nonresponse. Objective and subjective discrepancies, as well as conditioners, taken together explain why a given alumni would be satisfied or dissatisfied with his or her graduate experience. These explanatory variables are also hypothesized, as we have elaborated in chapter 2, as linked to the decision to cooperate with a survey request. Each subsection elaborated below will operationalize one conceptual block of the MDT model by specifying and assessing indicators drawn from the GSES and NGS datasets. The first subsection tackles the outcome variables, student satisfaction, hypothesized as the cause for nonresponse. The following subsections will deal with objective, subjective and conditioners. The final subsection addresses the issue of measurement error and ignorability by specifying variables that will control for item nonresponse.

Figure 3-1 MDT-Based Alumni Satisfaction Empirical Model



Adapted from (Michalos, 1985)

3.3.2.1 Outcome Variables: Satisfaction with Graduate Experience

The GSES data set contains fifty-five satisfaction questions spanning all aspects of graduate studies, from supervisory relationship, department, and graduate studies office, to university services, graduate student association, teaching experiences, and financial support. Instead of focusing on any one question, all satisfaction questions were submitted to a multiple correspondence analysis[44]. The intent was two-fold: first, alumni need to be grouped according to their similarity on satisfaction questions; as hypothesized in chapter, alumni who share similar satisfaction scores will also share a common response pattern. Second, item nonresponse to satisfaction questions needed to be included as a meaningful trait of survey response; unable to completely disentangle inapplicable questions from unanswered questions, it was deemed preferable to use the item nonresponse as valid data. Using all satisfaction categories (very satisfied, satisfied, dissatisfied, very dissatisfied), including no opinion and not stated, the multiple correspondence analysis (MCA) extracted seven clusters. These seven clusters constitute a global measure of satisfaction with one's degree program.

Multiple correspondence analysis results are presented in Table 3-6. The number of cases and the percentages of the combined 2000-02 alumni pool for each cluster are indicated in the last two columns. A short cluster description was extracted from cross-tabulations of basic socio-demographic and program characteristics with cluster membership. Of the seven clusters, two are of particular interest to our research: cluster two and four.

[44] Our technique is partially based on similar studies such as (Blasius & Thiessen, 2001; Puyo-Savary, 2004). The software utilized for multiple correspondence analysis was SPAD 5.6 from Decisia.

Table 3-5 Global Graduate Experience Satisfaction Cluster Membership

Clusters	Cluster Description	N	%
GSES 2000-2002 Combined			
1	Composed of doctoral alumni only with a significant high proportion of visa students, and science-related departments, married with children, who reported research activities and income, and who are significantly more likely to stay an extended period at Waterloo. They are for the most part satisfied with their degree program except for GSO registration procedures and the finance office.	116	14.3
2	Characterized by alumni who are entirely very satisfied with every aspect of their degree program particularly their experience towards their supervisor and department. No MAcc alumni. This cluster will be labelled "elated".	154	19.0
3	Predominantly composed of master's with research alumni who are essentially satisfied with their degree program. No MAcc alumni	186	22.9
4	Characterized by alumni who are dissatisfied to very dissatisfied with each aspect of their graduate experience, but particularly and predominantly directed at their supervisory experience. No MAcc alumni. This cluster will be called "irate".	54	6.7
5	Predominantly composed of master's with research alumni who are essentially satisfied with their degree program. No MAcc alumni	153	18.8
6	Residual category.	18	2.2
7	Predominantly master's of accounting students (MACC program). This cluster is marked by a large number of missing and/or no opinion questionnaire items. Beyond missingness, the experience is largely satisfactory. No Doctoral alumni.	131	16.1
	Total	812	100

Cluster two, the "elated group", regroups alumni whose experience was generally qualified as "very satisfied"; cluster four, regroups those who responded "dissatisfied" and/or "very "dissatisfied" to the majority of the questionnaire items. The test-value reported for each questionnaire item (not shown, see Appendix C) clearly indicates that "experience with supervisor" differentiated these two clusters from all others in terms of graduate experience; and differentiates them from each other in terms of satisfaction level. Unfortunately, doctoral students were not represented in either of these two clusters, likely due to their uniqueness with respect to their experience with comprehensive examinations. For that reason, measures

of satisfaction were singularly derived from two facets of graduate experience: experience with supervisor(s), and experience with the department. These two facets (not shown) are likely to be the last components of a degree program that alumni will have experienced before graduation (see Appendix C).

Table 3-6 Other Global Graduate Experience Satisfaction Cluster Membership

Clusters	Cluster Description	N	%
GSES 2002 Factorial Experiment			
1	Composed predominantly of alumni satisfied with their graduate experience.	79	72.5
2	Residual category. Predominantly missingness, the remainder no opinion	6	5.5
3	Characterized by alumni who are entirely very satisfied with every aspect of their degree program particularly their experience towards their supervisor and department. No MAcc alumni. This cluster will be labelled "elated".	20	18.3
4	Residual category. Missingness on all questionnaire items	4	3.7
	Total	109	100
GSES 2000-02 & NGS 1995 Combined			
1	Characterized by alumni who are very satisfied with every aspect of their degree program particularly their experience towards their supervisor and department. No MAcc alumni. This cluster will be labelled "elated".	1932	20.4
2	Composed predominantly of alumni satisfied with their graduate experience.	5782	61
3	Characterized by alumni who are dissatisfied to very dissatisfied with each aspect of their graduate experience, but particularly and predominantly directed at their supervisory experience. No MAcc alumni. This cluster will be called "irate".	899	9.5
4	Residual category. Missingness on global satisfaction measures	866	9.1
	Total	9479	100

Two other global measures of satisfaction, shown in Table 3-6, were also extracted using multiple correspondence analysis. First, the dataset from the factorial experiment on survey protocols yielded four clusters. These clusters were derived using the same questionnaire items used in the global satisfaction measures in the combined GSES 2000-02 dataset. The cluster sizes are small which affect data analytical strategies. More importantly,

MCA did not retrieve any irate clusters. Cluster three regroups alumni who had a very satisfactory experience; cluster one, the largest, regroups satisfied alumni. Second, the NGS-GSES datasets combined also yielded four clusters. Cluster one regroups "very satisfied" alumni, while cluster three regroups "dissatisfied" alumni. These clusters were derived from a limited set of questionnaire items, those of which are common to both GSES and NGS.

3.3.2.2 Conditioner Variables

The selection of conditioner variables was informed primarily by the objective of reproducing Michalos' model. To that end, socio-demographic and program variables were drawn from the GSES questionnaire. However, in order to accommodate the findings of the literature review, as well as to account for the hierarchical nature of the organization, the model was furnished with a measure of student involvement, and a series of contextual-level conditioners. As a result, the model now has three sets of conditioners:

Table 3-7 Indicators of Socio-Demographic Conditioners

Contextual Conditioners		**Indicators**
Sex		Women alumni
Income		Total income from all sources summed over five years
Ethnicity		Self-reported visible minority status
Social Support		
a)	Family	Marital status
		Number of dependants
b)	Peers	Number of Graduate House visits
Education		Master's no research (MAcc) alumni
		Doctoral alumni
		Science-related degree program

Socio-demographic and program conditioners. Socio-demographic variables such as sex and income mapped directly to questions already present in the survey. Reported incomes from all sources (see section ten of the GSES questionnaire) were summed together to arrive at alumni income over a period of five years. Ethnicity was crudely operationalized as self-reported visible minority status. Social support as a measure of satisfying relationships with friends and family was crudely operationalized as family status (marital status and number of dependants) and Graduate House visits (presumably one visits to exchange with peers). The variables age and social esteem were not part of the questionnaire, and could not therefore be included in the student-level satisfaction model.

Education was operationalized as "degree type" in three distinct categories. The category "master's no research" included masters of accounting (MAcc) only. This particular program was devised as preparation for accounting exams, and consists of an eight month course schedule. The remaining portion of masters programs was categorized as masters with research. Admittedly, programs such as masters of engineering (MEng), masters of architecture (March) and masters of environmental studies (MAES) could arguably be grouped into a separate category. These three degree programs in particular require a research project to be completed within one year, as opposed to the regular masters stream which requires a research thesis to be completed within two years. However, despite their differences, both require the presence of a professor who will supervise and/or advise during the completion of the research project. Doctoral studies were coded directly. Finally, in order to account for the distinction between science and art/social science related programs a dichotomous variable "science-related programs" was added to the model.

Involvement conditioners. In addition to socio-demographic conditioners, the graduate student satisfaction model also included, following the outcome of the literature review, the dimension of student involvement. Three aspects of student involvement were captured by the questionnaire. First, research and teaching involvement activity was captured by a battery of research thesis accomplishments such as conference attendance, paper presentation and paper publications; and using questionnaire items found in section nine, "teaching experience". A filter question posed at the beginning of sections eight and nine of the GSES questionnaire served to identify who had a research thesis and teaching assistantship roles as part of their degree program. Second, degree program involvement was operationalized using the part-time registration variable located in section six of the GSES questionnaire. Finally, the third aspect of involvement, organizational involvement, was measured using the variable "interest in participating in educational fair".

Table 3-8 Involvement Cluster Membership

Clusters	Type of Engagement		Cluster Description	GSES 2000-02		GSES Factorial	
	Roles	Activity		N	%	N	%
1	Both	Some	Mixed involvement; research conference attendance	413	51	66	60.1
2	Both	All	Involvement in all aspects	84	10.3	0	0
3	Thesis	None	No involvement beyond research thesis	111	13.7	19	17.4
4	None	None	No roles provided; no activity reported (MAcc)	204	25.1	24	22.0
			Total	812	100	109	100

All variables, except part-time registration, pertaining to student involvement were analyzed using multiple correspondence analysis. The outcome, shown in Table 3-8,

produced three distinct clusters of student involvement. The first and second clusters both regrouped students who had research and teaching roles, but only students grouped around cluster two were fully involved in all activities that such roles permit. Cluster three regroups students who were not given any teaching assistantships, had a degree program which encompassed a research thesis, but chose not to partake in any additional research activities whatsoever. Cluster four are those students, predominantly from the master's no research program (MAcc), who by definition hold no teaching/research roles or activities. The bulk of the respondents are located along cluster one, the mixed involvement category.

Figure 3-2 MCA Cluster Loadings on Involvement, GSES 2000-02

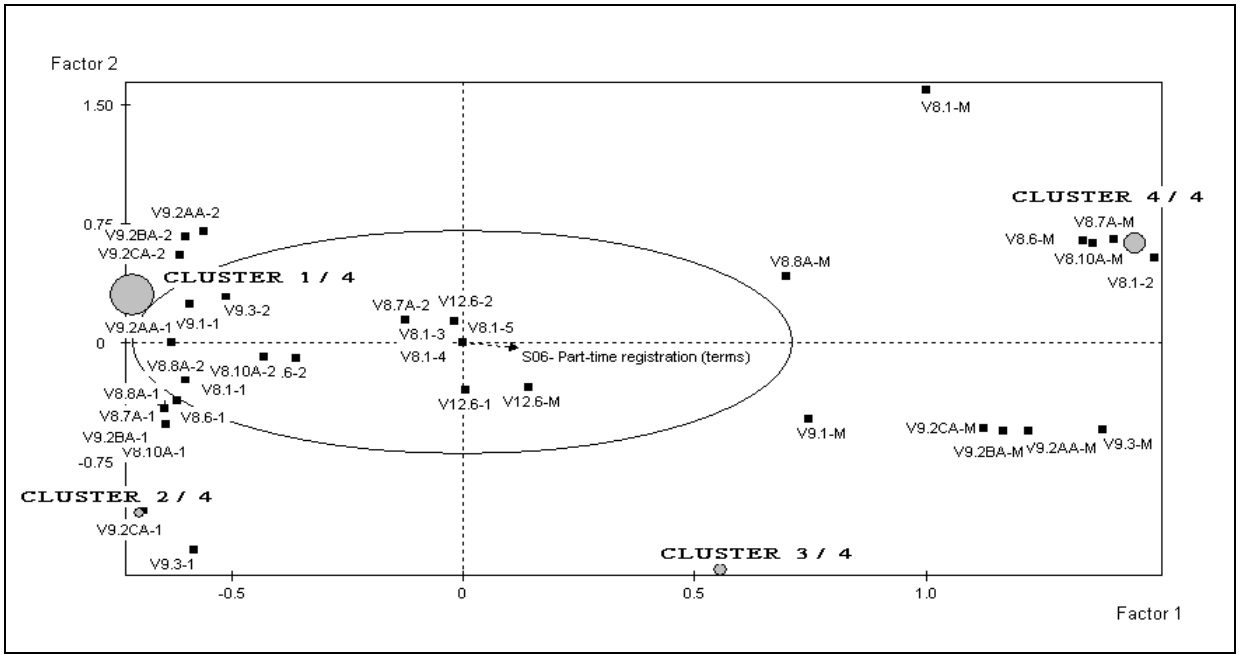


Figure 3-2 locates clusters and variables according to their relation to the first two factors of the multiple correspondence analysis. The first factor differentiates variables according to their value related to teaching roles. As already indicated above, cluster three

and four have no such roles and differ in this respect from cluster one and two. Factor two differentiates between clusters exhibiting missingness on teaching activity: cluster two and three have responded fully whereas cluster one and four have not. The presence of cluster one with cluster four seems to suggest that students grouped under cluster one have skipped several questions pertaining to teaching activities. Both of these clusters are similar with respect to missingness on participation in Trace programs, teaching certificate and department seminar. The correlation circle, and the regression line, "S06-Part-time registration" indicates where the quantitative variable "part-time" is located in the factorial graph. The direction of the regression line indicates that the more a student is registered part-time the more he or she will converge towards cluster three, no involvement. As such, the variable part-time, which is our only measure of involvement in the NGS database, is a valid but coarse substitute for student involvement at the teaching/research level.

Table 3-9 Indicators of Contextual Conditioners

Contextual Conditioners			**Indicators**
1.	Crowding		Department alumni count
2.	Involvement		
	a)	Departmental research focus	Graduate to undergraduate ratio
	b)	Experience with faculty	Full-time Faculty to all students ratio
	c)	Departmental resources	Average departmental budgets
3.	Cohesiveness		
	a)	Peer influence	Proportion of doctoral students
			Proportion of Macc students
			Average GPA
			Average deviation from normal program duration
	b)	Diversity	Proportion of women alumni
			Proportion of visa students
			Average alumni age

Contextual conditioners. Contextual or departmental variables were drawn from alumni population frame data and from reports provided by the Institutional Analysis & Planning Office at the University of Waterloo. Three sets of variables were constructed as shown in Table 3-9. The first, department alumni count, serves as a crude indicator of department size; the latter would be a truer indication of crowding within a given department, but the sheer size of alumni being pumped out at each convocation period is a reasonable substitute. A second set of contextual variables was constructed to tap into the concept of departmental involvement. Three variables were constructed: 1.) student-faculty ratio derived from full tenure professor count divided by the sum of undergraduate and graduate alumni counts; 2.) research orientation derived from graduate alumni count divided by undergraduate alumni count; 3.) departmental budget calculated from a combination of budget monies, total basic income units (BIU's), and full-time equivalent faculty (FTE's) units for each faculty between the year 2000 and 2002[45]. Departmental budgets were calculated for each year, and then averaged for the three-year period. A final third set assembles indicators of peer influence and peer diversity, both themselves concepts of departmental cohesiveness. Peer influence was constructed from program-related variables such as proportions of doctoral alumni, master's no research, average GPA, average deviation from normal program duration. Peer diversity, or student body composition, incorporated indicators such as the proportion of women alumni, and visa students as well as average alumni. age. The survey fatigue dimension, elaborated in chapter 2, could not be operationalized in this dataset. It is suspected however, that either some departments, offices, or the library system are

[45] The formula utilized to calculate departmental budgets from faculty budgets is as follows:
Departmental budget $ = ((Faculty budget/BIU) X (BIU/FTE)) x Departmental full-time professor count

conducting simultaneous, if only sporadic surveys themselves. These might have an unknown impact on nonignorable nonresponse.

3.3.2.3 Subjective Discrepancies

As we have already alluded to earlier, there are virtually no questionnaire items that would qualify as a subjective discrepancy as conceptualized by Michalos (1985). However, some variables harbour an implicit discrepancy and might conceivably be used as subjective discrepancies. Three such discrepancies can be extracted from the questionnaire. The first can be operationalized using the variables, "satisfaction with authorship issues on publications" (section 2, Q3) and "financial support was distributed in a fair and equitable manner" (section 10, Q3); both suggest a discrepancy pertaining to equity. Alumni, in responding to these questions, may recall the gap between what they had and what they thought they deserved. The exact nature of equity is difficult to ascertain, even amongst theorists of equity, but for our purposes equity is defined relatively as: "[when] the actual allocation of outcomes coincides with p's beliefs about how they should be distributed" (Cook (1975: 376) cited in Michalos, 1985). Both variables are coded as dichotomies.

Table 3-10 Indicators of Subjective Discrepancies

Discrepancies	Indicators
Had-Have by Now	Program completion longer than expected
Had-Deserve	Not satisfied with supervisor on authorship issues
Had-Deserve	Financial support distribution not fair or equitable
Had-Needed	Conference expenses partially/not funded during degree program
Had-Wanted	University expectations before enrolment (see table below)
Had-Wanted	Coop program expectations before enrolment (see table below)
Had-Wanted	Research expertise expectations before enrolment (see table below)

A second discrepancy, operationalized by "length of time spent in your graduate program", suggests, following Michalos (1985), the notion of a cognitive dissonance between what one had and what one expected to have by now. That assumes, of course, that one has incorporated, to one extent or another, a measure of expectancy related to degree completion time. Finally, a third discrepancy which only superficially tapped into what the alumni had and needed, was operationalized as "resources not adequately available for TA duties".

Table 3-11 Degree Program Expectations Cluster Membership

Clusters	Cluster Description	N	%
GSES 2000-2002 Combined			
1	Missing: no expectations given	99	12.2
2	Program-related expectancies: coop program, other	48	5.9
3	Department-related expectations: supervisor and research expertise	38	4.7
4	University-related expectancies: reputation, location	627	77.2
	Total	812	100
GSES 2002 Factorial Experiment			
1	University-related expectancies: reputation, location	80	73.4
2	Program-related expectancies: coop program, other	17	15.6
3	Department-related expectations: supervisor and research expertise	6	5.5
4	Missing: no expectations given	6	5.5
	Total	109	100

A fourth type of discrepancy, shown in Table 3-11, captures, albeit imperfectly, the gap between what one had and what one wanted. This discrepancy will be operationalized through the coded open-ended questionnaire item, "why did you decide to study at Waterloo?". Answers provided by alumni, already coded in the GSES data, were submitted to a multiple correspondence analysis to regroup each reason given into meaningful clusters.

The cluster tentatively named "university-related expectations" regrouped several reasons such as reputation, and location. The cluster "program-related expectancies" regrouped alumni who considered coop programs as the reason for choosing Waterloo. The cluster "department-related expectations" regrouped supervisor and research expertise as reasons for studying at Waterloo. A fourth cluster captured those who skipped the question entirely. Admittedly, these clusters do not provide, explicitly or implicitly, a measure of what they actually thought they had. However, the discrepancy emerges in relation to satisfaction clusters elaborated earlier. If any of these variables are strongly related to the irate satisfaction cluster, for example, it might be possible to posit, implicitly, that they did not get what they wanted.

Unfortunately, the questionnaire does not supply any items that might help us construct discrepancies related to what one needs (in any real depth), or discrepancies related to the best one ever had, what one expected to have, or to what others had. The latter, social comparison discrepancy, in particular, is unfortunate because it explains most of the variance pertaining to satisfaction after the had-want aspiration discrepancy (Michalos, 1985: 390). Iverson (1991) 's "relative model", will alleviate this impasse. This modelling strategy will be discussed in chapter five.

3.3.2.4 Objective Discrepancies

The impetus for including objective discrepancies in our model stems, like Michalos, from our basic realist ontological assumption about the nature of graduate experience: "[…] there is a world relatively independent of this or that person, containing things with more or

less objectively measurable properties, which are more or less objectively comparable" (Michalos, 1985). It suggests that the evaluation of one's graduate experience is not an entirely subjective affair; it is partly determined by the organizational context within which the alumni had to contend with during his or her stay. According to MDT, these objective differential aspects of life within an organization are likely to have a negative impact on the alumni's evaluation of satisfaction with their degree program.

Table 3-12 Indicators of Objective Discrepancies

Discrepancies	Indicators
Had-Needed	Conference expenses partially/not funded during degree program
Had-Future	Reported loans at graduation
Had-Future	Unemployed at graduation
Had-Have by Now	Normal program duration (graduate calendar)
Had-Best Ever Had	Alumni who was registered as a visa student
Had-Best Ever Had	Canadian students from other universities (not an UW alumni before enrolment)

Four objective dimensions were formed by assembling what, following MDT concepts developed in chapter two, may be conceptualized as potential candidates for objective, salient, experiential discrepancies. Table 3-12 shows the linkages between MDT objective discrepancies and their respective indicators. First, the discrepancy "had-needed to have" is captured, albeit minimally, by monetary needs encountered for any students wishing to attend/participate in/present to research conferences. The desire to present one's research coupled with the objective reality of travel costs qualifies as a real source of frustration likely to be salient in the alumnus' mind. Second, the discrepancy had-expected to have in the future" would be severely magnified if the alumnus is unable to capitalize on his or her

investment. The reality of outstanding student loans and/or being unemployed at graduation will likely thwart one's anticipated outlook for the future. Third, the discrepancy "had-should have by now" is likely to be salient when students exceed institutional norms with respect to time-to-completion; these norms, enforced by the beloved extension forms, are a constant reminder as to where the student should be in terms of progress. Should the reality of time-to-completion delays be attributable to factors other than the students themselves, graduation day may mean an unappealing retrospective glance at where they should have been by now. Fourth, the gap between what they had and the best they ever had in terms of graduate experience might be most salient amongst international students, and perhaps students whose previous degree was not from Waterloo.

3.3.2.5 Controls for Item Nonresponse

Item nonresponse to conditioners and discrepancies raises the possibility that missing cases will be related to the dependent variables of the satisfaction model (see section 2.1.3). It is conceivable, for example, that missingness on socio-demographics such as sex or visible minority is related to dissatisfaction clusters; item nonresponse is the strategy to cope with threats to anonymity. If an increase in response rate is associated with increases in item nonresponse, the conditioning variable "visible minority" will not show any nonignorable nonresponse; it will show, however, nonignorability in item nonresponse. In addition, missing cases are treated, in logistic regression, listwise. Deletion of entire alumni records based on a few missing cases is tantamount to reducing the overall response rate of the sample. Sample truncation, however, is likely to be nonignorable as well. While data imputation is a possible option, it would have foiled our attempt to detect bias on item

nonresponse. Consequently, item nonresponse needs to be accounted for on both conceptual and statistical grounds.

The issue, however, is that accounting for missingness on each and every variable would not only multiply the number of variables severalfolds (assuming dummy coding missingness), it would render the statistical model unstable. A compromise was reached by following Cohen & Cohen (1983)'s recommendation to submit all dummy codes to a factorial analysis[46]. The result is to arrive at the lowest amount of clusters describing item missingness, and yet remain conceptually intelligible.

Table 3-13 Conditioner and Discrepancies Missingness Cluster Membership

Clusters	Cluster Description	N	%
GSES 2000-2002 Combined			
1	No missingness	453	55.8
2	Missingness on discrepancies	327	40.3
3	Missingness on socio-demographics	32	3.9
	Total	812	100
GSES 2002 Factorial Experiment			
1	Missingness on unemployment	15	13.8
2	No missingness	65	59.6
3	Missingness on remaining discrepancies	29	26.6
	Total	109	100
GSES 2000-02 & NGS 1995 Combined			
1	Missingness on socio-demographics	1142	12.0
2	No missingness	8223	86.7
3	Missingness on discrepancies	114	1.2
	Total	9479	100

[46] Following Cohen and Cohen (1983), missingness of less than 5%-10% of cases was ignored provided that no significant differences (alpha = 0.5) were found between dummy codes and satisfaction clusters.

Table 3-13 presents the results of the multiple correspondence analysis on missing dummy codes. The separation of dummy codes into distinct cluster membership is far from perfect. However, judging by the test-value which quantifies how each dummy codes differentiates itself in relation to clusters, it is possible to ascertain two distinct groupings. Generally speaking, there will be one cluster regrouping missingness on conditioners (e.g.: married and children for the GSES), and another regrouping missingness on discrepancies. Variables for which dummy codes were analyzed had their missing cases subsequently mean imputed. The "no missingness" cluster became the reference category.

3.3.3 Data Analysis

As we alluded in the introduction, nonresponse bias analyses do not follow the traditional methodological and analytical aims commonly found in social sciences. Instead of orienting research towards the explaining of variations in outcomes variables (dependant) by predictors (independents), the intent is to focus on the effects of nonresponse error on the descriptive and analytical statistics used to describe or account for alumni satisfaction. Since these statistics are performed on predictors of satisfaction that are hypothesized as most topically salient, and as a result most likely to generate nonresponse error, one is likely to encounter variations in means, proportions, and regression coefficients with a change in response rate. In effect, predictors of satisfaction are the search for, and the expression of, nonresponse error. This final section of the quantitative component of nonresponse error has two objectives. First, it will establish the basic analytical and modelling strategies required to ensure the assessment of nonignorability. Second, strategies to bolster statistical validity will be discussed.

3.3.3.1 Analytical Strategies

By definition, predictors of satisfaction that are not related to the causes of nonresponse, and conversely, predictors of nonresponse that are not related to satisfaction will not cause nonresponse error. The previous sections elaborated a series of predictors of satisfaction that are considered generative of topic saliency behaviour. Unfortunately, not all these variables can be assessed as causes of nonresponse. Only a select few are included in student records. As a first step, these variables, mostly enrolment variables, will be assessed for their representativeness against alumni population values. Further tests will conducted to ascertain whether enrolment variables exhibiting representational bias are also indirectly related with other strong predictors of satisfaction. The second step is to construct logistic models for each very satisfied and very dissatisfied cluster using all variables described earlier[47]. A concomitant relationship between enrolment variables exhibiting representational bias and any of the two satisfaction clusters will establish that nonresponse is missing at random. Bias estimates on enrolment variables can be calculated using population means and proportions.

[47] When the nominal variable under study has several categories, there is a strong case to be made against the use of separate logistic regression models, and to opt instead for a single multinomial regression model. Given identical variables (e.g. coding, reference category, etc), both regression models will produce similar results; multinomial regression, however, will produce better estimations of standard errors, and of goodness of fit measures (Simonoff, 2003: 429-430). Our decision to use separate logistic models rested on conceptual grounds. We wished to assess the degree to which a variable differentiated "being very satisfied" or "being very dissatisfied", when compared to "all other" satisfaction levels. In a multinomial model, one would normally assess the degree to which a variable differentiates a given satisfaction level from each other individual level. Admittedly, the variable of interest could have been recoded into three categories: "very satisfied", "very dissatisfied", and "all others" as the reference category, and submitted to a single multinomial model. Since the same predictors were used for both "very satisfied" and "very dissatisfied" logistic models, this particular multinomial model would have been more efficient and yielded all the advantages enumerated above.

In addition, these same logistic models will be tested, sample size permitting, under varying response rate conditions. Logistic models testing the effects of student characteristics on nonresponse error will be compared with the NGS dataset; effects of departmental characteristics on nonresponse error will be tested on departmental response rate variations; the effects of survey protocol on nonresponse error will be tested against factorial group response rate variations. In all cases, a formal assessment of nonignorability can be devised using Rubin's definition of nonignorability. According to Rubin nonresponse is nonignorable when "respondent and nonrespondent with exactly the same values of variables observed for both have systematically different values of variables missing for the nonrespondent" (Rubin, 1987: 202). To that end, two types of hierarchical analytical models[48] are required.

First, a contextual analysis will determine whether a change in response rate is accompanied by a change in the probability of being satisfied, or of being dissatisfied, with one's degree program. Thus, response rate variations are the main effect that determines the probability of a given alumnus being very satisfied or very dissatisfied, once all other determinants of satisfaction have been controlled for. To paraphrase Raudenbush (2002: 141), a contextual effect is the expected difference in the satisfaction probabilities of two alumni who are identical on all subjective, objective and program-related variables, but who belong to different response rate groups. This first analytical model will establish that nonresponse is missing at random, and potentially nonignorable if can be demonstrated that satisfaction was the determinant of survey response behaviour. It will certainly establish that

[48] For a literature review on hierarchical analyses, see (D. A. Hoffmann, 1997; S. W. Raudenbush & Bryk, 2002; Van den Eeden & Hüttner, 1982)

reports relying on descriptive statistics will remain biased unless the contextual bias can be controlled for.

However, contextual analyses only tap into one of two types of nonresponse error. It also possible that a change in response rate will mediate or moderate the initial impact between a given predictor and one of the satisfaction clusters. This type of effect requires a slopes-as-outcomes model (S. W. Raudenbush & Bryk, 2002), or the mediational model (D. Hoffmann & Gavin, 1998). Like the contextual model, logistic models are built using all subjective, objective and program-related variables; unlike contextual models, response rates act as the main determinant of coefficient slope variations. This second analytical model will establish that nonresponse is missing at random to the extent that the cause of nonresponse specifies the relationship between a predictor of satisfaction and a given satisfaction cluster. Unless the cause of nonresponse can be determined, nonresponse is nonignorable. The form resistant hypothesis, which stipulates that the relationship between variables is resistant to response rate variations, will not hold.

3.3.3.2 Nonresponse Error Inferences and Statistical Validity

Evidence of nonresponse error, as observed through the variations in means, proportions and logistic coefficients across variations in response rates, is severely weakened by issues of statistical validity. Variability in these statistics may not be a reflection of true changes in the response pool as the response rate increases, but rather an artefact of the significance testing strategy itself. Two types of threats to validity must be dealt with. First, as illustrated in Figure 3-3, type I errors occur when significant differences are found

amongst our significance test results (e.g. chisquare, t-tests, etc), and yet, no real differences occur in sample representativeness despite changes in response rates. Type II errors occur when significance tests report no difference when, in fact, real differences in sample representativeness do occur as response rates changes. If we follow convention, the probabilities of obtaining a correct conclusion are set at $\beta = 0.2$ for type II errors and $\alpha = 0.05$ for type I errors (J. Cohen & Cohen, 1983). This implies that type I errors are four times more important than type II errors (Lipsey, 1998). There is no doubt that type I errors are problematic: incorrectly asserting nonresponse error will, as a consequence of initiating compensation strategies, introduce errors on an otherwise systematic-error-free distribution. Type II errors are equally problematic if one wishes to infer results to populations other than the one under study, or infer error-free results from collated samples. A closer look at how each type of error will emerge in our study is needed.

Figure 3-3 Type I and Type II Errors in Significance Testing

	Changes in sample representativeness with changes in response rates	
	Real differences	No differences
Significant differences found (reject H_0)	Correct conclusion Probability = 1 - β (power)	Type I error Probability = α
Statistical test results		
No significant differences found (accept H_0)	Type II Error Probability = β	Correct conclusion Probability = 1 - α

Adapted from (Lipsey, 1998: 41)

Type I errors, that is to say asserting nonresponse error when there is in actuality none, are likely to occur because of the quasi-exploratory nature of our study. Despite the stipulation of a limited set of theoretical hypotheses, the elaboration of these through Michalos' multiple discrepancies model increased severalfold the number of operational hypotheses to be tested. Since a model will be tested for each very satisfied and very dissatisfied cluster, the number of significance tests as now reached excessive proportions. The analytical strategy is in danger of high investigationwise error rates. The probability of finding at least one significant variation in the means, proportions or logistic coefficients is quite high[49]. To minimize type I errors, two remedies proposed by (J. Cohen & Cohen, 1983: 166-176) will be utilized. First, a "protected t" procedure will diminish the investigationwise error rate by using sets of predictors as the determinants of error, not the number of predictors themselves. The significance of the predictor set will determine whether the significance tests within each set are to be considered significant or not. Changes in the deviance parameter, along with the changes in degrees of freedom, will determine whether a given set of predictors made a significant contribution to the model. Second, Bonferroni correction will be utilized on the array of tests for means and proportions variations. The correction divides the desired type I error rate (alpha) by the number of comparisons performed. Admittedly, this conservative strategy may lead to further type II errors; it does immunize, however, against unwarranted exuberance.

Type II errors, the inability to detect real changes in the composition of the sample for changes in response rates, are likely to occur if the sample size is too small, the type I

[49] In fact, for twenty significance tests conducted with a type I error set 0.05, the probability of finding one random significant test is 100%. See (J. Cohen & Cohen, 1983)

error rate is too small, or if the magnitude of the nonresponse error to be detected is too small (J. Cohen & Cohen, 1983; Lipsey, 1998). Equation 1 below illustrates their interrelationships:

[Equation 1]: $\chi_c^2 \leq \frac{(\beta_1 - \beta_2)^2}{(s.e._1)^2 + (s.e._2)^2}$ see (Allison, 1999)

The statistical significance of the test depends ultimately on whether the results of the equation is greater than or equal to chisquare critical χ_c^2 (alpha with one degree of freedom). The ability to detect significance depends on the relative nonresponse effect on the logistic coefficient $(\beta_1 - \beta_2)^2$ for a given variation in response rate. Sample size, which dictates, in part, the sampling error of each logistic coefficient (i.e. $(s.e._1)^2$ and $(s.e._2)^2$) will also determine whether differences are detectable or not. The attempt to detect variations at low response rates (e.g. low sample size) will be foiled because of the large sampling error in relation to the net effect of nonresponse on logistic coefficients. Conversely, small relative changes in response rates are likely to produce small differences in nonresponse error. Worse still, large numbers of predictors in a given analytical model will segment the variance in ever-smaller portions diminishing as a result the power of each Wald significance test. In all cases, the ability to detect real variations in sample representativeness is limited by the power of the significance test. Type II errors are likely to be present in situations such as our use of factorial experiment data, where the sample is quite small, and the relative variation in response rate is restricted to a small range. In that particular case, the type I error rate was

loosened from 0.05 to 0.1. In all other situations, collating three years worth of GSES data will help in increasing sample size, and render analyses more sensitive.

3.4 Qualitative Approach

In addition to quantitative data sets, the assessment of bias will draw from a series of fieldwork interviews. The purpose is to collect data on the subjective component of topic saliency. Traditionally, these data are collected during household survey fieldwork or through the researcher's attribution of saliency codes on specific population groups. Our approach has the advantage of widening the scope of inquiry by observing respondent's behaviour as they fill out the survey. Such an approach allows a better assessment of the thinking pattern, heuristic or salient, in relation to the survey material. This section will begin by describing the alumni sample utilized to recruit participants; it will be followed by a description of the interview process, as well as ethical and validity considerations. The section concludes with data analysis strategies.

3.4.1 Alumni Sample

Participants for the cognitive interviews were recruited from the population of recent alumni who graduated during the fall 2003 convocation, and whose home address was Kitchener-Waterloo area. The choice of these particular students, as opposed to ABD students still enrolled at Waterloo rested on two criteria. First, the exit survey contains questions that pertain to graduate students who have completed all requirements of their program. While ABD students currently enrolled at Waterloo may be able to answer most of the survey questions, their answers may reflect their current preoccupations with the program

as opposed to a net reflection of their entire program. It is important for this study to assess whether net satisfaction or a particular dimension of the program was most salient to them. ABD students may not be able to provide this information. Second, ABD students are still under the purview of the University structure. Their on-going position under the authority of the University may not allow the distance needed to candidly answer to a satisfaction survey. ABD students may unduly shy away from answering or may edit their answers altogether. The recruitment of recent alumni would remove both of these sources of measurement error.

A sample of twenty volunteers was to be recruited for the cognitive interviews. The number of interviews was a compromise between available resources (time, money), analytical exigencies, and probability of response. The intent is not representativeness per se, nor is it to conduct comparisons between categories of respondents. The objective is to collect as many and as diversified experiences as possible. Personal interviews, however, are costly in terms of interview time, transcription time, and participant stipend. Twenty interviews is the maximum the resource budget could afford. Also, in the past, an average of 30% of graduate students (approx 120 students) had a home address in the K-W area. Using the Graduate Student Exit Survey response rate of 35% as the worst-case scenario (approx. 40 students), twenty interviews seemed to be an achievable goal.

3.4.2 Data Collection Strategy

All alumni who graduated during the fall convocation of 2003 and who currently lived in the K-W area were sent at their home address a recruitment letter outlining the purpose of the study. The letter briefly described the research project, the kinds of questions

asked, and contact information should they wish to participate. A compensation of $20 was offered; should the participant, for any reasons whatsoever, choose to terminate the interview, he or she would be compensated on a pro-rated basis. No compensation was offered for any travel expenses incurred. No follow-ups were conducted. Participants were to make contact at the researcher's phone number or e-mail address. Interviews were conducted on campus at the department of sociology.

3.4.3 Interview Protocol

Cognitive interviews were broken down into two survey response processes:

Cover Letter. The participants were presented with an envelope containing a cover letter and the Graduate Student Exit Survey questionnaire. Observational notes on how the participants utilized the survey material were taken; specifically, notes were taken on whether participants scanned the questionnaire prior to reading the cover letter. Retrospective probes administered after reading the cover letter assessed the level of recall and comprehension of the cover letter generally but also key terminology such as "experience at Waterloo". Finally, participants were asked to recall their experience at Waterloo in an attempt to assess the breadth and scope of what was recalled, and to assess whether it matched with the survey topic.

Questionnaire. The participants were asked to "think aloud" (to describe what they were thinking) as they read each survey question and attempted to decide on a satisfaction score. Observational notes were taken on question items that seemed a burden, more salient (participant was more talkative), or threatening. At points, concurrent probes were used to

inquire further on the source of the expectations (if any) that were used during their determination of their satisfaction score. At the end of the interview, a series of retrospective questions were asked to ascertain if the survey proved to be at odds with the cover letter message, if it was lengthier than expected, and if the participant felt comfortable in entrusting these answers to the GSO.

3.4.4 Ethical Considerations

Great care was taken to ensure a voluntary participation based on informed consent. Graduate students were informed of the nature of the interview, and the kind of questions they were liable to answer. Participants were asked permission for tape-recording their answers. Prior to starting the interview, they were under no obligation to answer any questions, and they were free to terminate the interview without censure.

Interview participants were also assured anonymity. A numbered recruitment mailing list was devised to serve as an index to all material pertaining to the participant. No names, addresses, or phone numbers were transcribed on any of the material collected (including questionnaire, tapes, logs). The data were not distributed to anyone beyond the supervisory committee. No data were distributed to the Graduate Studies Office. This fact was clearly stated in the recruitment letter, the consent forms, and the feedback letter. The release of findings did not include any names, department or faculty references. Upon completion of the study, the mailing list was destroyed. All interview material was kept in a secure location at the researcher's residence.

3.4.5 Validity and Reliability

The use of cognitive interviews within the context of a study on nonresponse bias does impose some limitations on the interpretation of the data. Ordinarily, exit surveys are self-administered, not face-to-face interviews, and are filled out at home with all the contextual implications that may follow (timing of the request, the participant state of mind, etc). During a cognitive interview, the context in which the respondent answers questions on the survey is artificial. Cognitive interviews are low in "experimental realism" and low in "mundane realism" (Nachmias-Frankfort & Nachmias, 1996). While the experience of the survey may look real, the cognitive interview imposes an artificial amount of mental effort by asking the participant to think aloud, all the while being interrupted by concurrent probes. (Ericson, 2003)'s research on the validity of verbal protocol provides some evidence that verbalized data through the use of concurrent probes is consistent with the participants' thoughts; think-aloud reports are valid despite the artificiality of having to voice one's thoughts. The issue of realism might be more an aspect of context than of thinking-aloud in and of itself.

There are however, other limitations that are more serious and do threaten the validity of the study. One of the main methodological issues plaguing face-to-face interviews and cognitive interviews alike is the bias introduced by the presence of the experimenter. One of the subtle difficulties in cognitive interviews is to avoid impressing on the participants any conscious or subconscious expectations of findings. Upon a request from the participant to clarify one's retrospective or concurrent probe, it would be easy to communicate expectations. A detailed script has been written to guide the researcher through the cognitive

interview. Following the script reliably and keeping clarification short to the point of suspending the probe may help alleviate any bias in the verbal protocol of participants.

3.4.6 Data Analysis

Data to be drawn from the cognitive interviews came from several sources. First, verbal statements were audiotaped and then digitized into MP3 format. These statements were extracted through repeated listening to the tapes themselves. While the tapes themselves were not transcribed into text form, statements, when appropriate, were extracted and recorded. Extracted statements were written verbatim; pauses were indicated as (pause), three points (…) indicated the exclusion of some words, and square brackets indicated the intervention of the interviewer. Second, observational notes were taken on whether the participants looked at the questionnaire prior to reading the cover letter, and whether they incorrectly skipped any section or questions while filling out the questionnaire. Third, participants were instructed to highlight on the questionnaire any questions that presented any difficulties in terms of comprehension, recall, mapping difficulties or question threats. Fourth, data from the questionnaire were collected to match, when necessary, verbal statements provided by the participants.

Verbal statements were analyzed using two concomitant strategies. The first strategy was to proceed by identifying and cataloguing verbal statements in terms of topic saliency, cognitive burden, and question threat[50]. A salient feature of the questionnaire or the cover letter was identified by the breadth and scope of the statements in terms of the degree of

[50] Forsyth et al. (1999), basing their survey appraisal on Tourangeau (2000)'s response model, elaborated a detailed list of indicators susceptible to generating measurement error. These indicators will form the framework from which cognitive burden will be assessed.

details, the number of events and the intensity of the experience depicted. Cognitive burden, that is to say the level of cognitive effort needed to accomplish a given task, was identified indirectly through its effect on response behaviour. During the verbalization of the participants' thoughts, any explicit miscomprehension of the question, difficulties in remembering the event in question, difficulty in mapping one's answer to the Likert scale were logged as an indication of cognitive burden. Question threat was measured as an explicit hesitance on the part of the participant. Delays in answering (hovering over Likert scale) was probed further as to whether the participant felt the question to be sensitive. Any answers in the affirmative were logged as an instance of question threat.

The second strategy was to link verbal statements identified as exemplars of saliency, cognitive burden or question threats into theoretical frameworks. The basic analytical thrust followed the illustrative method (Neuman, 2004: 329-330), whereby key statements are extracted to illustrate and to provide evidence for the theoretical concepts elaborated for each phase of the survey response process: Petty and Cacioppo (1986)'s persuasion model and Groves (2000)'s leverage-saliency theory for the cover letter, Foddy (1993)'s sense-making interactional approach for measurement errors. No further analyses were conducted on the verbal statements themselves. Typical of cognitive interviews in survey research, verbal statements were taken at face value, that is to say, as depicting real appraisals of the cover letter and/or the questionnaire[51]; no attempts were made to code verbal statements at the word level such as protocol analysis (Ericson & Simon, 1993). Moreover, a case study approach (Yin, 2003) whereby verbal statements would be grouped in order to perform

[51] See for example Billings-Gagliardi et al. (2004), Low (1999), Nolin (1996), Sudman and Bradburn (1996)

analytical comparisons using salient variables found in the previous chapters, was abandoned; the number of interview cases was too limited to perform intelligible comparisons.

3.5 Conclusion

The sequential mixed method research design associated with the study of nonignorable nonresponse to satisfaction surveys was constructed from a series of compromises. The Graduate Student Exit Survey 2000-2002 was selected for pragmatic reasons. It was readily available, was supported by a rich source of frame variables, was compatible with Statistics Canada NGS survey, and dealt with the same population group as cognitive interviews. Unfortunately, the design and implementation of the GSES, and the restriction of the public version of the NGS data set, both greatly hampers the kind of analyses to be performed. Chapter Four will begin data analysis with a closer look at the socio-demographic predictors of nonignorable nonresponse.

Chapter 4
Influences of Student Characteristics on Nonresponse Error

4.1 Introduction

The concept of ignorability is of immediate concern for university administrators who wish to use alumni survey data to better understand their population. It presupposes that, regardless of the response rate, the survey can accurately reflect the composition and the sentiments of the student population. Yet, based on the literature review and general hypotheses developed in chapter two, there are sufficient reasons to question the presumption of ignorability. Satisfaction surveys might be particularly prone to topic intensity effects: alumni who are most satisfied or most dissatisfied with their graduate experience are much more likely to base their cooperation with a survey request on their satisfaction level. The remainder, presumably satisfied, are not likely to find the survey topically salient, and as a result, less likely to cooperate. In all cases, satisfaction is an important reason for choosing to answer or not. In addition, these topic saliency effects are likely to be located in certain socio-demographic and program-related variables. Gender, race, and length of stay, to name only a few, are likely to introduce a satisfaction bias. The net effect would be nonignorable nonresponse.

Because of topic saliency effects, it is likely that respondents are not representative of the alumni population. This will have a serious impact on the reporting of alumni satisfaction. It is expected that the probabilities of being very satisfied and being very dissatisfied are likely to diminish with an increase in response rate. This change in

probabilities will not only alter the values on univariate statistics (e.g. means and proportions), but will also affect bivariate and multivariate statistics as well. Assuming a categorical predictor of satisfaction (e.g. dummy coded) within the context of a logistic regression model, it is likely that the odds of being very satisfied or very dissatisfied on a given category may not remain constant in relation to the odds in the reference category. Consequently, a change in response rate will alter the odds ratio of that categorical predictor on being very satisfied or on being very dissatisfied. In sum, nonignorable nonresponse implies that a change in response rate would alter our understanding of alumni satisfaction.

Our objective in this chapter is to compare the University of Waterloo's Graduate Student Exit Survey (GSES) results against the National Graduates Survey (NGS) collected by Statistics Canada. Considering the potentially damaging effects of topic saliency on survey results, two issues will be addressed. On the assumption that the University of Waterloo is, on average, not significantly at variance from other Canadian universities in relation to three experiential clusters, overall program satisfaction, satisfaction with supervisor, and satisfaction with department, we need to ascertain whether the responses collected from the combined convocation 2000-02 are similar to what one may expect of Canadian alumni generally. Secondly, would our understanding of alumni change if topical saliency effects were diminished by a presumably better survey? To test these two issues, this study will begin with an assessment of representativeness on common enrolment variables. From these results, the study will formally test whether GSES results are ignorable or nonignorable with respect to student characteristics.

4.2 Methodological Considerations

This study on the impact of topically salient alumni characteristics on nonresponse error is seriously limited by the external database used to compare satisfaction scores. Comparison with an external database was required because the GSES implementation did not see fit to date-stamp surveys received, so as to distinguish early from late respondents, could not differentiate between non-contacts, and refusals because of anonymity concerns, and did not conduct any additional post-survey research on nonrespondents[52]. As previously elaborated in chapter three, the National Graduate Survey (NGS) is a good candidate for this comparison, but the inability to use SRDC restricted data sets forced our hand in using only the public version of the dataset. The public version, although still useful, opens up unbridgeable problems of coverage. There is simply no possibility of isolating University of Waterloo students, or of isolating Ontario students (PhD and Master's) for that matter. Consequently, there is no ability to discern whether any significant findings found are due to the effect of a better survey, or simply due to population differences. For this study, the comparison will be made against the entire Canadian alumni population. Differences in

[52] The inability to discern between various types of nonresponse, namely non-contacts from refusals, may confound results to the extent that types of nonresponse may be related to variables in our model. Consequently, we are limited in our ability to attribute topic saliency to any nonignorable findings. There is a greater probability, for example, that alumni who hold multiple degrees from the University of Waterloo stand a greater chance of being contacted, simply because they have provided/maintained accurate contact addresses with the Graduate Studies Office. In turn, these same alumni are also more likely to refuse a satisfaction survey request because they are less likely to encode discrepancies, or to perceive large deviations from expectations during their degree program. As a result, the proportion of refusals to non-contact is expected to be higher amongst returning alumni. This unaccounted proportional difference between types of nonresponse would most likely introduce coefficient variations in the variable "not an UW alumni before enrolment", but also potentially in other variables in our models, including the intercept. Thus, the reader should bear in mind that the presence of significant variations in model coefficients and intercepts might be the product of shifting proportions of refusals to non-contacts, and not necessarily the outcome of topic saliency.

survey periods (i.e. NGS two years after graduation) and survey mode (i.e. NGS phone surveys) further cloud comparison with GSES data.

In addition to these challenges, the NGS dataset was lacking many of the variables found in the GSES survey, and so we could not reproduce our original GSES 2000-02 overall satisfaction clusters tabulated in Table 3-5. In an attempt to emulate overall alumni satisfaction, the GSES and NGS datasets were combined, and four satisfaction items, common to both databases, "satisfaction with teaching quality", "satisfaction with computing equipment", "net satisfaction with university", and "net satisfaction with choice of degree program" were submitted to a multiple correspondence analysis. The analysis yielded four satisfaction clusters (see Table 3-6); each case in the combined databases was assigned membership in one of the four satisfaction clusters.

Table 4-1 GSES and GSES-NGS Satisfaction Clusters Cross-tabulation

GSES-NGS Overall Satisfaction Clusters	GSES Overall Satisfaction Clusters							
	1	2 (VS)	3	4 (VD)	5	6	7	Total
1 (VS)	19.7%	31.8%	12.1%	0.6%	27.4%	0.0%	8.3%	100.0%
2	12.6%	19.0%	27.7%	3.5%	15.1%	2.5%	19.6%	100.0%
3 (VD)	13.2%	7.9%	26.3%	25.0%	13.2%	3.9%	10.5%	100.0%
4	14.7%	6.3%	13.7%	17.9%	28.4%	3.2%	15.8%	100.0%

Chisquare = 151.8, $p < 0.001$, N=812. Cross-tabulation reflects GSES cases only.

To assess whether there is continuity and validity in the concept of "overall alumni satisfaction" linking the original GSES seven satisfaction clusters (table 3-5) and the four GSES-NGS satisfaction clusters (table 3-6), the two cluster sets were cross-tabulated, and assessed for their symmetry. Table 4-1 reveals that the GSES-NGS "very satisfied" cluster maps, albeit imperfectly, with the GSES "very satisfied" cluster; some of the GSES-NGS "very satisfied" cases, however, map incorrectly to GSES cluster five. The GSES-NGS "very

dissatisfied" cluster spans across the GSES "very dissatisfied" cluster and GSES cluster three. In sum, the GSES-NGS clusters have limited validity and continuity when compared with our original seven GSES satisfaction clusters. Because the NGS-GSES "very satisfied" and "very dissatisfied" clusters are muddled, they may not bring forth the kind of topic intensity effects generative of nonresponse error. Thus the inability to detect any topic saliency effects might be confounded by the muddled salience of the GSES-NGS satisfaction clusters.

In terms of similarities on alumni characteristics, alumni's gender, visible minority status, science program, loans, and deviation from normal program duration were immediately usable without translation. The variable employment was coded as "employed" for any reports of employment after graduation. The variables marital status and dependents, although present in the NGS, could not be used due to a different time reference. The GSES asks to report marital status one year prior to graduation; the NGS reports marital status two years after graduation. Finally, and perhaps more problematic, international student registration was not coded in the NGS survey. It was derived from the country of origin prior to being admitted to university. Since the Master's of Accounting program is unique to the University of Waterloo, it was recoded as master's no research; the NGS "university diploma or certificate above bachelor level" were assigned to the master's no research category. Science degree was derived the NGS first field of study major groups, which is itself a recoded variable that merges "university student field of study codes (USIS) and community college and trade-vocational field of study codes" (Statistics Canada, 1995a: 10)

Data analysis strategies will follow the guidelines already elaborated in chapter three. We need to ascertain whether survey respondents are any different from the alumni population of 2000-2002 on program-related and socio-demographic characteristics. Second, we must ascertain if any of these characteristics are also related to the satisfaction model proposed in chapter three. Third, we need to know if our understanding of student experience (according to the satisfaction model) would change if a higher response rate were obtained. A comparative analysis of alumni satisfaction using GSES and NGS data sets will determine the contextual and moderational effect of response rates on our logistic model of student satisfaction.

To protect against type I errors, logistic models were built around the protected-t statistical technique. Sets of predictors were entered in successive blocks. For each entry, two pieces of information are reported: First, the block chi-square indicates whether the block made a significant contribution to the model beyond what was already included. Significance was calculated from the differences in model deviances and degrees of freedom. Second, the increased McFadden r-square analog provides an overall measure of the contribution to the explained "variance" to the model. Following the protect-t logic, all significant predictors within a given block are to be discarded if the block itself did not make a significant contribution to the model.

4.3 Nonresponse across Student Characteristics

This section will assess the degree to which survey respondents are representative of the graduate alumni population of 2000-2002 in terms of socio-demographic and enrolment

variables provided by the Graduate Student Office. Significant differences between population and survey results on a given set of variables would signal that cooperation behaviour was not random; directly or indirectly, enrolment variables, such as alumni gender or degree program, would be related with survey cooperation. Representational bias might be indicative of a direct effect of the enrolment variables on cooperation behaviour. It may also be indicative of an interaction with other topically salient variables.

The variables provided by the Graduate Student Office were limited in number but do provide an excellent starting point to assess the representativeness of the survey. The enrolment variables to be compared are alumni gender, visa status, degree program, department, and graduation status. The variable "degree program" was collapsed into two dummy codes: a one-year professional degree (MAcc) and a doctoral degree variable; master's degree program is the reference category. Departments were recoded into whether or not they were part of a natural science-related department (e.g. mathematics, chemistry, etc). Comparison will be performed using the chi square test of difference (Healey, 1996; Rodeghier, 1996). The null hypothesis will be rejected with a statistical error of 5%. Failure to reject the null hypothesis of no difference implies that the survey respondents are deemed proportionally similar to the population.

Table 4-2 presents the chi square results for the six enrolment variables. Women alumni were significantly overrepresented in the combined 2000-02 GSES response pool. Although significance levels varied between convocation periods, women alumni seem to respond more than men as a general rule. This finding corroborates other alumni studies referenced in chapter two, and may lead to nonignorable nonresponse.

Table 4-2 Sample Representativeness

Predictors	Alumni Population N	Alumni Population %	Survey Respondents N	Survey Respondents %	Bias	Sig.
Women alumni						
Convocation 2000	656	43.1%	254	48.8%	5.7%†	0.068
Convocation 2001	717	42.1%	262	42.7%	0.6%	0.837
Convocation 2002	734	39.1%	285	43.2%	4.1%	0.160
Total	2107	41.4%	801	44.8%	3.4%*	0.049
Student with visa status						
Convocation 2000	656	9.6%	244	12.7%	3.1%	0.100
Convocation 2001	717	7.9%	256	10.9%	3.0%†	0.077
Convocation 2002	734	9.3%	275	14.5%	5.2%***	0.003
Total	2107	8.9%	775	12.8%	3.9%***	0.000
Master's of accounting alumni						
Convocation 2000	656	14.8%	257	12.8%	-2.0%	0.379
Convocation 2001	717	13.8%	264	14.0%	0.2%	0.922
Convocation 2002	734	15.4%	283	11.7%	-3.7%†	0.082
Total	2107	14.7%	804	12.8%	-1.9%	0.137
Doctoral alumni						
Convocation 2000	656	20.9%	257	24.1%	3.2%	0.201
Convocation 2001	717	17.3%	264	17.0%	-0.3%	0.915
Convocation 2002	734	14.4%	283	17.0%	2.6%	0.228
Total	2107	17.4%	804	19.3%	1.9%	0.164
Withdrawal/Failed to register						
Convocation 2000	715	8.3%	257	1.9%	-6.4%***	0.000
Convocation 2001	807	11.2%	266	0.8%	-10.4%***	0.000
Convocation 2002						
Total	1522	9.8%	523	1.3%	-8.5%***	0.000
Science- related department						
Convocation 2000	656	55.5%	253	56.9%	1.4%	0.647
Convocation 2001	717	59.6%	249	60.6%	1.0%	0.726
Convocation 2002	734	58.7%	272	59.2%	0.5%	0.874
Total	2107	58.0%	774	58.9%	0.9%	0.605

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Should women respondents, unlike women nonrespondents, have found the survey topic salient for reasons of equity issues and discrimination, then we might expect

nonignorable nonresponse. If representational bias is related to the variables of interest but women respondents are similar to women nonrespondents, the survey responses will be biased but nonresponse will remain ignorable.

Visa students were also significantly overrepresented amongst the GSES survey respondents. The international student registration population stands at 8.9% for the alumni of 2000-2002 combined; the survey results show a composition of 12.8%. This overrepresentation was observed for all convocation years, although convocation of 2000 did not reach significance below 0.1 levels. These findings flow with the hypothesis that visa students may have found the survey topic salient; in comparing their Waterloo experience with their post-graduate expectations had they stayed in their home country, visa students may, simply by being admitted, be predisposed to being satisfied, and likely to reciprocate.

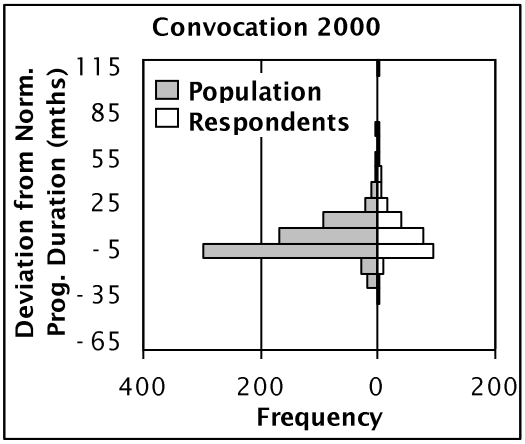
Withdrawal/failed to register students were significantly underrepresented in the GSES survey. The combined 2000-2001 alumni population was composed of 9.8% withdrawal students; the survey collected 1.3%. The alumni population of 2002 did not include withdrawal students. These findings are perhaps surprising in the dramatically low response rate for that population group but perhaps not in terms of topic saliency. The questionnaire items were primarily worded in a manner that indicated a completed program. Some satisfaction questionnaire items were explicit: “Dissertation/thesis submission”; others less so: “appropriateness of degree requirement”. The cover letter was addressed to “dear alumnus”. Moreover, there may have been a fair amount of resentment and/or social desirability problem in revealing one’s status as withdrawal. On balance, the decision to participate might have been perceived as an unwitting cost. More research is required to

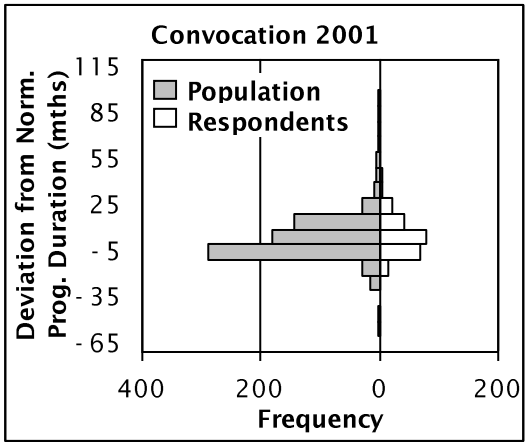
assess if these hypotheses can be integrated with the current literature on student retention, whether psychological or organizational integration models (Andres & Carpenter, 1997; Delucchi, 2003; Elliott & Healy, 2001) no significant representational biases were found amongst degree program and department related variables. However, while their combined 2000-02 total showed no significant bias, alumni registered in the Master's of Accounting program were significantly underrepresented in the convocation of 2002.

A glance at all convocation periods would seem to suggest a pattern of under-representation. Doctoral alumni response patterns seem to suggest a non-significant overrepresentation bias. In both instances, this might be an indirect effect of program duration on cooperation behaviour. As stated in chapter two, length of stay was associated with cooperation behaviour. Alumni registered in science-related departments also revealed a pattern of overrepresentation. The bias, however, never reached the 0.1 level.

Table 4-3 shows the differences in deviation from normal program duration between respondents and the alumni population for each convocation period. Deviation from normal program durations were calculated as the difference between program duration and institutional expectations of degree completion: master's of accounting was coded as 12 months, all remaining master's programs as 24 months, and doctoral degree programs as 60 months (five years). A breakdown of deviation from normal program duration by convocation periods reveals that respondents were generally those who had an extended stay at university by an average of two months, or roughly one additional term.

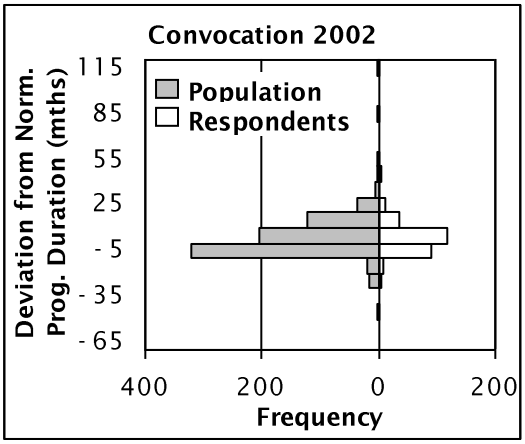
Table 4-3 Representativeness on Deviation from Normal Program Duration (Months)

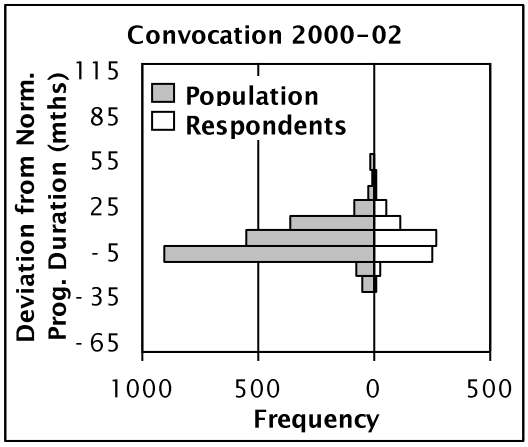




Population		Respondents	
Mean	-0.829	Mean	1.311
Median	-5	Median	-2
St. Dev	14.064	St. Dev	15.340
N	656	N	251
		Bias	2.140*

Population		Respondents	
Mean	0.227	Mean	2.386
Median	-1	Median	-1
St. Dev	14.767	St. Dev	14.420
N	717	N	278
		Bias	2.159†





Population		Respondents	
Mean	-0.951	Mean	-0.387
Median	-1	Median	-1
St. Dev	11.768	St. Dev	12.325
N	734	N	274
		Bias	0.564

Population		Respondents	
Mean	-0.512	Mean	1.019
Median	-1	Median	-1
St. Dev	13.570	St. Dev	14.052
N	2107	N	753
		Bias	1.531***

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 4-4 Frame Correlations

Predictors	1	2	3	4	5
1. Women alumni					
Population					
Respondents					
Difference in correlation (z score)					
2. Student with visa status					
Population	-0.009				
Respondents	-0.085*				
Difference in correlation (z score)	0.075†				
3. Deviation from avrg program duration (mths)					
Population	0.005	-0.138***			
Respondents	-0.004	-0.070†			
Difference in correlation (z score)	0.008	-0.069			
4. Science degree					
Population	-0.189***	0.040†	0.086***		
Respondents	-0.221***	0.168***	0.116**		
Difference in correlation (z score)	0.034	-0.129**	-0.030		
5. Master's of accounting alumni					
Population	0.074***	-0.125***	-0.126***	-0.487***	
Respondents	0.075*	-0.133***	-0.133***	-0.469***	
Difference in correlation (z score)	-0.001	0.009	0.007	-0.023	
6. Doctoral alumni					
Population	-0.051*	-0.008	-0.127***	0.140***	-0.190***
Respondents	-0.061†	0.089*	-0.082*	0.116**	-0.187***
Difference in correlation (z score)	0.011	-0.097*	-0.045	0.024	-0.003

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The overall distribution of deviations, compared between population and respondents, was equivalent. With the exception of the convocation period of 2000, the standard deviations were similar, as was the distribution skewness shown here as the relationship between mean and median. No outliers were detected; minimums and maximums (not shown) were also equivalent. There is, thus, some support for the hypothesis that university residency does generate a response differential. Those alumni who had a longer stay tended

to answer in greater numbers than those who did not. It remains to be seen if program residency is also related to satisfaction levels.

Table 4-4 provides an indication as to whether enrolment variables relate to each other with the same strengths and direction in the response pool as in they do in the population frame. Differences between correlations were calculated by z score transformations (Goudy, 1976; Snedecor & Cochran, 1989)[53]. A closer look at significant correlation differences between population and respondents reveals a bias pattern stemming from visa students: visa respondents are predominantly male doctoral alumni enrolled in science-related departments. A probable cause for this bias may lie in the nature of the enrolment of visa students. Visa students are permitted to enrol in a Canadian university only through the acquisition of a student visa. Those who are not returning to Waterloo or are not applying for permanent resident status have already moved to their countries of origin. The composition of responding visa students may be biased towards doctoral students (long program residency) who have applied for permanent residency and, still residing in Canada, are more likely to respond to the survey. All remaining correlations amongst frame variables were comparable to those found in the population.

4.4 Nonresponse Error across Student Characteristics

Analyses of representational bias on the combined 2000-02 convocations indicated that women alumni, visa-registered alumni, and alumni who completed their studies over a period of time longer than normal institutional expectations were all overrepresented in the

[53] Pearson correlations were first converted to z scores: $z = (1/2)[\ln(1+r) - \ln(1-r)]$. Test of significance was obtained by subtracting z scores ($D = z1-z2$) and by dividing the difference with their common standard error: $\text{sqrt}\ (1/(n_1-3) + 1/(n_2-3))$.

response pool. Interaction between enrolment variables revealed that visa doctoral males registered in science-related departments are particularly overrepresented. This section will assess whether these variables directly or indirectly cause nonresponse error.

Under ideal circumstances, the determination of nonresponse error on satisfaction scores is conducted through an analysis of early-late respondents, or better yet, on an analysis of post-survey interviews of nonrespondents. The advantage of either of these techniques is to permit an analysis on all variables included in the questionnaire. Since we are conducting an analysis on a secondary data set that precludes such possibilities, two complementing strategies will be followed: first, adjustment logistic models will be used to assess the bias incurred by the overrepresentation of variables found in the previous section; second, a comparison will be made with the NGS dataset.

4.4.1 GSES Nonresponse Error Adjustment Models

The primary function of adjustment models is to provide an estimate of satisfaction bias when enrolment variables found to be overrepresented in the response pool are adjusted to match population means and proportions. The idea is to build a regression model with enrolment variables included, and then to substitute respondents' means on enrolment predictors with population means[54]. This statistical adjustment gives an indication of the estimated satisfaction bias should a missing at random condition prevail. The degree to which overrepresentation on enrolment variables will have an impact on the estimated satisfaction bias depends on the magnitude of both, the representational bias, and the impact of enrolment variables on being very satisfied or being very dissatisfied. Since it is not

[54] This statistical adjustment strategy is based on approaches developed by Duncan (1968) and Goyder (1981)

expected that enrolment variables will play a major role in the explanation of satisfaction, it is not expected that their overrepresentation will introduce a substantial satisfaction bias.

To facilitate statistical adjustments, all models were constructed with predictors centered on their respective respondent grand-mean. When predictors (x) are all equivalent to their means, they all reduce to zero (e.g. $\beta\left(x-\bar{x}\right)=0$, including all frame interactions. The remaining parameter, the intercept, represents the predicted log-odds of being very satisfied or being very dissatisfied when the means are equivalent to the respondent pool. To obtain an estimation of satisfaction bias on a given cluster, the representational bias of a given enrolment variable, $\bar{x}_{pop}-\bar{x}_{resp}$, can be substituted for the respondents' mean. Table 4-6 to Table 4-11 present the logistic regression results for each very satisfied and very dissatisfied cluster on all three experiential dimensions. The bullet "●" indicates variables contained in the student records. Table 4-5 tabulates the unadjusted and adjusted probabilities.

Table 4-5 Predicted Bias from Adjustment Logistic Models

	Unadjusted			Adjusted	
		95% CI			
Dependent Variables	P(Y=1)	Lower	Upper	P(Y=1)	Bias
Overall graduate experience					
Cluster 2 (Very Satisfied)	12.85%	10.06%	16.28%	13.19%	-0.34%
Cluster 4 (Very Dissatisfied)	1.75%	0.96%	3.19%	1.82%	-0.06%
Experience with research supervisor					
Cluster 2 (Very Dissatisfied)	5.13%	3.50%	7.45%	4.86%	0.27%
Cluster 3 (Very Satisfied)	24.14%	20.80%	27.83%	24.20%	-0.05%
Experience with the department					
Cluster 3 (Very Dissatisfied)	3.59%	2.38%	5.38%	3.94%	-0.35%
Cluster 5 (Very Satisfied)	15.75%	12.83%	19.19%	15.76%	-0.01%

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

As expected, overrepresentation on women and visa-registered alumni, and those alumni who completed their studies later than expected norms, did not result in any substantial bias in any experiential clusters. At most, one can anticipate biases hovering around one third of a percentage point. The reason stems from a combination of minor representational biases, and weakly related predictors. Women alumni, for example, were overrepresented by only 3.4 percentage points; a glance at all models reveals that being a woman alumnus does not matter in relation to some experiential clusters, and when it does, the relation is moderate (i.e. odds ratio of 2.4 on being dissatisfied with overall graduate experience, odds ratio of 2.1 on being very dissatisfied with supervisory experience). The same can be said for visa-registered alumni, and deviation from average program duration. It is not that conditioners have no role to play in the determination of alumni satisfaction. Their contribution to the explained variance of satisfaction clusters, however, is modest: increases in the McFadden r-square analog range from 2.8% to 10.6%. Thus, as long as the form resistant hypothesis holds, the estimated bias on enrolment variables will remain essentially ignorable.

Part of the explanation for ignorability lies in the fact that net bias effects on satisfaction clusters sum to triviality. A closer examination of the frame interactions with overrepresented enrolment variables establishes just how complex these relationships can be. The impact of women alumni on being very dissatisfied with overall graduate experience is a case in point. The net effect is a positive odds ratio of 2.4; women alumni are two and a half times more likely to be very dissatisfied than men. And yet, not all women alumni agree: women who frequent the Graduate House (peer support) and women who have reported

being a member of a visible minority would be less likely to report being very dissatisfied overall with their degree program. Similar mixed effects can be seen on the likelihood of being very dissatisfied with supervisory experience. Perhaps nonignorability manifests itself in the details of satisfaction, but taken as a whole, enrolment variables seem to point to ignorability.

The secondary function of adjustment models is to ascertain which predictors, following Michalos' multi-discrepancies theory (MDT), are topically salient to the respondent pool. Admittedly, the number of discrepancy variables makes MDT-based models more suited for very dissatisfied clusters; nonetheless, involvement and prior-enrolment expectations (aspiration theory) should provide explanation for being very satisfied with one's degree program. As a general assessment of goodness of fit, two indicators are useful to this task. First, the McFadden r-square analog provides an indication of the predictive power of subjective, objective and program-related variables to explain the initial variance[55] in a given satisfaction cluster. A look at each model reveals a weak to moderate proportional reduction in initial variance, ranging from 11.5% to 37.5%. Second, the lambda statistic[56] provides an indication of how well the model classifies cases correctly. While the McFadden statistic gives an indication of explained variance, it does not establish whether predictors can predict cluster membership.

[55] The concept of variance in a logistic model, expressed as the initial -2LL0 likelihood, is a measure of dispersion not identical but conceptually analogous to the "index of qualitative variation" (Healey, 1996) used to describe variation within categorical variables. Contrary to the IQV, -2LL0 likelihood logs the respective proportions, does not yield an index ranging from 0 to 1, and its value is dependent upon sample size. The initial -2LL0 likelihood is calculated as D0 = 2(NY=1ln(NY=1/N) + NY=0ln(NY=0/N)) (Menard, 2002). Maximum variance or heterogeneity can be expected when the observed probabilities of being in a particular cluster P(Y=1) are equal to the probabilities of not being in that cluster P(Y≠1).

[56] See (Menard, 2002: 27-35)

Table 4-6 Very Satisfied Graduate Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-1.914***	
Subjective Discrepancies (Block 1: χ^2 = 18.710** , R^2 = 0.024)			
Program completion longer than expected	0.205	0.149	0.019
Not satisfied with superv. on authorship issues			
Financial support distribution not fair/equitable	-0.473†	-0.279	-0.029
Resources not adequately available for TA duties	-0.076	-0.137	-0.011
University expectations before enrolment	0.874***	1.112***	0.126
Coop program expectations before enrolment			
Research expert. expectations before enrolment	-0.232	0.746	0.043
Objective Discrepancies (Block 2: χ^2 = 3.766 , R^2 = 0.028)			
Conference expenses partially/not funded	0.157	0.590†	0.047
Reported loans at graduation	-0.248	-0.436*	-0.057
Reported unemployment at convocation	0.086	-0.418	-0.038
• Deviation from avrg program duration (mths)	0.009	-0.008	-0.029
• Student with visa status	0.231	-0.310	-0.027
Not an UW alumni before enrolment	0.212	-0.143	-0.018
Conditionners (Block 3: χ^2 = 61.473***, R^2 = 0.106)			
• Women alumni	-0.300	-0.322	-0.043
Visible minority status	-0.461*	-0.473†	-0.057
Married during degree program	0.040	0.401	0.051
With dependents during degree program	0.075	0.368	0.038
• Master's of accounting alumni			
• Doctoral alumni	-1.672***	-2.264***	-0.240
• Science degree	0.349†	0.210	0.027
Financial support (K $)	-0.014***	-0.013*	-0.101
Frequent visits to the Graduate House	0.357*	0.181	0.024
Full research or teaching involvement	0.326	0.520	0.043
No research or teaching involvement	-0.218	-0.728	-0.067
Item Nonresponse (Block 4: χ^2 = 5.576† , R^2 = 0.113)			
Missingness on program characteristics	-0.450*	-0.616*	-0.082
Missingness on socio- demographics	-0.511	-0.718	-0.038
Frame Interaction (Block 5: χ^2 = 42.349***, R^2 = 0.167)			
Women x Reported unemployment at convocation		-1.700*	-0.075
Women x With dependents during degree program		1.339*	0.066
Women x Financial support (K $)		0.024*	0.091
Women x No research or teaching involvement		-1.924*	-0.087
Visa x Deviation from avrg program duration (mths)		-0.070*	-0.073
Visa x No research or teaching involvement		-5.337†	-0.145
Visa x Missingness on program characteristics		1.613*	0.065
N Dur. x Reported unemployment at convocation		-0.085**	-0.114
Full Model: χ^2 = 131.874* df = 30, McFadden R^2 = 0.167, λ = 0.006**			

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Table 4-7 Very Dissatisfied Graduate Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-4.025***	
Subjective Discrepancies (Block 1: χ^2 = 93.754***, R^2 = 0.236)			
Program completion longer than expected	1.040***	1.370***	0.200
Not satisfied with superv. on authorship issues	3.032***	3.109***	0.228
Financial support distribution not fair/equitable	1.586***	1.488***	0.178
Resources not adequately available for TA duties	1.062**	0.452	0.041
University expectations before enrolment	0.278	0.155	0.020
Coop program expectations before enrolment			
Research expert. expectations before enrolment			
Objective Discrepancies (Block 2: χ^2 = 13.449* , R^2 = 0.270)			
Conference expenses partially/not funded	1.348***	0.994†	0.091
Reported loans at graduation	0.747**	0.840*	0.127
Reported unemployment at convocation	0.917**	0.768	0.080
• Deviation from avrg program duration (mths)	-0.006	-0.047**	-0.197
• Student with visa status	0.214	0.130	0.013
Not an UW alumni before enrolment	0.256	0.300	0.045
Conditionners (Block 3: χ^2 = 21.800* , R^2 = 0.325)			
• Women alumni	0.218	0.866*	0.133
Visible minority status	0.236	0.595	0.082
Married during degree program	0.128	0.274	0.040
With dependents during degree program	-0.353	-0.630	-0.074
• Master's of accounting alumni			
• Doctoral alumni	0.074	-0.477	-0.058
• Science degree	0.837*	1.642***	0.245
Financial support (K $)	0.000	-0.003	-0.027
Frequent visits to the Graduate House	0.476†	0.352	0.053
Full research or teaching involvement	0.086	-0.897	-0.085
No research or teaching involvement	-0.066	-0.317	-0.034
Item Nonresponse (Block 4: χ^2 = 1.057 , R^2 = 0.328)			
Missingness on program characteristics	-0.598†	-0.076	-0.012
Missingness on socio- demographics	0.735	1.013	0.061
Frame Interaction (Block 5: χ^2 = 18.705***, R^2 = 0.375)			
Women x Financial support distribution not fair/equitable		-1.489†	-0.089
Women x Visible minority status		-2.154*	-0.146
Women x Frequent visits to the Graduate House		-1.585*	-0.119
N Dur. x Reported unemployment at convocation		0.081*	0.125
Full Model: χ^2 = 148.764* df = 26, McFadden R^2 = 0.375, λ = 0.278***			

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Table 4-8 Very Dissatisfied Departmental Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-3.291***	
Subjective Discrepancies (Block 1: χ^2 = 22.918***, R^2 = 0.059)			
Program completion longer than expected	0.367	0.509	0.074
Not satisfied with superv. on authorship issues	1.862***	1.678***	0.123
Financial support distribution not fair/equitable	1.398***	1.055**	0.126
Resources not adequately available for TA duties	0.843*	0.344	0.031
University expectations before enrolment	0.101	0.540	0.070
Coop program expectations before enrolment	0.573	0.864	0.063
Research expert. expectations before enrolment	-0.218	0.410	0.027
Objective Discrepancies (Block 2: χ^2 = 11.513† , R^2 = 0.089)			
Conference expenses partially/not funded	0.515	0.187	0.017
Reported loans at graduation	0.505†	0.360	0.054
Reported unemployment at convocation	0.803*	0.636	0.066
• Deviation from avrg program duration (mths)	-0.017	-0.040*	-0.168
• Student with visa status	-0.879	-1.087†	-0.110
Not an UW alumni before enrolment	0.055	0.139	0.021
Conditionners (Block 3: χ^2 = 20.296* , R^2 = 0.142)			
• Women alumni	0.225	0.205	0.031
Visible minority status	0.365	0.465	0.064
Married during degree program	0.043	0.723†	0.106
With dependents during degree program	-1.030†	-1.328*	-0.157
• Master's of accounting alumni	-0.931	-0.633	-0.065
• Doctoral alumni	-0.459	-0.823	-0.100
• Science degree	0.564†	0.980*	0.146
Financial support (K $)	-0.003	0.000	0.000
Frequent visits to the Graduate House	0.566†	0.454	0.069
Full research or teaching involvement	0.131	0.014	0.001
No research or teaching involvement	-0.206	-0.419	-0.045
Item Nonresponse (Block 4: χ^2 = 2.604 , R^2 = 0.148)			
Missingness on program characteristics	-0.168	0.540	0.082
Missingness on socio- demographics	0.434	0.696	0.042
Frame Interaction (Block 5: χ^2 = 2.387 , R^2 = 0.182)			
Full Model: χ^2 = 70.188* df = 26, McFadden R^2 = 0.182, λ = 0.096**			

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Table 4-9 Very Satisfied Departmental Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-1.677***	
Subjective Discrepancies (Block 1: χ^2 = 36.639***, R^2 = 0.046)			
Program completion longer than expected	0.107	-0.323	-0.048
Not satisfied with superv. on authorship issues	-1.160*	-0.630	-0.047
Financial support distribution not fair/equitable	-0.936***	-0.574*	-0.070
Resources not adequately available for TA duties	-1.550***	-0.815*	-0.076
University expectations before enrolment	0.383†	0.335	0.045
Coop program expectations before enrolment	-1.340*	-0.536	-0.040
Research expert. expectations before enrolment	0.553	0.411	0.028
Objective Discrepancies (Block 2: χ^2 = 12.488† , R^2 = 0.061)			
Conference expenses partially/not funded	-0.817*	-0.819*	-0.077
Reported loans at graduation	-0.227	-0.148	-0.023
Reported unemployment at convocation	0.216	0.046	0.005
● Deviation from avrg program duration (mths)	0.011†	0.007	0.030
● Student with visa status	-0.041	-0.077	-0.008
Not an UW alumni before enrolment	0.511**	0.284	0.043
Conditionners (Block 3: χ^2 = 40.399***, R^2 = 0.112)			
● Women alumni	-0.306†	-0.245	-0.038
Visible minority status	-0.368†	-0.124	-0.018
Married during degree program	-0.066	-0.836**	-0.126
With dependents during degree program	0.468*	0.952**	0.115
● Master's of accounting alumni	-2.209***	-1.532*	-0.162
● Doctoral alumni	0.468*	0.270	0.034
● Science degree	0.363†	-0.316	-0.048
Financial support (K $)	0.003	0.000	0.000
Frequent visits to the Graduate House	0.563**	0.320	0.050
Full research or teaching involvement	0.363	0.081	0.008
No research or teaching involvement	-0.040	0.148	0.016
Item Nonresponse (Block 4: χ^2 = 9.919** , R^2 = 0.124)			
Missingness on program characteristics	-0.999***	-0.769**	-0.120
Missingness on socio- demographics	-0.544	-0.860	-0.053
Frame Interaction (Block 5: χ^2 = 15.750***, R^2 = 0.144)			
Visa x Reported loans at graduation		1.602*	0.070
Visa x Doctoral alumni		-1.963*	-0.091
N Dur. x Doctoral alumni		0.031*	0.075
Full Model: χ^2 = 115.195* df = 28, McFadden R^2 = 0.144, λ = 0.000**			

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

[a.] Not entered, differs for each computation

Table 4-10 Very Dissatisfied Supervisor Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-2.918***	
Subjective Discrepancies (Block 1: χ^2 = 90.895***, R^2 = 0.161)			
Program completion longer than expected	0.858***	0.828**	0.124
Not satisfied with superv. on authorship issues	2.723***	3.330***	0.251
Financial support distribution not fair/equitable	0.857***	0.916**	0.112
Resources not adequately available for TA duties	0.486	0.180	0.017
University expectations before enrolment	0.616*	1.179*	0.157
Coop program expectations before enrolment			
Research expert. expectations before enrolment	0.206	1.065	0.072
Objective Discrepancies (Block 2: χ^2 = 6.619 , R^2 = 0.172)			
Conference expenses partially/not funded	1.025**	0.627	0.059
Reported loans at graduation	0.331	0.275	0.043
Reported unemployment at convocation	0.572†	0.575	0.061
• Deviation from avrg program duration (mths)	0.011	0.026*	0.112
• Student with visa status	-0.136	-0.186	-0.019
Not an UW alumni before enrolment	0.323	0.637*	0.097
Conditionners (Block 3: χ^2 = 33.009***, R^2 = 0.231)			
• Women alumni	0.481*	0.727*	0.114
Visible minority status	-0.572*	-0.583	-0.083
Married during degree program	-0.450†	-0.681†	-0.103
With dependents during degree program	-1.017*	-1.180*	-0.143
• Master's of accounting alumni			
• Doctoral alumni	0.067	0.319	0.040
• Science degree	0.239	0.308	0.047
Financial support (K $)	-0.004	-0.012†	-0.110
Frequent visits to the Graduate House	0.579*	0.202	0.031
Full research or teaching involvement	-0.347	-1.176*	-0.114
No research or teaching involvement	0.356	0.627	0.069
Item Nonresponse (Block 4: χ^2 = 5.143† , R^2 = 0.240)			
Missingness on program characteristics	-0.225	0.047	0.007
Missingness on socio- demographics	0.855†	1.622**	0.100
Frame Interaction (Block 5: χ^2 = 40.538***, R^2 = 0.312)			
Women x Not satisfied with superv. on authorship issues		-2.409**	-0.090
Women x Resources not adequately available for TA duties		-1.936*	-0.092
Women x Science degree		1.248*	0.093
Women x No research or teaching involvement		-2.032**	-0.108
N Dur. x Financial support distribution not fair/equitable		0.065*	0.092
N Dur. x Doctoral alumni		-0.057**	-0.138
Full Model: χ^2 = 176.205* df = 29, McFadden R^2 = 0.312, λ = 0.156†**			

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Table 4-11 Very Satisfied Supervisor Experience Cluster Membership

Predictors	Zero Order β	Full Model β	Full Model Z (β)
Intercept	__a	-1.145***	
Subjective Discrepancies (Block 1: χ^2 = 52.502***, R^2 = 0.054)			
Program completion longer than expected	0.059	0.025	0.004
Not satisfied with superv. on authorship issues	-3.445***	-2.534***	-0.200
Financial support distribution not fair/equitable	-0.322	0.032	0.004
Resources not adequately available for TA duties	-0.482†	-0.081	-0.008
University expectations before enrolment	0.747***	0.685*	0.096
Coop program expectations before enrolment	-1.847**	-0.748	-0.059
Research expert. expectations before enrolment	0.032	0.101	0.007
Objective Discrepancies (Block 2: χ^2 = 10.087 , R^2 = 0.065)			
Conference expenses partially/not funded	-0.132	0.431	0.043
Reported loans at graduation	-0.432**	-0.389*	-0.063
Reported unemployment at convocation	0.195	0.211	0.024
● Deviation from avrg program duration (mths)	-0.001	-0.008	-0.036
● Student with visa status	0.569*	0.467†	0.051
Not an UW alumni before enrolment	0.227	0.031	0.005
Conditionners (Block 3: χ^2 = 27.630** , R^2 = 0.093)			
● Women alumni	-0.383*	-0.253	-0.042
Visible minority status	-0.189	-0.309	-0.046
Married during degree program	0.290†	0.099	0.016
With dependents during degree program	0.269	0.096	0.012
● Master's of accounting alumni			
● Doctoral alumni	0.486**	0.314	0.041
● Science degree	0.795***	0.643**	0.103
Financial support (K $)	0.000	-0.008*	-0.077
Frequent visits to the Graduate House	0.128	0.068	0.011
Full research or teaching involvement	0.499*	0.315	0.032
No research or teaching involvement	-0.294	-0.256	-0.029
Item Nonresponse (Block 4: χ^2 = 8.821* , R^2 = 0.102)			
Missingness on program characteristics	-0.430**	-0.091	-0.015
Missingness on socio- demographics	-1.378*	-1.807**	-0.118
Frame Interaction (Block 5: χ^2 = 12.200** , R^2 = 0.115)			
N Dur. x Research expert. expectations before enrolment		-0.088*	-0.071
N Dur. x Reported unemployment at convocation		-0.057*	-0.095
Full Model: χ^2 = 111.239* df = 26, McFadden R^2 = 0.115, λ = 0.048**			

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Only one model, very dissatisfied with overall graduate experience, had a significant lambda value: for that particular model, predictors improve our capacity to classify alumni

cases by 27.8%. In sum, predictors of alumni satisfaction are not well suited to explain the initial variance or the cluster membership. It cannot be said that, taken as a whole, subjective, objective and program-related variables get at the essential reasons why responding alumni reported being very satisfied or very dissatisfied.

Two additional statistics can establish which variables did explain, within the bounds of the model overall explanatory value, the reason for reporting being very satisfied and being very dissatisfied. The standardized log-odds coefficient, $Z(\beta)$ (Menard, 2002: 51-53), ranks predictors according to the relative impact of each on satisfaction clusters. In addition, a comparison between the magnitude of the zero-order coefficient and the multivariate coefficient provides the means to assess the effective net impact of a given predictor; closely matched crude zero-order coefficients and multivariate coefficients indicate an unaltered, robust direct relationship with satisfaction clusters. A glance at logistic models reveals a convergence of the two statistical indicators. The primary reasons for being very dissatisfied are authorship issues, program completion expectations, income fairness, reporting loans at graduation. Science and women alumni are salient to overall program dissatisfaction. The primary reason for being very satisfied is the meeting of university-related expectation before enrolment (reputation, location, etc). Science and alumni with dependents are also shown to be salient for supervisor and department experiences respectively.

Two points can be drawn from the convergence of these statistics. First, these predictors virtually guarantee that one will report being very satisfied or being very dissatisfied on virtually all experiential clusters. Those who held university-related expectations such as reputation, for example, are likely to be very satisfied simply by

graduating. Second, these predictors represents the strongest, most robust, direct net reasons for reporting being very satisfied and being dissatisfied. They are strong predictors, and remain strong despite being incorporated in a multivariate environment. Since these variables were chosen for their potential topic saliency, and since these variables are not likely to be present in any frame or student records, they represent a real and serious threat of nonresponse error. However, a simple sensitivity analysis conducted on being dissatisfied with overall graduate experience established that only a large variation in representational bias would actually cause significant bias. For example, a 5% representational bias on authorship issues raised the adjusted probabilities from 1.8% to 2.1%. A 27% representational bias was required to obtain a significant change in probabilities.

Beyond issues of model fit, it is understood that topic saliency effects may also reveal themselves through the presence of outliers in the response pool. These alumni, at variance with other responding alumni, may unduly influence the value of coefficients. These coefficients are likely to change as more and more cases are incorporated in the response pool. Two analyses of residuals statistics were utilized: the standardized Cook's distance, DBETA, which provides an estimate of the overall changes in coefficient estimates if a given influential case were to be deleted, and DFBETA, an estimate of the change in a specific predictor's coefficient if a given influential case were to be deleted. Cases showing a DBETA larger than one, and/or a DFBETA larger than +/- 1 were considered influential. A review of residuals of all models did not reveal the presence of influential cases on any specific predictor. Dissatisfaction clusters, however, all exhibited large DBETA values but only a few cases were influential: two cases with DBETA < 1.2 on department satisfaction,

six cases with DBETA < 1.4 on overall dissatisfaction. None of these significantly impacted satisfaction models upon removal. Undoubtedly, the main issue remain misclassified cases, which exhibited normalized residuals larger than +/- 3.

4.4.2 Comparisons with National Graduate Survey

This section will further assess whether nonresponse should be construed as ignorable or nonignorable. A combination of GSES and NGS contextual analyses will be conducted to estimate the level of expected bias in satisfaction scores for a change in response rate. Further analyses will be conducted to assess whether estimates of predictor impact on satisfaction scores will themselves vary for a change in response rate.

Data analysis will proceed according to the strategies elaborated in chapter three and according to specific methodological considerations discussed in section 4.2. A model was constructed for each dataset; one for the GSES, and one for the NGS. Each model has its predictors centered on its GSES grand-mean. The intercept represents the log-odds of being very satisfied or being very dissatisfied when predictors are at their respective grand means. Significance tests between model log-odds ratios were performed by dividing the squared difference between log-odds coefficients by the sum of the two coefficient variances squared (Allison, 1999; J. Cohen & Cohen, 1983). The computation is compared with a given critical chisquare value.

Before any analysis proceeds, it should be noted that predictors in each logistic model are poor predictors of the initial "variance" or log likelihood observed in each satisfaction cluster. In some instances, NGS, which boasts a higher response rate, explains less of the

initial variance; it is conceivable that an increase in response rate brought an increase of sampling noise leading to a heteroscedastic pattern. At the very best, a given logistic model will explain up to 4% of the initial variance; at worst, only 1.4%. None are useful predictors for classifying cases in any cluster groups. As a consequence, any analyses hoping to demonstrate the presence of nonignorability should bear in mind that important variables are undoubtedly missing from our models. As a consequence, the net effect of nonresponse error, if present at all, is likely to be small.

4.4.2.1 Impact of Nonresponse on Satisfaction Scores

A first look at the differences in satisfaction scores between the two surveys reveals a slight but significant positive bias in the NGS dataset. Controlling for all objective discrepancies and program-related variables, GSES alumni respondents have a probability of being very satisfied with their overall degree program of 18.6%, while NGS alumni have a probability of 21.7%. The difference is marginally significant. An identical, albeit non-significant, trend is detected on the probability of being very dissatisfied. A probability of 8.46% is estimated for GSES alumni respondent, and a probability of 9.28% is estimated for NGS alumni. This trend is in reverse to what was hypothesized. If topic intensity was the prime motivator for survey cooperation, the probability on both clusters should have been higher in the GSES dataset. The comparison with NGS, presumably a better survey, should have incorporated more satisfied scores and, as a result, should have reduced initial probabilities observed in the GSES dataset. Admittedly, the "context", which was construed as a change in response rate, could have been confounded with a host of other issues mentioned earlier. At best, it could be concluded that being very dissatisfied, or being very

satisfied, is not conducive to cooperation behaviour. Since cooperation and contactibility are confounded, this reversed trend might also be the expression of hard to reach alumni.

A closer look at the "very satisfied" residuals for each model reveals a series of misclassified observed cases, estimating what should have been a "very satisfied" alumnus into the "other" (e.g. not very satisfied) category. They all exhibit normalized residuals larger than +/- 3 but none, to the exclusion of a single case in the GSES, are influential cases as indicated by the standardized Cook's distance (DBETA) and the constant's estimated change if the case was to be deleted (DFBETA). Indeed an inspection of NGS residuals simply reveals more of these same types of misclassified cases. The offending GSES case was removed from the model, and the model recalculated. Only minor changes occurred: the McFadden r-square analog increased slightly to 3.3%, the intercept decreased to -1.489, and the normal deviation from average program duration became non-significant. Upon inspection of the data, that particular case was at variance from other cases in relation to program duration. In sum, predictors explained poorly why a given alumnus would report being very satisfied with their overall graduate experience. This inability to explain their experience only gets worse as more and more of these types of alumni are incorporated in the response pool.

An identical situation occurs upon inspection of the "very dissatisfied" residuals. There is a large number of cases ($N > 30$) exceeding a value of +/- 3 on the normalized residual statistics. A single influential case on the GSES residuals exhibits an extreme value on the standardized Cook's distance (DBETA = 27.5); none are found in the NGS residuals. Removal of the influential cases on the GSES dataset only uncovered more, albeit smaller

impact, influential cases. The intercept increased to -2.262; women and master's no research alumni, who were marginally significant, are no longer significant. The McFadden r-square analog increased slightly to 4.4%. It is reasonable to conclude that the model poorly explains why alumni are very dissatisfied with their graduate experience. The muddled clustering of the very dissatisfied alumni, elaborated earlier, may explain, in part, why the model had such a poor fit. Nevertheless, as was observed with the "very satisfied" cluster, more of these misclassified cases are present in the NGS dataset. This would explain the decline in the McFadden r-square analog.

4.4.2.2 Impact of Nonresponse on Predictors of Satisfaction

The previous sections reported a marginal impact of nonresponse on satisfaction scores. Both adjustment models and contextual models showed a (very) slight increase in the probability of being very satisfied on the overall graduate experience cluster. Adjustment models have also highlighted predictors most salient to the respondent alumni. This section will assess if the impact of predictors on satisfaction scores varies when compared with responses obtained from a, presumably, better survey with higher response rates. Comparisons with the NGS dataset, shown in Table 4-12, revealed two significant differences in the log-odds of being very satisfied with one's overall degree program. The first change can be observed on the deviation from normal program duration. While the GSES odds of being very satisfied increase by 1.3% for each month's deviation from the average, the NGS odds of being very satisfied decrease by 0.9% for each month's deviation from the GSES average.

Table 4-12 GSES and NGS Logistic Model on Very Satisfied Cluster

Predictors	GSES			NGS			Coefficent Difference	Chi-Square for Difference
	β_{GSES}	s.e.	$z\beta_{GSES}$	β_{NGS}	s.e.	$z\beta_{NGS}$		
Intercept	-1.473***	0.095		-1.285***	0.053		0.188	2.987†
Subjective Discrepancies								
Objective Discrepancies	(Blk 1: χ^2 = 7.003 , R^2 = 0.009)			**(Blk 1: χ^2 = 66.507***, R^2 = 0.008)**				
Reported loans at graduation	-0.249	0.199	-0.045	-0.087	0.068	-0.016	0.162	0.593
Reported unemployment at convocation	-0.109	0.276	-0.014	0.068	0.075	0.012	0.177	0.383
• Deviation from avrg program duration (mths)	0.013	0.008	0.066	-0.009***	0.002	-0.064	-0.022	7.118**
• Student with visa status	0.370	0.267	0.045	0.212	0.143	0.015	-0.158	0.272
Conditionners	(Blk 2: χ^2 = 6.402 , R^2 = 0.017)			**(Blk 2: χ^2 = 136.490***, R^2 = 0.025)**				
• Women alumni	-0.104	0.190	-0.019	-0.348***	0.061	-0.067	-0.244	1.495
Visible minority status	-0.095	0.218	-0.016	0.064	0.085	0.008	0.159	0.462
• Master's no research alumni	-0.412	0.387	-0.051	-0.141	0.116	-0.016	0.271	0.450
• Doctoral alumni	0.103	0.230	0.015	0.114	0.087	0.014	0.011	0.002
• Science degree	0.188	0.225	0.034	0.468***	0.064	0.082	0.280	1.433
Part-time studies (proxy for no involvement)	-0.017	0.448	-0.002	-0.342***	0.095	-0.057	-0.325	0.504
Item Nonresponse	(Blk 3: χ^2 = 0.251 , R^2 = 0.017)			**(Blk 3: χ^2 = 28.071***, R^2 = 0.028)**				
Missingness on program characteristics	-0.126	0.471	-0.009	-0.056	0.084	-0.007	0.070	0.021
Missingness on socio-demographics	0.303	0.825	0.011	-0.879***	0.191	-0.070	-1.182	1.948
Frame Interaction	(Blk 4: χ^2 = 11.028* , R^2 = 0.031)			**(Blk 4: χ^2 = 23.742***, R^2 = 0.031)**				
Women alumni x program duration (months)	0.006	0.014	0.015	0.013***	0.004	0.050	0.007	0.231
Norm. Duration x doctoral program	-0.009	0.015	-0.026	0.016**	0.005	0.036	0.025	2.500
Norm. Duration x visible minority status	0.058**	0.019	0.115	0.007	0.006	0.017	-0.051	6.552*
Norm Duration x Part-time studies	-0.007	0.028	-0.010	0.009*	0.004	0.045	0.016	0.320
GSES Model: χ^2 = 24.684† df = 16, McFadden R^2 = 0.031, λ = 0.000								
NGS Model: χ^2 = 254.810* df = 16, McFadden R^2 = 0.031, λ = 0.000**								

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 4-13 GSES and NGS Logistic Model on Dissatisfied Cluster

Predictors	GSES β_{GSES}	GSES s.e.	GSES $z\beta_{GSES}$	NGS β_{NGS}	NGS s.e.	NGS $z\beta_{NGS}$	Coefficent Difference	Chi-Square for Difference
Intercept	-2.382***	0.151		-2.279***	0.073		0.103	0.377
Subjective Discrepancies								
Objective Discrepancies	(Blk 1: χ^2 = 3.716 , R^2 = 0.007)			**(Blk 1: χ^2 = 21.711***, R^2 = 0.004)**				
Reported loans at graduation	0.200	0.260	0.021	0.068	0.085	0.013	-0.132	0.233
Reported unemployment at convocation	0.198	0.352	0.014	0.076	0.095	0.014	-0.122	0.112
• Deviation from avrg program duration (mths)	-0.007	0.011	-0.020	0.002	0.002	0.015	0.009	0.648
• Student with visa status	-0.184	0.575	-0.013	0.277	0.189	0.021	0.461	0.580
Conditionners	(Blk 2: χ^2 = 11.551† , R^2 = 0.030)			**(Blk 2: χ^2 = 25.058***, R^2 = 0.008)**				
• Women alumni	-0.343	0.276	-0.036	-0.322*	0.127	-0.066	0.021	0.005
Visible minority status	0.630*	0.273	0.060	0.090	0.110	0.012	-0.540	3.366†
• Master's no research alumni	-1.116†	0.632	-0.080	0.138	0.189	0.017	1.254	3.614†
• Doctoral alumni	0.036	0.312	0.003	-0.086	0.127	-0.011	-0.122	0.131
• Science degree	-0.216	0.288	-0.022	-0.285***	0.089	-0.053	-0.069	0.052
Part-time studies (proxy for no involvement)	-1.432	1.170	-0.074	-0.006	0.109	-0.001	1.426	1.473
Item Nonresponse	(Blk 3: χ^2 = 0.195 , R^2 = 0.031)			**(Blk 3: χ^2 = 45.931***, R^2 = 0.016)**				
Missingness on program characteristics								
Missingness on socio-demographics	0.559	1.090	0.012	0.841***	0.141	0.070	0.282	0.066
Frame Interaction	(Blk 4: χ^2 = 4.801 , R^2 = 0.040)			**(Blk 4: χ^2 = 87.769***, R^2 = 0.031)**				
Women alumni x MAcc	-1.592	1.242	-0.057	0.640*	0.306	0.040	2.232	3.045†
Women alumni x visa status	-0.121	0.803	-0.004	-0.914*	0.373	-0.038	-0.793	0.802
Women alumni x unemployment	-0.675	0.725	-0.024	0.671***	0.165	0.082	1.346	3.277†
Visa x Part-time studies	-0.005	6.740	0.000	2.630***	0.480	0.079	2.635	0.152
Norm. Duration x MAcc	0.023	0.048	0.014	-0.020**	0.007	-0.081	-0.043	0.786
GSES Model: χ^2 = 20.262 df = 17, McFadden R^2 = 0.040, λ = 0.000								
NGS Model: χ^2 = 180.469* df = 17, McFadden R^2 = 0.031, λ = 0.019**								

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 4-14 GSES and NGS Logistic Model on Dissatisfied Cluster with Influential Cases Removed

Predictors	GSES β_{GSES}	GSES s.e.	GSES $z\beta_{GSES}$	NGS β_{NGS}	NGS s.e.	NGS $z\beta_{NGS}$	Coefficent Difference	Chi-Square for Difference
Intercept	-2.513***	0.218		-2.279***	0.073		0.234	1.036
Subjective Discrepancies								
Objective Discrepancies	(Blk 1: χ^2 = 3.716 , R^2 = 0.007)			**(Blk 1: χ^2 = 21.711***, R^2 = 0.004)**				
Reported loans at graduation	0.216	0.256	0.014	0.068	0.085	0.013	-0.148	0.301
Reported unemployment at convocation	0.324	0.352	0.014	0.076	0.095	0.014	-0.248	0.463
• Deviation from avrg program duration (mths)	-0.004	0.010	-0.007	0.002	0.002	0.015	0.006	0.346
• Student with visa status				0.277	0.189	0.021		
Conditionners	(Blk 2: χ^2 = 11.551† , R^2 = 0.030)			**(Blk 2: χ^2 = 25.058***, R^2 = 0.008)**				
• Women alumni	-0.277	0.261	-0.018	-0.322*	0.127	-0.066	-0.045	0.024
Visible minority status	0.444†	0.268	0.026	0.090	0.110	0.012	-0.354	1.493
• Master's no research alumni				0.138	0.189	0.017		
• Doctoral alumni	0.178	0.310	0.009	-0.086	0.127	-0.011	-0.264	0.621
• Science degree	0.048	0.272	0.003	-0.285***	0.089	-0.053	-0.333	1.354
Part-time studies (proxy for no involvement)	-3.712	2.850	-0.118	-0.006	0.109	-0.001	3.706	1.688
Item Nonresponse	(Blk 3: χ^2 = 0.195 , R^2 = 0.031)			**(Blk 3: χ^2 = 45.931***, R^2 = 0.016)**				
Missingness on program characteristics								
Missingness on socio-demographics				0.841***	0.141	0.070		
Frame Interaction	(Blk 4: χ^2 = 4.801 , R^2 = 0.040)			**(Blk 4: χ^2 = 87.769***, R^2 = 0.031)**				
Women alumni x MAcc				0.640*	0.306	0.040		
Women alumni x visa status				-0.914*	0.373	-0.038		
Women alumni x unemployment	-0.281	0.723	-0.006	0.671***	0.165	0.082	0.952	1.648
Visa x Part-time studies				2.630***	0.480	0.079		
Norm. Duration x MAcc				-0.020**	0.007	-0.081		
GSES Model: χ^2 = 17.420† df = 10, McFadden R^2 = 0.036, λ = 0.000								
NGS Model: χ^2 = 180.469* df = 17, McFadden R^2 = 0.031, λ = 0.019**								

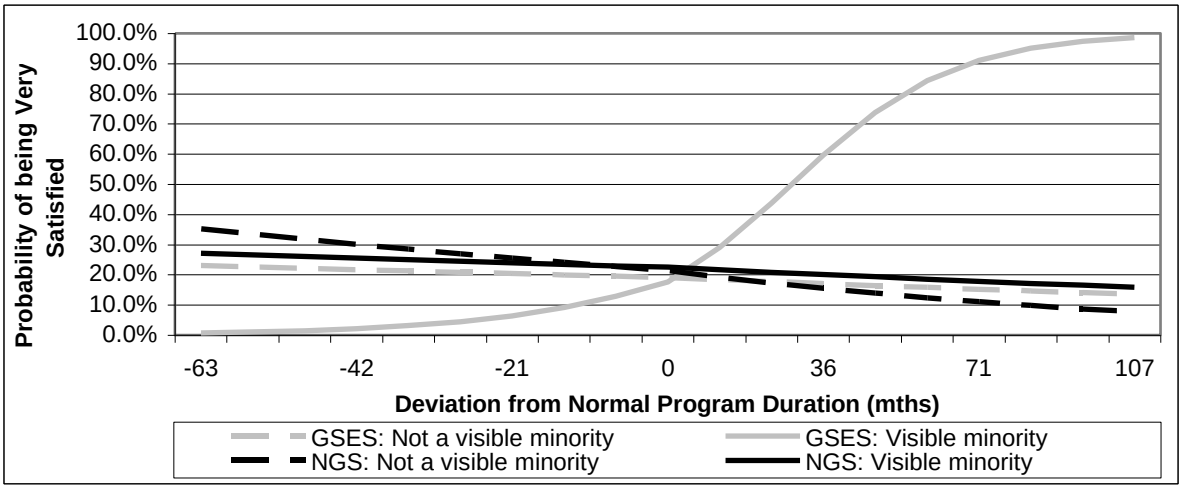
† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Although small, this sign change in the log-odds introduces a large difference in probabilities between the two models. Controlling for all other predictors, at maximum deviation from the average (107 months), the GSES probability of being very satisfied is estimated at 48%, the NGS at 9.5%. GSES respondents, other things being equal, would seem to assess their relative time-to-completion departure from institutional expectations as a positive contribution to their overall experience at Waterloo. The reverse was hypothesized: it was expected that the developmental discrepancy, what I should have by now, would produce a negative impact on the probability of being very satisfied. The longer the stay, the more tuition, the greater number of extension forms, and the more pressure to conform to expectations, were all thought to induce costs in the minds of the alumni. It is conceivable that length of stay, observed to be topically salient, introduced not cost-analysis benefits but a sense of identification in some alumni's mind, and these are the alumni who were more likely to cooperate. That bifurcation between cost-analysis and identification is likely to occur between those alumni who had a substantive relationship with their organization (OCB), from those who maintained only an instrumental one.

The second significant change in the log-odds of being very satisfied can be observed on the interaction variable visible minority by deviation from normal program. We can see, as shown in Figure 4-1, a clear difference between the two datasets. For the GSES dataset, visible minority alumni whose time-to-completion was longer than average have a greater probability of being very satisfied with their degree program, whereas for all other groups, in both datasets, the opposite is true. The situation gets paradoxical when it is revealed that GSES visible minority alumni are also more likely to report being very dissatisfied with their degree program (see Table 4-13). Although only marginally significant differences were

found, GSES visible minority alumni are 88% more likely than all other alumni to report being very dissatisfied; NGS reports they are only 9.4% more likely to be very dissatisfied than all others. Taken together, these two pieces of evidence, barring potential issues of coverage, flow with our original hypothesis on topic intensity. For visible minority alumni, the propensity to cooperate with a survey request may have been initiated by their overall graduate experience. The results are tentative but it is clear that, if comparisons with NGS are at all valid, the form resistant hypothesis does not hold for these alumni. This is an important finding considering the equity issues that visible minority students bring forth.

Figure 4-1 Interaction between Visible Minority and Deviation from Normal Program Duration in Very Satisfied Cluster



The remaining significant change in the log-odds of being very dissatisfied, although marginal, is more difficult to interpret. The odds of GSES master's no research alumni reporting being very dissatisfied decrease by 32% compared to all other alumni; for NGS master's no research the odds increase by 14%. Although this difference is marginally significant, it would appear that very dissatisfied master's no research alumni were

underrepresented amongst the GSES respondents. This is an important finding to the extent that it constitutes an indication of nonignorable nonresponse, and a refutation of the form resistant hypothesis. However, the problem of coverage between datasets looms very large. GSES master's no research alumni are essentially those enrolled in the master's of accounting program; NGS master's no research alumni reflect a much wider and diverse set of programs. It is extremely likely that coefficient differences are predominantly due to program differences and not response rate differences. For this reason these findings, although worthy of future research, cannot be considered valid.

It is interesting to note that reporting loans at graduation, graduating from a science degree, and women alumni were, in the previous sections, considered to be salient predictors of being very dissatisfied, and science degree salient with being satisfied with one's overall degree program. None of the coefficients on these predictors were significantly different when compared with the NGS dataset. Frame interaction on being very dissatisfied did reveal marginal differences, but the block chisquare on the GSES dataset remained non-significant.

Two additional tests are required to establish the source of coefficient variations. First, coefficient variations may be the result of influential cases that were either brought into the response pool at an early stage, affecting the GSES dataset, or at a later stage, affecting the NGS dataset. Both models on each cluster were tested for the presence of influential cases following the basic analytical strategy outlined in the previous section. Of interest are influential cases that may affect individual coefficients; six cases exhibited DFBETA larger than +/- 1 on master's no research, missingness on socio-demographics, and frame interactions for women alumni x macc, and visa x part time studies. Upon removal of these

cases, the GSES model changed substantially: it required the removal of visa, master's no research and missingness on socio-demographics and related frame interactions due to high collinearity between predictors. Table 4-14 displays the changes in the GSES model once influential cases are removed. Of note, both visible minority and the interaction women with unemployment, which were originally marginally significant, are now both non-significant. The influential cases are predominantly irate, Canadian-registered alumni, who completed their socio-demographic information in full, and where half came from the master's no research program, and a third were registered part-time. These cases were sufficiently influential to substantially decrease the odds of part-time studies, visible minority alumni, and increase the odds of unemployment, women, doctoral, and science degree on being very dissatisfied. In sum, those six cases are the source of the original marginal coefficient changes between the two models; they are the major contributor to nonignorability.

Secondly, a review of the coefficient differences in Table 4-12 and Table 4-13 indicates the presence of a general trend towards smaller coefficients in NGS models than GSES models. It is conceivable that the residual variance within the GSES model be much greater than the NGS model. As a result, differences in coefficients are not due to actual changes in the sample composition of a given predictor, but to the residual variation, or heterogeneity, in the GSES model compared to the NGS model[57]. To test this possibility, the analytical strategy proposed by Williams (2006a; 2006b) was performed using SPSS Plum (SPSS, 2003). Both datasets were merged into a single file. A dummy code representing the dataset was constructed. GSES predictors were directly entered in the model; NGS predictors

[57] This is precisely the point made by Allison (1999) when separate logistic models are used for group comparisons.

for which significant coefficient variations were found, were added as interaction terms (e.g. multiplied dataset dummy code). The dataset dummy code was entered as the scale factor, the heteroscedastic parameter to be estimated, NGS being the reference category. Results on the very satisfied cluster estimated a non-significant scale factor (sigma) of 0.380 (sig = 0.323); the scale factor on the very dissatisfied cluster was equally non-significant (sigma = 2.395, sig = 0.584). It appears that smaller NGS coefficients are not the result of residual variances.

4.5 Conclusion

Student characteristics have, for the most part, a weak influence on nonresponse error. Women alumni, visa-registered alumni, and alumni whose time-to-completion exceeded institutional norms were over represented in the respondent pool, but introduced little to no nonresponse error. Working under the missing at random assumption, adjustment models estimated the bias caused by overrepresentation on these variables to be well below one percentage point – well within sampling error. Comparison with the NGS dataset established that the form resistant hypothesis holds for these variables. As a consequence, overrepresentation on these variables can be considered ignorable.

These models have also pointed to several predictors that remained significantly and strongly related to satisfaction clusters, despite controlling for several subjective, objective, and program-related variables. Authorship issues, program completion expectations, income fairness and reporting loans, science degree and women alumni were all primary factors for being very dissatisfied; the meeting of university-related expectations, science degree and

alumni with dependents were salient to being very satisfied. Comparisons with the NGS dataset could only be performed on science degree, women alumni, and reported loans. None of the logistic coefficients on any of these three predictors was significantly different between datasets. Barring potential issues of coverage, these presumably salient predictors remained constant throughout. The form resistant hypothesis holds for these predictors as well.

Comparisons with the NGS dataset established that the form resistant hypothesis did not hold for those alumni who deviated from expected program duration, and for visible minority alumni. First, a small but significant change in the odds of being very satisfied was observed for deviation from average program duration; the GSES dataset reported a small but positive impact, whereas the NGS dataset estimated a small but negative impact. It was hypothesized that length of stay, perhaps a crude indicator of institutional identification, interacted with the developmental discrepancy, what I should have by now. Second, GSES visible minority alumni whose program duration exceeded institutional norms were much more likely to report being very satisfied than NGS respondents; once again, potentially an issue of length of stay. In addition, the impact of visible minority alumni on being very dissatisfied was marginally different across datasets; while exerting a positive impact on both datasets, the odds ratio were stronger in the GSES -- a potential issue of topic saliency.

In sum, while it could be said that student characteristics were not, by and large, the source of nonresponse error, there was evidence of errors stemming from missing at random nonresponse, and from nonignorable nonresponse. The results were severely limited by the number of enrolment variables and the public version of the NGS dataset. However, these

findings are sufficient to suggest further inquiry into the roles played in nonresponse error by length of stay within an organization (concept of semi-total institution perhaps) and racial inequalities.

Chapter 5
Influences of Departmental Characteristics on Nonresponse Error

5.1 Introduction

Departments provide the context that shapes much of graduate life at Waterloo. Through them flow many of the resources dispensed to the students, whether financial through grants and teaching assistantships, or research expertise through advisors and supervisors. They also provide a logical arena where similar peers interact with one another, (potentially) discuss their research, and generally learn the subtleties of collegial living. It is reasonable to conclude that departments will affect one's satisfaction with a given degree program. Consequently, if graduate experience is the basis for cooperation with an alumni survey, departments may shape both graduate experience and survey cooperation behaviour. This chapter focuses on the effects of departmental characteristics on nonresponse error.

To reiterate briefly, there are several reasons to believe that departments could be generative of nonresponse error. Departmental diversity in terms of the proportion of minority groups registered in a given department marginally contributes to satisfaction with one's major. Peer influences, such as average GPA, average program duration, proportion of doctoral students, may play a role in the integration and enhancement of student behaviours and outcomes. Both diversity and peer influences can be construed as indicators of departmental cohesiveness to the extent that there is more integration amongst students, and a better climate for exchange. Cohesiveness was closely associated with organizational

citizenship behaviour. Departmental engagement, measured as the focus of research, access to faculty, and budget allowances are, relatively speaking, more likely to induce organizational citizenship behaviour (e.g. engaged very satisfied alumni). Finally, department crowding, the density of registered students within the sphere of a single department, is likely to create less concern for the affairs of all, and presumably, less meaningful substantive relationships with peers, and as a result, (potentially) to weaken any possibility of organizational citizenship behaviour.

Our objectives in this chapter are to assess whether response rate disparities observed between departments substantially alter the representativeness of the combined alumni convocations of 2000-02, and, as a result, alter our understanding of alumni experience. To that end, we need to extend our analysis started in Chapter 4 by introducing departmental determinants of satisfaction that may also impinge on the decision to cooperate with a survey request. This chapter begins with an explication of the data analysis strategies and modelling decisions associated with hierarchical analysis. This is followed by the results of the hierarchical linear intercept, random and fixed models.

5.2 Methodological Considerations

The methodology deployed to analyze the departmental component of nonresponse error rests on a series of assumptions and analytical approaches that severely limits the conclusions of this study. Firstly, it should be reiterated that nonresponse bias is to be conducted on respondents only. Variations in departmental response rates within the response pool will be used as the estimate of nonresponse bias. The implication is the

determination of nonresponse bias as relative bias: we are assessing whether topic saliency (at the departmental level) explains, given the response pool, why certain departments were more successful than others in recruiting responses from alumni. The danger is the inherent lack of external validity; there is no certainty that observed nonresponse bias amongst respondents would be reproducible under higher alumnus response rates. There is also no certainty that observed bias, if any, is the result of departmental characteristic. Respondents may have self-selected themselves in such a way as to be homogenous with respect to satisfaction clusters. Ultimately, there is no certainty as to whether we are dealing, in fact, with topic saliency effects (cooperation dimension) or with contactibility effects.

Secondly, it should be noted that hierarchical model construction imposes upon the researcher additional constraints. A primary concern is the introduction of specification error by omitting variables that share their variance with other variables in the model, including within and between group variances. The solution is not to systematically remove variables simply because they were found to be non-significant at the alumnus level[58]. The cost, however, is the potential geometric explosion of parameters to be estimated for each random slope. The number of parameters to be estimated is $m(m+1)/2 + 1$, where m is the number of random level-1 predictors (see Raudenbush & Bryk (2002: 258). As a consequence, given the sample size per department, we may not be able to detect multiple random coefficients. Since each test of slope heterogeneity is equivalent to a chisquare test, we are also increasing the likelihood of experiment-wise significance error.

[58] Raudenbush & Bryk (2002: 258) would suggest removing a given level-1 (e.g. alumnus-level) variable if the fixed effect and the random effects are both non-significant. Since randomness amongst predictors is one of the factors under study, we are not at liberty to remove any variables at this stage.

A second concern is the presence of multicollinearity amongst departmental variables. A look at the correlation matrix between contextual variables shows that for this one set, variables are highly correlated with each other. The correlation between the variables "department size" and "proportion of master's of accounting alumni" was observed as 0.814; the latter was removed from the analysis. The variable science-related department was removed from the alumnus-level model because of its perfect correlation with group definition (the variable was originally derived from dept groupings). A third concern is the treatment of missing cases. The HLM6 software deletes missing cases listwise. Since eighteen alumni did not record their department, individual-level case count was reduced to 774 instead of 812 cases used in chapter four. A dummy-code expressing department missingness was introduced in the hierarchical models in order to maintain the number of cases analyzed in the previous chapter.

Finally, hierarchical analysis procedures had to be sequenced to ensure that certain conditions were established prior to the assessment of nonresponse bias, and to ensure model convergence when a large number of departmental level variables are entered in a given model. Hierarchical analysis proceeded in three distinct steps. First, since we would hypothesize that departmental characteristics, generative of topic saliency, will introduce nonresponse error in our satisfaction models, we should detect the presence of a hierarchical structure in the data. Each logistic model introduced in chapter four was tested for the presence of significant variance in the intercepts (contextual models) or slopes (moderational models) at the departmental level. Second, if there was significant between-department variance observed, a nonresponse bias model was elaborated using departmental response

rate as an explanatory variable. Third, the model was further fitted with departmental variables deemed to be sources of topic saliency. Because of the total number of parameters to be estimated, department-level variables were entered as sets of predictors. Sets were prioritized according to their anticipated importance in explaining the random variance. Each set was entered following the forward stepwise likelihood test (SPSS, 1999: 51-52). A set was removed if it was determined that a change in model log-likelihood was not significant; said differently, if it was determined that a given set of predictors did not make a significant contribution to the explanation of satisfaction score variance. A cut-point for removal was set for significance levels greater than 0.1. The drawback in using this model building strategy is the inability to detect potential suppression effects between sets (Menard, 2002).

5.3 Nonresponse across Departmental Groups

Early in the data collection process of the GSES, it became clear that response rates, aggregated at the department level, were quite varied. This variation has always left doubt as to the representativity of the overall response pool. Some departments seemed to have been much more successful than others in securing alumni survey responses. Since the alumni population was unevenly distributed across departmental clusters, overrepresentation of densely populated departments could easily signal their complete dominance in the response pool. In addition, regardless of the numbers of alumni who responded in each department, the composition of alumni in each department may also be significantly different from the original alumni population. Both forms of representational bias could potentially be a source of nonresponse error if any of these variables are related to satisfaction; both may partly be caused by topic saliency. To begin addressing the impact of the potential representational

bias on nonresponse error, the first section reports on the distribution of alumni responses across departments; the second will compare the composition of alumni respondents with frame variables already enumerated in chapter four.

5.3.1.1 Distribution of Nonrespondents across Departmental Groups

Table 5-1 provides an indication of the representativeness of departmental respondent groups with available frame data from the GSES 2000-2002 population combined. The first column reports on the distribution of the alumni population for each department. Four departments account for more than a third of the alumni population: School of Accountancy (MAcc), Electrical and Computing Engineering, Computer Science and Management Science. The third column, population weight, clarifies this point. The School of Accountancy, for example, accounts for 17.2% of the overall alumni population; together, all four top departmental groups account for 37% of the alumni population. The remainder of alumni are scattered around the remaining departments. The second column of Table 5-1 reports on the response rate aggregated at the departmental level. Departmental response rates can be as little as 9.1% and as high as 71%. On average, departments have a response rate of 37.7% with a standard deviation of 11.3%: slightly more than two-third of departments will have a response rate ranging from 26.4% to 49%. The final column, population missingness, however, also shows that the number of alumni missing from each department is fairly uniform – with the exception of School of Accountancy. A chisquare test performed between population and respondent counts (not shown here) produced no significant representation differences.

Three points of discussion can be drawn from these data. First, the alumni population is already skewed towards engineer-related departments and toward the Accounting department providing the MAcc eight-months program/courses. If topic saliency has little to no bearing on these departments, its effect, if present at all, will be greatly diluted. Second, while response rate variance is fairly wide, the bulk of departments exhibit response rates over a narrow range. As a result, topic saliency effects may remain too small to detect, unless the more extreme response rates are marginal because of topic saliency. Third, the size of the alumni pool within a given department does not predict the response rate one may expect for that department. Indeed, when alumni population counts and respondent counts are tabulated together (not shown), the resultant chisquare is non-significant. In sum, School of Accountancy aside, the proportion of missingness is more or less evenly distributed across all departments; not one department or a class of departments singularly define nonrespondents. Thus, on the assumption that departmental size, an indicator of crowding, would weaken topic saliency, we would have expected to see, other things being equal, lower response rates for higher populated departments. Admittedly, departmental size may still bear some impact on the decision of an alumnus to participate or not to a survey request. That variable will remain an important factor in our hierarchical logistic models.

Table 5-1 Nonresponse across Departments

Rank	Departments	Convocation 2000-02 Combined			
		Alumni Count	Response Rate	Population Weight	Population Missingness
1.	Fine Arts	11	9.1%	0.5%	0.5%
2.	French Studies	9	11.1%	0.4%	0.4%
3.	Applied Mathematics	27	25.9%	1.3%	0.9%
4.	Statistics and Actuarial Science	84	26.2%	4.0%	2.9%
5.	Economics	57	26.3%	2.7%	2.0%
6.	Chemical Engineering	66	27.3%	3.1%	2.3%
7.	Philosophy	22	27.3%	1.0%	0.8%
8.	Political Science	22	27.3%	1.0%	0.8%
9.	Biology	95	29.5%	4.5%	3.2%
10.	Germanic and Slavic Studies	43	32.6%	2.0%	1.4%
11.	Physics	24	33.3%	1.1%	0.8%
12.	School of Accountancy	363	34.2%	17.2%	11.3%
13.	Psychology	82	35.4%	3.9%	2.5%
14.	Systems Design Engineering	56	35.7%	2.7%	1.7%
15.	Electrical and Computing Engineering	171	36.3%	8.1%	5.2%
16.	Mechanical Engineering	95	36.8%	4.5%	2.8%
17.	English	61	37.7%	2.9%	1.8%
18.	Management Sciences	116	37.9%	5.5%	3.4%
19.	Architecture	13	38.5%	0.6%	0.4%
20.	Local Economic Development	26	38.5%	1.2%	0.8%
21.	Computer Science	120	40.8%	5.7%	3.4%
22.	Combinatorics and Optimization	29	41.4%	1.4%	0.8%
23.	Recreation and Leisure Studies	24	41.7%	1.1%	0.7%
24.	School of Urban and Regional Planning	48	41.7%	2.3%	1.3%
25.	Kinesiology	46	43.5%	2.2%	1.2%
26.	Civil Engineering	80	43.8%	3.8%	2.1%
27.	Earth Sciences	82	43.9%	3.9%	2.2%
28.	Chemistry	68	44.1%	3.2%	1.8%
29.	History	34	44.1%	1.6%	0.9%
30.	Geography	54	44.4%	2.6%	1.4%
31.	School of Optometry	15	46.7%	0.7%	0.4%
32.	Pure Mathematics	10	50.0%	0.5%	0.2%
33.	Health Studies and Gerontology	25	52.0%	1.2%	0.6%
34.	Sociology	8	62.5%	0.4%	0.1%
35.	Environment and Resource Studies	21	71.4%	1.0%	0.3%
	Total	2107	100.0%	36.7%	63.3%
	Department not stated = 38			38.5%	

5.3.1.2 Composition of Nonrespondents across Departmental Groups

Departmental response rates do not tell the whole story. While higher response rates are generally indicative of better representativeness, presumably because they incorporate a more diversified portion of the response pool, it is entirely conceivable, as Groves & Couper (1998) aptly demonstrated, that higher rates may also favour certain population groups over others. Using the frame variables enumerated in chapter four, aggregated at the departmental level, we can assess the degree to which the composition of each department reflects the alumni population. Admittedly, the number of variables remains small, but they will allow the testing of three departmental characteristics hypothesized as sources of a topic saliency response mechanism.

A glance at Table 5-2 reveals no large differences between the average population proportions and means against those obtained in the response pool. The largest difference was observed in the proportion of women alumni: the population proportion averaged over all departments is 45.7%, while respondents were 50.8%, an overrepresentation bias of 5.1%. The standard deviation of the bias provides a measure of dispersion of the bias around the overall average of departments. It can be seen that the bias within departments is not clustered around the average. The average deviation of program duration and the proportion of women alumni exhibit large swings of bias amongst departments. However, a chisquare test between population and respondents counts across departments shows no significant departures in representativeness[59].

[59] Admittedly, the reliability of the chisquare test was greatly diminished due to the high number of cells with counts below five.

Table 5-2 Department-Level Predictors Representativeness

Contextual Predictors	Population	Respondents	Bias
Departmental Crowding			
Proportion of alumni in population			
Mean	0.029	0.027	-0.001
Standard Deviation	0.031	0.028	0.006
Departmental Cohesiveness: Peer Influences			
Proportion of doctoral alumni			
Mean	0.177	0.202	0.025
Standard Deviation	0.135	0.162	0.075
Proportion of MAcc alumni			
Mean	0.024	0.024	-0.001
Standard Deviation	0.144	0.140	0.003
Average deviation from normal program duration			
Mean	0.133	0.124	-0.009
Standard Deviation	5.350	5.181	4.396
Departmental Cohesiveness: Diversity			
Proportion of women alumni			
Mean	0.457	0.508	0.051
Standard Deviation	0.216	0.299	0.165
Proportion of visa alumni			
Mean	0.098	0.134	0.036
Standard Deviation	0.122	0.132	0.097

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

5.4 Nonresponse Error across Departmental Groups

The previous section did not reveal any significant departures of representativeness at the departmental level; neither the distribution nor the composition of alumni across departments differed significantly from known population values. While significant findings would have alerted us to the possibility of representational bias at the departmental level, lack of findings at the aggregate level may still result in satisfaction bias if any of these variables are operative at the alumnus level. In fact, as was briefly enumerated in the introduction, it is expected that these variables are causes of nonignorable nonresponse. Thus, these variables still warrant our attention, and must be included in our hierarchical

models. This assessment of nonresponse error will begin with an analysis of the impact of nonresponse on satisfaction scores; it will be followed by the impact of nonresponse on predictors of satisfaction scores.

5.4.1 Impact of Departmental Nonresponse on Satisfaction Scores

The assumption of the logistic regression models, developed in chapter four, was that observations of alumni satisfaction levels are independent from each other; no dependence should be present amongst alumni due to departmental groupings. In an institutional setting, such as a university, an effect of the department on the alumni would be likely. It is reasonable to expect that satisfaction scores have some of their variance traceable back to the effect of departmental groupings, that is to say there is between-department variance. In addition, if satisfaction (e.g. its intensity) was the basis for self-selection in response to a survey request, we would expect that the probability of being very satisfied or being very dissatisfied would vary with departmental response rates. Since these rates have been observed varying anywhere between 9.1% to 71.4%, it is reasonable to expect between-department variance in satisfaction scores. Thus, whether through natural variance, or through self-selection, the assumption of independence is likely not to hold.

In this section, we wish to analyze between-department variations in the intercept for the presence of nonresponse bias, and subsequently for bias caused by topic saliency. Since the intercept is the outcome variable under study, logistic models developed in chapter four must be designed according to the assumptions of contextual analysis. This analytical model requires that the intercept be left random, and all other predictors were fixed. All predictors,

including departmental level ones, were grand-mean centered to their respective grand means. This facilitates the interpretation of the intercept, but more importantly ensures that department-level estimates correspond to the contextual impact and not the expected difference between two departments (S. W. Raudenbush & Bryk, 2002: 139-141).

5.4.1.1 Nonresponse Bias of the Intercept

Before any hierarchical nonresponse bias analysis can proceed, we must first ascertain whether the intercept varies significantly across departments. A base model was constructed by regressing the very satisfied and very dissatisfied clusters on all three experiential dimensions using all available student-level predictors[60]. The left-most column of Table 5-3 through Table 5-8 shows the results of the student characteristic base model. The intercept estimates the average probability of a given alumnus being very satisfied or very dissatisfied for all student-level predictors at their respective grand-means. The statistic of interest, Intercept-Tau, estimates the degree to which the probability of being very satisfied or very dissatisfied varies between departments. The estimated log-odds of being very dissatisfied with one's overall graduate experience are, on average, estimated at -3.780; the probability is estimated at $1/(1+e^{-(-3.780)}) = 2.2\%$. These probabilities do vary significantly from department to department. We would expect that 95% of departments' estimated

[60] Readers accustomed to multi-level analysis might have anticipated an intercept-only model as the base model. Several considerations guided our choice of base model. It should be recognized that the intercept-only would only provide an uncontrolled estimate of the intercept and the corresponding between-department variance; as such, its value would simply be to indicate the presence or absence of a hierarchical structure. A better model, in line with our contextual analysis, is to reproduce the logistic model already established in chapter four, but with its intercept left random.

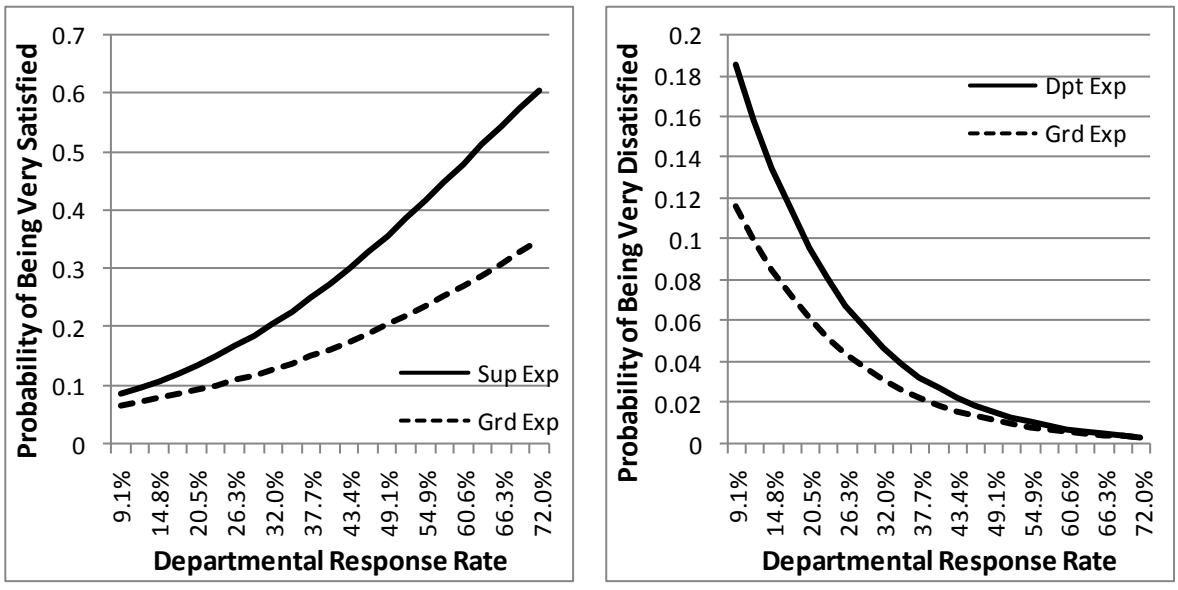
probabilities[61] fall within 0.4% and 11.2%. The log-odds of being very satisfied with one's overall graduate studies are -1.687; the corresponding probability is estimated as [$1/(1+e^{-(-1.687)})$] 15.6%; 95% of departmental probabilities will fall between 4.0% and 45.3%. The remaining experiential clusters also exhibit equally strong and significant between-group variances. We can conclude that individual department means are not reducible to the grand mean (intercept) and that significant variations are exhibited between departments.

The reported reliability is the average of individual reliabilities calculated for each department. It provides an indication of how reliable are, on average, intercept estimates of each department based on computing a regression coefficient separately for each department (S. W. Raudenbush & Bryk, 2002: 79). Individual reliabilities, computed for each department, are an estimate of the intercept (a) divided by the OLS estimates Y = a + e. An average reliability estimate near the value of one signals that departmental means are close to their OLS estimates. Lower average estimates of reliability indicate large error (e) values, stemming, in part, from the sample size of the department. Since level-2 predictors utilize departmental intercepts in their calculations, their reliability will also be affected (see (S. W. Raudenbush & Bryk, 2002: 46, 66, 72, 79, 115). Reliability estimates of departmental means for each dependent variable are moderate, ranging from 0.309 to 0.545. This is observable from both the extreme values taken by the dependent variables within each department and by the wide variations in sample size of each department.

[61] The range of estimated log-odds values among departments is calculated as the z = intercept ± $1.96\sqrt{\text{Tau}}$; probabilities are calculated as $1/(1+e^{-z})$.

Once it has been determined that model intercepts do vary significantly between departments, we are now in a position to assess whether observed between-department variations can be explained by variations in departmental response rates. Said differently, we need to ascertain if, given identical attributes on subjective, objective and program-related variables, alumni belonging to departments with different response rates also differ in their satisfaction scores. As we have stated in the previous section, if satisfaction-based response behaviour was indeed the operating basis for self-selection, then, departments with higher response rates should also correlate with an alumnus' probability of being very satisfied or very dissatisfied with their graduate experience.

Figure 5-1 Contextual Logistic Results for Intercept Models



Several logistic models showed significant contextual effects with respect to departmental response rates. The interpretation of these effects, however, is difficult to analyze because departmental response rate was centered at its grand mean; for an average

department (37.7%), the centered response rate is equal to zero, at maximum response rate (71.4%) it is equal to -37.7%; at minimum response rate, it is equal to 28.6%. As a result, the reader would have to take into account sign changes when calculating the actual impact of the contextual coefficient on the probability of being satisfied or very satisfied. To facilitate the interpretation, each logistic model that showed significant contextual effects was graphed to show expected variations in the probability of being very satisfied or very dissatisfied for a unit change in departmental response rate, once all subjective, objective and program-related attributes have been controlled for.

The left side of Figure 5-1 graphs the probability of being very satisfied with supervisory and graduate experience. An increase in departmental response increases the probability of an alumnus being very satisfied with both supervisory and graduate experiences. Thus, two alumni with exactly the same values on subjective and objective discrepancies will hold different probabilities of being satisfied by virtue of being in two different departments with different response rates. While it is too early to attribute any causality, it would seem that departments that were most successful in getting response from alumni are also related with very satisfied alumni. The right side of Figure 5-1 graphs the probability of being very dissatisfied with departmental and graduate experience. Contrary to expectations, in both cases, the probability of being very dissatisfied (of an alumnus) diminishes as the departmental response rate increases. It would appear that being dissatisfied is either not as salient as was believed, or that the overall response pool is overrepresented with satisfied alumni. The graph does reiterate what was shown in the left

side, that very satisfied alumni seem to be overrepresented amongst departments with higher response rates.

A closer look at the change in the statistic Tau in Table 5-3 through Table 5-8 provides an indication of the degree to which the variable departmental response rate explains the original between-department variance found earlier. In all cases, not all of the original variance was explained; the between-department variance remained significant despite the introduction of departmental response rate, albeit only marginally significantly so for students' graduate experience. This fact was corroborated by the change in the step chisquare; in both satisfaction clusters the departmental variable produced only a marginal improvement over the original unfitted student characteristic model. Nevertheless, the introduction of departmental response rate to the latter cluster did explain 40.6% of the initial variance. Other clusters only showed a modest decrease in variance. Satisfaction with graduate experience indicates a variance explained of 15%, satisfaction with supervisory experience, 28.4%, and being very dissatisfied with departmental experience, 16.2%.

5.4.1.2 Departmental Characteristics Bias of the Intercept

The trend indicates that alumni within departments that succeeded in getting higher response rates are less likely to very dissatisfied, and more likely to be very satisfied, above and beyond controlling for other departmental characteristics. The key issue at this stage is to ascertain if the nonresponse bias observed in the previous section can be explained by any of the four departmental characteristics elaborated in previous chapters. To repeat: we would hypothesize that alumni within departments whose proportion of very satisfied or very

dissatisfied is greatest will be most likely to find the topic salient, and as a result, to respond in greater numbers. This concentration of very satisfied and very dissatisfied is likely to occur within departments that are less crowded, more engaging, more diverse, and subject to positive peer influences. These four departmental predictors of alumni satisfaction are, as elaborated in chapter two, likely to create a topic saliency effect on cooperation behaviour.

The final column in Table 5-3 through Table 5-8 provides an indication of the degree to which departmental characteristics explain the effect of departmental response rates. Table 5-7, being very satisfied with supervisory experience, shows that the effect of departmental response rates is reduced by 19%. This reduction is attributable to both department size and department engagement (assigned budget); both variable coefficients sign are within expectations; both variable sets made a significant contribution to the original student characteristic model. Table 5-3, being satisfied with overall graduate experience, also points in the same direction. The effect of departmental response rate was reduced by 40% below significance. This time the reduction was attributable to department crowding and peer influences (grade point average). Being very dissatisfied with one's graduate and department experience, Table 5-4 and Table 5-6 respectively, also indicated a reduction of the impact of departmental response rate. This time the effect registered closer to 50%; Table 5-4 attributes this reduction to peer influences (proportion of doctoral alumni); and Table 5-6, to department diversity. Although the remaining two models did not register any departmental bias, some of the between-department intercept variance was explained by departmental engagement.

If there is any conclusion to be drawn from these data, it is certainly that nonresponse bias is present amongst alumni respondents. Departments that seem most successful in getting responses from alumni are also the same ones that house alumni most predisposed to answer being very satisfied. This bias does have its source in departmental characteristics hypothesized to produce topic saliency; this bias, however, is not completely explained by the same characteristics. Moreover, there doesn't seem to be any overriding trend as to which characteristic will be operative. Departmental crowding is inversely related with the probability of being very satisfied, and most likely to be a source of non-cooperation. Departmental cohesiveness expressed here under peer influences and department diversity did play a role in explaining nonresponse bias, but may benefit from further conceptual analysis. Finally, departmental engagement was also related to alumni satisfaction, most notably directed towards supervisory, and to a less extent departmental experiences. It did not play a major role in reducing nonresponse error, but considering the convoluted nature of the departmental response rate, its contribution should remain suspect.

Table 5-3 Contextual Model for Very Satisfied Graduate Experience Cluster

Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-1.687	0.187***	-1.728	0.182***	-1.603	0.153***
Dept. Missingness						
Response rate			3.227	1.805†	1.927	1.576
Dept. Crowding						
Alumni count					-0.004	0.002*
Dept. Engagement						
Graduate to undergraduate ratio						
Graduate/undergraduate to faculty ratio						
Assigned budget (M$)						
Dept. Peer Influences						
Proportion of doctoral alumni					0.819	1.161
Proportion of MAcc alumni						
Average grade point average					0.155	0.065*
Avg. deviation from normal program duration					0.020	0.028
Dept. Diversity						
Proportion of women alumni						
Proportion of visa alumni						
Average alumni age						
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	78.575*** df=23		81.696*** df=24		94.084*** df=28	
Step Chi-square, df	78.575*** df=23		3.121† df=1		12.389* df=4	
Tau (Intercept)	0.585***		0.497***		0.125*	
Reliability	0.543		0.507		0.243	

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 5-4 Contextual Model for Very Dissatisfied Graduate Experience Cluster

Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-3.780	0.324***	-3.787	0.314***	-3.897	0.323***
Dept. Missingness						
Response rate			-6.139	3.247†	-9.102	3.490*
Dept. Crowding						
Alumni count						
Dept. Engagement						
Graduate to undergraduate ratio						
Graduate/undergraduate to faculty ratio						
Assigned budget (M$)						
Dept. Peer Influences						
Proportion of doctoral alumni					3.673	2.081†
Proportion of MAcc alumni						
Average grade point average					0.022	0.119
Avg. deviation from normal program duration					0.085	0.051
Dept. Diversity						
Proportion of women alumni						
Proportion of visa alumni						
Average alumni age						
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	125.161*** df=23		127.916*** df=24		134.300*** df=27	
Step Chi-square, df	125.161*** df=23		2.755† df=1		6.384† df=3	
Tau (Intercept)	0.758*		0.450†		0.247	
Reliability	0.343		0.255		0.165	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5-5 Contextual Model for Very Satisfied Dept. Experience Cluster

Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-1.746	0.195***	-1.747	0.195***	-1.652	0.179***
Dept. Missingness						
Response rate			0.068	1.826		
Dept. Crowding						
Alumni count						
Dept. Engagement						
Graduate to undergraduate ratio					0.678	0.324*
Graduate/undergraduate to faculty ratio					-0.007	0.011
Assigned budget (M$)					-0.240	0.107*
Dept. Peer Influences						
Proportion of doctoral alumni						
Proportion of MAcc alumni						
Average grade point average						
Avg. deviation from normal program duration						
Dept. Diversity						
Proportion of women alumni						
Proportion of visa alumni						
Average alumni age						
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	82.075*** df=26		82.077*** df=27		91.919*** df=29	
Step Chi-square, df	82.075*** df=26		0.002 df=1		9.842* df=3	
Tau (Intercept)	0.560***		0.560***		0.240**	
Reliability	0.545		0.545		0.370	

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

Table 5-6 Contextual Model for Very Dissatisfied Dept. Experience Cluster

Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-3.339	0.270***	-3.375	0.268***	-3.287	0.234***
Dept. Missingness						
Response rate			-6.635	3.004*	-9.923	2.796**
Dept. Crowding						
Alumni count						
Dept. Engagement						
Graduate to undergraduate ratio						
Graduate/undergraduate to faculty ratio						
Assigned budget (M$)						
Dept. Peer Influences						
Proportion of doctoral alumni						
Proportion of MAcc alumni						
Average grade point average						
Avg. deviation from normal program duration						
Dept. Diversity						
Proportion of women alumni					2.438	1.164*
Proportion of visa alumni					-1.208	1.992
Average alumni age					0.395	0.141**
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	69.586*** df=26		74.797*** df=27		89.342*** df=30	
Step Chi-square, df	69.586*** df=26		5.210* df=1		14.545** df=3	
Tau (Intercept)	0.499***		0.418**		0.000†	
Reliability	0.309		0.273		0.000	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5-7 Contextual Model for Very Satisfied Supervisor Experience Cluster

Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-1.038	0.158***	-1.102	0.147***	-0.980	0.122***
Dept. Missingness						
Response rate			4.458	1.554**	3.622	1.244**
Dept. Crowding						
Alumni count					-0.008	0.003**
Dept. Engagement						
Graduate to undergraduate ratio					0.353	0.291
Graduate/undergraduate to faculty ratio					0.004	0.013
Assigned budget (M$)					0.235	0.106*
Dept. Peer Influences						
Proportion of doctoral alumni						
Proportion of MAcc alumni						
Average grade point average						
Avg. deviation from normal program duration						
Dept. Diversity						
Proportion of women alumni						
Proportion of visa alumni						
Average alumni age						
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	70.321*** df=25		78.228*** df=26		93.223*** df=30	
Step Chi-square, df	70.321*** df=25		7.907** df=1		14.996** df=4	
Tau (Intercept)	0.430***		0.308***		0.049†	
Reliability	0.537		0.463		0.148	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 5-8 Contextual Model for Very Dissatisfied Supervisor Experience Cluster

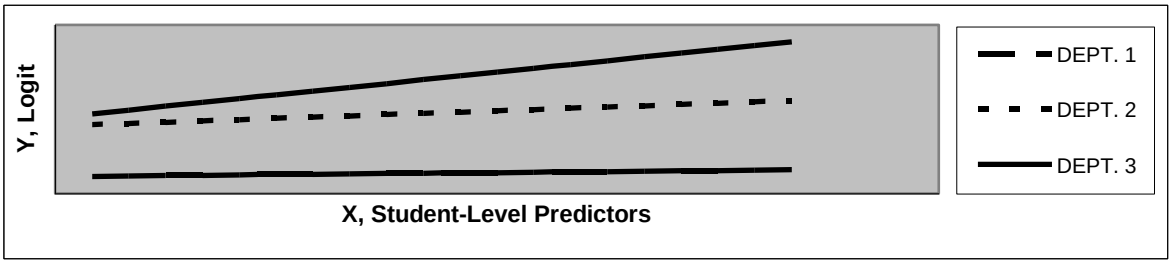
Departmental Characteristics	Student Char.		Response Rate		Dept. Char.	
	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.	Unit Coeff.	Robust s.e.
Intercept	-2.725	0.228***	-2.726	0.229***	-2.691	0.226***
Dept. Missingness						
Response rate			0.118	2.196		
Dept. Crowding						
Alumni count						
Dept. Engagement						
Graduate to undergraduate ratio					-1.145	0.527*
Graduate/undergraduate to faculty ratio					-0.018	0.014
Assigned budget (M$)					0.069	0.130
Dept. Peer Influences						
Proportion of doctoral alumni						
Proportion of MAcc alumni						
Average grade point average						
Avg. deviation from normal program duration						
Dept. Diversity						
Proportion of women alumni						
Proportion of visa alumni						
Average alumni age						
Dept. Nonresponse						
Skipped dept. identification						
Model Chi-square, df	127.013*** df=24		127.021*** df=25		132.923*** df=27	
Step Chi-square, df	127.013*** df=24		0.008 df=1		5.902 df=3	
Tau (Intercept)	0.514**		0.516**		0.341**	
Reliability	0.396		0.396		0.318	

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

5.4.2 Impact of Departmental Nonresponse on the Predictors of Satisfaction

The logistic models elaborated in chapter four not only assumed the same intercept value for each department, they also assumed that the predicted estimates, such as sex, visible minority, or reporting loans at graduation, were identical for all alumni regardless of home department. If, for whatever reasons, one would suspect that departments have the power to shape graduate experiences, they will undoubtedly have an impact on these same predicted estimates. In fact, if we were to construct a logistic model for each department using the same predictors found in the last chapter, we should see slightly different sets of means and predictor estimates, as shown in Figure 5-2. For some departments the impact of a given predictor will be much weaker (dept 1) or will be much stronger (dept 3). Depending on the characteristics of a given department, the models will vary between departments. In addition, as we have alluded in the previous section, if satisfaction was the basis for survey cooperation behaviour, salient predictors of satisfaction are also likely to be moderated by a given departmental characteristic. It is precisely this variation in the slopes of our logistic models that is in need of explanation.

Figure 5-2 Illustration of Slope Heterogeneity



In this section, we will push the analysis further by inquiring whether the impact of subjective, objective, and program-related predictors of satisfaction varies according to

departmental response rates; subsequent analyses will be conducted to determine whether this bias can be explained by the four departmental characteristic elaborated in the introduction. To that end, this section will first elaborate random-coefficient models, and proceed with fitted slopes-as-outcomes models[62], otherwise known as moderator models, for each significant random coefficient observed.

Before we proceed with the analysis, the logistic models developed in chapter four must undergo several changes. First, because we are now interested in the slopes, all alumnus-level predictors are group-centered; departmental-level predictors are left grand-mean centered. Group-centering ensures a more accurate estimate of slope heterogeneity (D. Hoffmann & Gavin, 1998; S. Raudenbush, 1989; S. W. Raudenbush & Bryk, 2002: 143-149). The usual strategy is to subtract an individual's score from the mean of the reference group in question, otherwise known as the relative model (Iverson, 1991). Admittedly, this transposition at the empirical level will impose a shift at the conceptual level. A relative model implies that an alumnus is somehow evaluating his or her experience according to the mean of his or her home department. While it does replicate Michalos (1973)' social comparison discrepancies, it nonetheless opens up difficulties as to who are the relevant others used as a comparative reference group, and through which mechanism this comparison is performed. While comparative judgements are much more likely to be based on observed behaviour than reference group means, in case of subjective discrepancies, it does give credence to Blalock (1984)'s contention that we are dealing with "social telepathy". Despite

[62] See Raudenbush and Bryk (2002: 75-85).

these difficulties, the relative model can be understood as a very rough estimate of the process of social comparison within a given department.

5.4.2.1 Nonresponse Bias of the Slopes

First, to test for the presence of slope heterogeneity, each slope in each of the logistic models introduced in the previous section, and now re-designed for moderational analysis, was systematically analyzed. For every slope left random, the model was tested for convergence, for its chisquare significance in relation to the closed model, and for slope Tau significance. If all three conditions were met, another slope was left random, in addition to those already known to be heterogeneous across departments. This procedure was done iteratively until all slopes were systematically tested. Only two slopes exhibited heterogeneity across departments, and both were only marginally significant. However, while the logistic model is significant with respect to an intercept-only model, the contribution of slope heterogeneity to the model was not significant. The step chisquare (not shown) was observed at 3.344 with two degrees of freedom for the departmental experience model (Table 5-9), and 5.099 with two degrees of freedom for supervisory experience (Table 5-10); only the latter was marginally significant. We are, as a consequence, unable to proceed any further with nonresponse bias analysis, and any subsequent analysis of departmental characteristics.

The question is how to interpret these results. With the current data set, it would appear that the lack of slope heterogeneity implies that the impact of predictors of being very satisfied or being very dissatisfied are essentially the same for every department. Regardless of which department a given alumnus may come from, their perceived discrepancy in relation

Table 5-9 Random Model for Very Satisfied Departmental Experience Cluster

Predictors	Fixed Effects		Random Effects	
	Unit Coeff.	Robust s.e.	Variance	Relia-bility
Intercept	-1.556	0.196***	0.789***	0.663
Subjective Discrepancies (χ^2 = 27.073***, R^2 = 0.012)				
Program completion longer than expected	-0.251	0.244		
Not satisfied with superv. on authorship issues	-0.769	0.527		
Financial support distribution not fair/equitable	-0.421	0.301		
Resources not adequately available for TA duties	-1.032	0.402*		
University expectations before enrolment	0.331	0.345		
Coop program expectations before enrolment	-0.405	0.707		
Research expert. expectations before enrolment	0.109	0.533		
Objective Discrepancies (χ^2 = 7.393 , R^2 = 0.015)				
Conference expenses partially/not funded	-0.652	0.403		
Reported loans at graduation	-0.045	0.227		
Reported unemployment at convocation	-0.148	0.316		
• Deviation from avrg program duration (mths)	0.012	0.008		
• Student with visa status	-0.506	0.412	1.154†	0.251
Not an UW alumni before enrolment	0.222	0.232		
Conditionners (χ^2 = 31.633***, R^2 = 0.029)				
• Women alumni	-0.588	0.232*		
Visible minority status	0.050	0.252		
Married during degree program	-0.896	0.292**		
With dependents during degree program	0.995	0.320**		
• Master's of accounting alumni	-1.489	0.797†		
• Doctoral alumni	0.319	0.297		
• Science degree				
Financial support (K $)	-0.001	0.004		
Frequent visits to the Graduate House	0.247	0.222		
Full research or teaching involvement	0.223	0.327		
No research or teaching involvement	0.114	0.357		
Item Nonresponse (χ^2 = 8.686* , R^2 = 0.033)				
Missingness on program characteristics	-0.783	0.292**		
Missingness on socio- demographics	-0.794	0.597		
Full Model: χ^2 = 78.129* df = 27, McFadden R^2 = 0.035**				

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

Table 5-10 Random Model for Very Dissatisfied Supervisor Experience Cluster

Predictors	Fixed Effects		Random Effects	
	Coeff.	Robust s.e.	Variance	Relia-bility
Intercept	-2.702	0.259***	0.616***	0.447
Subjective Discrepancies (χ^2 = 82.134***, R^2 = 0.040)				
Program completion longer than expected	0.668	0.262*		
Not satisfied with superv. on authorship issues	2.856	0.413***		
Financial support distribution not fair/equitable	0.922	0.437*		
Resources not adequately available for TA duties	0.178	0.430		
University expectations before enrolment	0.838	0.373*		
Coop program expectations before enrolment				
Research expert. expectations before enrolment	1.083	0.566†		
Objective Discrepancies (χ^2 = 3.284 , R^2 = 0.042)				
Conference expenses partially/not funded	0.422	0.514		
Reported loans at graduation	0.282	0.268		
Reported unemployment at convocation	0.609	0.357†		
• Deviation from avrg program duration (mths)	0.009	0.011		
• Student with visa status	-0.065	0.398		
Not an UW alumni before enrolment	0.285	0.397	1.018*	0.228
Conditionners (χ^2 = 35.632***, R^2 = 0.059)				
• Women alumni	0.672	0.339*		
Visible minority status	-0.639	0.329†		
Married during degree program	-0.498	0.343		
With dependents during degree program	-1.200	0.412**		
• Master's of accounting alumni				
• Doctoral alumni	0.468	0.462		
• Science degree				
Financial support (K $)	-0.014	0.006*		
Frequent visits to the Graduate House	0.170	0.333		
Full research or teaching involvement	-1.297	0.513*		
No research or teaching involvement	0.416	0.425		
Item Nonresponse (χ^2 = 4.988† , R^2 = 0.062)				
Missingness on program characteristics	0.343	0.371		
Missingness on socio- demographics	1.222	0.676†		
Full Model: χ^2 = 131.139* df = 25, McFadden R^2 = 0.064**				

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

[a.] Not entered, differs for each computation

to average in their department will have the same predicted impact on satisfaction scores. Either alumni are not very sensitive to variations in discrepancies or the data itself does not allow these variations to be detected significantly. Both interpretations are possible. Alumni may simply be operating from a zone of tolerance where variations in discrepancies will not result in any real changes in satisfaction levels. Equally possible is that the sample size within departments is simply too small to allow the detection of slope heterogeneity. In the case of the latter, it may be concluded that the form resistant correlation cannot be operative under low population counts, and for low response rates.

5.5 Conclusion

We have assessed whether departmental response rate could explain any significant variations in the between-department slopes and means from logistic models elaborated in chapter 4. Four of the six models under study exhibited nonresponse bias in the intercept. Of these, departmental characteristics explained a portion of the effect of departmental response rate on the probability of an alumnus being very satisfied or very dissatisfied. Although no clear trend did emerge from these departmental characteristics deemed to be generative of topic saliency, it would appear that departmental crowding, cohesiveness and engagement are strong contenders. Thus, a missing at random condition does prevail in these data, and barring the possibility of controlling for this error, a nonignorable condition will remain. In addition, no significant between-department variations were found amongst the slopes on any of the six models under study. Because the department sample size may have been insufficient to properly test this proposition, and considering the number of indicators

utilized, it cannot be ascertained whether the form resistant correlation holds or not. It does seem that the form resistant correlation is simply more difficult to detect.

Chapter 6
Influences of Survey Protocol on Nonresponse Error

6.1 Introduction

As was elaborated earlier, topic saliency can be enhanced or thwarted by survey design choices. Questionnaire design, survey mode, survey protocol and survey administration, through their impact on the probability of non-contact and non-cooperation, may bring into the response pool alumni who have different attitudes and experiences than nonrespondents. This chapter focuses on one aspect of survey design. We will examine the impact of survey protocol on nonresponse error.

During the spring 2002 convocation, two changes were made to the survey protocol of the graduate student exit survey. These changes were implemented using a 2x2 factorial experiment where all graduating alumni where randomized into one of four groups. The first change was to postpone the mailing date from two weeks after convocation (e.g. end of June, early July) to two months after convocation (e.g. early September). A change in the timing of the mailing date was expected to counteract some of the burden associated with dislocation and/or time constraints. The second change was to substitute the final reminder phone message with the sending of a dean-signed letter. As already mentioned in Chapter 3, it was expected that a final reminder letter would improve contact rates by circumventing any gatekeeper effect and by including overseas alumni in the final reminder. It was also expected that this final reminder would improve cooperation rate by providing a more coherent and persuasive argument to send the survey. The two changes were cross-tabulated,

July mailing with phone message being the control group. It was hoped that one or a combination of these changes in the implementation strategy might yield a higher response rate.

In terms of topic saliency, we would expect that a July mailing would align itself with the salience of having recently graduated, and the salience of attending the graduate convocation held at the end of June. However, the realities of having to deal with recent graduation, top-of-mind events such as relocating, finding employment, or attending to student loans are likely to drown the saliency of a survey request. Thus, it is most likely that topic saliency effects, if any, are topic intensity effects. Only the very dissatisfied and the very satisfied alumni would see this survey request as salient, that is to say, above the noisy realities of recent graduation. We would also expect that a letter reminder, while enhancing the saliency of the topic generally, would incorporate more alumni experiences by its ability to speak directly to each alumnus, and by its ability to make a case for survey cooperation (e.g. helpfulness, authority, etc). Conversely, the potentially ineffective phone reminder message is likely to draw on the most available and predisposed of alumni. Thus, we would expect that the combined July mailing and phone reminder, the control group, are most likely to be conducive to satisfaction-based cooperation behaviour. All other permutations will foil, in one way or another, topic intensity effects.

Our objective in this chapter is to address two essential questions. Are survey respondents representative of the alumni population of spring 2002 despite changes in the implementation of the graduate student exit survey? Would a change in the implementation of the graduate student exit survey also change our understanding of alumni satisfaction? To

test these propositions, we will begin with an analysis of the sample representativeness on socio-demographic and program-related variables. We will conclude this chapter with contextual and moderator effects of nonresponse on the satisfaction clusters elaborated in chapter three.

6.2 Methodological Considerations

Before proceeding with the results of this study it is important to reiterate its methodological procedures and limitations. This study will focus on the factorial experiment conducted during the spring/summer convocation of 2002. All graduating alumni (n=331) were randomized into one of four groups, elaborated above. A total of 109 responses were received. Analyses are to be conducted on the respondents, and assessments will be made as to the relative impact of factorial groups on nonresponse error. Three limitations flow from these analyses. First, the spring/summer convocation of 2002 population lacks external validity. It cannot be said that they are necessarily representative of alumni generally. Second, results from this study are based on respondents. Since the initial population satisfaction scores are unknown, it is uncertain whether nonresponse error will be found in the first place. Indeed, lack of nonresponse error findings may not signify that survey protocols had no impact on nonresponse error. The population under study may be homogenous with respect to the variables of interest. Or worse, respondents, regardless of survey protocol, are all alike with respect to the variables of interest[63]. Third, the inability to discern between non-contacts and refusals means that one cannot be certain that nonresponse error reflects topic saliency effects, as opposed to contactibility effects. As a result, while

[63] This empirical possibility was initially developed by (Keeter et al., 2000)

some factorial groups may show nonresponse error and others may not, the reasons for this variability may be confounded by contactibility.

Data analysis will proceed in the same fashion as stipulated in chapter three. An assessment of sample representativeness will be conducted to better understand how respondents differ from nonrespondents. Variables that are not part of the frame variables will be tested for relative overrepresentation across factorial groups; with the consequence that the initial bias is not known. Contextual and moderational logistics models will be developed to assess the level of nonresponse error introduced by survey protocol changes, and introduced by the over or underepresentativeness within factorial groups. The ability to detect nonresponse error across factorial groups, however, is severely limited by the low sample count in each group. Following Hox (2002: 177-179, 184-186), four hierarchical groups with an average sample count of 27, an anticipated intraclass correlation of approx. 0.15[64], and a critical region (alpha) set to 0.1, the power of the intercept-only model would be 0.46. This means that slightly less than ten times out of twenty we would fail to reject the null hypothesis (ignorability) when significant between-group variations are present (nonignorability) – as anticipated in the population. Thus, it is likely that only large effects will be detected. It also implies that the failure to detect an effect does not signify ignorability. It simply suggests that for the response pool on hand nonignorability effects

[64] An approximate value for the anticipated between-group variance is difficult to interpret. It is partly based on the anticipated effect of the survey protocol of cooperation behaviour; it is also based on the anticipated heterogeneity of the population with respect to satisfaction scores. We have used the three conventions of Raudenbush and Liu (2000) cited in (Hox, 2002: 184): small=0.05, medium=0.1, large=0.15 for intraclass correlations.

must be large to be statistically significant. This situation is conceivable if the rate of decline of nonresponse error in relation to response rate (see Figure 2-2) is non-linear.

6.3 Nonresponse across Factorial Groups

Whether the modified survey protocol did, in fact, magnify or dilute topic saliency remains an open question. It is clear that these two factors, mailing dates and reminder format, did produce an impact on the probability of obtaining a response to the GSES survey request. Response rates were calculated with return to sender envelopes counted as eligible (AAPOR, 2006). These were eligible alumni, but wrong contact information was on file at the time of the mailing period. The results of the factorial experiment conducted during the spring/summer convocation of 2002 are presented in Table 6-1.

Table 6-1 Response Rate by Factorial Groups

	July Mailing		September Mailing		Reminder Letter Total	September Mailing Total	Grand Total
	Phone	Letter	Phone	Letter			
Sample size	83	83	83	82	165	165	331
Completed surveys	30	34	18	27	61	45	109
Returned to sender	3	4	2	0	4	2	9
Response rate	36.1%	41.0%	21.7%	32.9%	37.0%	27.3%	32.9%

Source: GSES 2002 Spring Convocation

The highest response rate (41%) was obtained when surveys were mailed two weeks after convocation (July mailing), and a dean's letter was used as a final reminder. The worst (21.7%) response rate was with a September mailing period and with a phone message reminder. It was expected that a dean's letter would yield a better response rate than otherwise obtained in the previous convocation years (approx. 35%). A reminder letter,

which presumably enhanced cooperation and contact rates, yielded a response rate of 37%. It was not expected, however, that a September mailing would create such a drop in the response rate, yielding only 27.3%. It would seem to suggest that alumni are easier to reach, and/or that alumni find the survey more salient, in July. Considering that a July-Phone implementation strategy was already in place for the 2000-2001 convocations, it also suggests that only minor improvements in the response rates have been achieved with a reminder letter. It is instructive, however, that minor changes in the survey protocol did produce response rate variations ranging from 21.7% to 41%.

6.3.1 Sample Representativeness on Select Demographic Groups

As noted by (Gallagher et al., 2005), a change in survey design, whether through survey modes, incentives, or in our case reminder types, is likely to attract certain population groups over others. We have already seen that women alumni, visa students and those who took longer than anticipated to complete their degree program are more likely to respond. If their response was based on topic saliency, these population groups should be overrepresented in factorial groups that magnify saliency effects. We have hypothesized that the control group, July mailing with phone message reminder, is most likely to incorporate topic intensity effects. As a result, these population groups, if driven by topic intensity, are likely to be overrepresented in the control group. All other factorial groups should foil topic intensity effects, and thus not show bias in sample representativeness. A comparison of respondents with population counts across factorial groups is shown in Table 6-2.

Table 6-2 Sample Representativeness on Enrolment Variables

	Factorial Groups			
	September		July	
	Phone	Letter	Phone	Letter
Degree Program Variables	21.7%	32.9%	36.1%	41.0%
Conditionners				
Women alumni				
Sample	39.8%	37.8%	36.1%	37.3%
Respondents	44.4%	46.2%	53.3%	38.2%
Bias	4.6%	8.4%	17.2%*	0.9%
Science degree				
Sample	74.7%	70.7%	62.7%	68.7%
Respondents	72.2%	73.9%	62.1%	64.5%
Bias	-2.5%	3.2%	-0.6%	-4.2%
Doctoral alumni				
Sample	18.1%	15.9%	20.5%	18.1%
Respondents	11.1%	19.2%	13.8%	23.5%
Bias	-7.0%	3.3%	-6.7%	5.4%
Objective Discrepancies				
Deviation from avrg prog. duration (mths)				
Sample	-1.8	0.8	2.4	1.2
Respondents	2.5	0.0	1.5	3.1
Bias	4.2	-0.8	-1.0	1.9
Student with visa status				
Sample	12.0%	13.4%	8.4%	13.3%
Respondents	11.1%	8.3%	18.5%	23.5%
Bias	-0.9%	-5.1%	10.1%†	10.2%†

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

A glance at Table 6-2 provides marginal support for the contention that July mailing, and July mailing with phone reminder should introduce representational bias amongst select population groups. First, women alumni are overrepresented across all factorial groups. It would appear that women alumni are, as already seen in chapter four, more likely to respond. Since we cannot distinguish between non-contact and cooperation, we cannot ascertain if this likelihood is related to contactibility, availability or topic saliency. Nevertheless, the representational bias is significant, and most prominent, within the control group. This does

not confirm the presence of topic intensity effects. It does confirm, however, the necessary but insufficient outcome of satisfaction-based behaviour. If this overrepresentation is distinctive in terms of topic intensity, further analysis on the control group should reveal the presence of nonresponse error.

Second, Table 6-2 also reveals that visa students are marginally overrepresented in the July mailing but not in the September mailing factorial groups. Because visa students were never contacted by phone, only through reminder letters, we would have expected to see overrepresentation within final reminder factorial groups, not within mailing groups. Yet, it is also possible that overrepresentation in July mailing is an artefact of visa students' still living in Canada; visas may have expired in September. Thus, while it appears that visa students are overrepresented in salient July factorial groups, these uncontrolled confounding factors undermine the possibility that response was based on satisfaction behaviour. Nevertheless, it is worth knowing if this overrepresentation, by survey contact or by topic saliency, is conducive of nonresponse error.

Finally, it was unexpected to see no significant representational bias on the deviation from normal program duration variable. Although respondents in the September mailing with phone message deviate the most from time-to-completion, it has not reached statistical significance. The other factorial groups reveal chaotic variations in sample representativeness, whereas we would have expected greater deviations in the control group. This lack of support for our expectation may also be an artefact of non-contact. However, it was expected that science-related departments and doctoral degree would not reveal

representational bias. These populations did not reveal bias in chapter four, and were not expected to be amongst population groups who would find the survey topic salient.

Further analyses on sample representativeness were conducted to assess whether particular combinations of alumni responded more (or less) to a given factorial group. Table 6-3 shows the correlation between each variable pair. Population and respondent correlations were compared using Fisher's z transformation (J. Cohen & Cohen, 1983: 53-55). The table shows respondents within the September-Phone factorial group were composed of visa students whose normal program duration was longer than those found in the original sample. Respondents within the September-Letter factorial group contained more women in non-science-related departments. Finally, respondents in the July-Letter factorial group contained doctoral alumni who spent more time to complete their studies than those found in the original sample.

Both tables taken together, it could be said that clearly, women and visa alumni are overrepresented in some factorial groups. Once cross-tabulated with all other enrolment variables, however, these results do not amount to any specific trends not already observed in chapter four. From the data, one might be inclined to suggest that visa doctoral science alumni on the one hand, and women arts alumni on the other, are slightly and chaotically overrepresented across factorial groups. The inability to distinguish non-contact from refusal is one confounding factor. Equally important is the investigationwise error rate. At forty tests, the likelihood of randomly finding one significant test is extremely high. Using the conservative Bonferroni correction ($\alpha = 0.100/40 = 0.0025$), none of the tests would be significant. At best, it could be said that doctoral alumni who spent more time to complete

Table 6-3 Enrolment Variables Correlations across Factorial Groups

Predictors (September Mailing – Phone Message Reminder)	1	2	3	4	5
1. Women alumni					
Population					
Respondents					
Fisher's *z*					
2. Student with visa status					
Population	0.002				
Respondents	0.040				
Fisher's *z*	-0.135				
3. Deviation from avrg program duration (mths)					
Population	-0.072	-0.177			
Respondents	-0.414†	0.354			
Fisher's *z*	1.271	-1.895†			
4. Science Degree					
Population	-0.150	-0.210†	0.272*		
Respondents	-0.194	0.219	0.598*		
Fisher's *z*	0.161	-1.549	-1.419		
5. Doctoral alumni					
Population	-0.062	0.019	-0.493***	0.057	
Respondents	0.040	-0.125	-0.361	-0.175	
Fisher's *z*	-0.363	0.514	-0.559	0.831	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6-3 Enrolment Variables Correlations across Factorial Groups (Cont'd)

Predictors (September Mailing – Letter Reminder)	1	2	3	4	5
1. Women alumni					
Population					
Respondents					
Fisher's *z*					
2. Student with visa status					
Population	-0.012				
Respondents	-0.302				
Fisher's *z*	1.221				
3. Deviation from avrg program duration (mths)					
Population	-0.161	-0.140			
Respondents	-0.309	-0.163			
Fisher's *z*	0.615	0.088			
4. Science Degree					
Population	-0.328**	0.253*	0.246*		
Respondents	-0.736***	0.218	0.307		
Fisher's *z*	2.353*	0.139	-0.253		
5. Doctoral alumni					
Population	-0.132	0.221*	-0.198†	0.132	
Respondents	-0.167	-0.142	-0.477*	0.011	
Fisher's *z*	0.149	1.469	1.272	0.486	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6-3 Enrolment Variables Correlations across Factorial Groups (Cont'd)

Predictors (July Mailing – Phone Message Reminder)	1	2	3	4	5
1. Women alumni					
Population					
Respondents					
Fisher's *z*					
2. Student with visa status					
Population	-0.048				
Respondents	0.269				
Fisher's *z*	-1.391				
3. Deviation from avrg program duration (mths)					
Population	-0.026	-0.203†			
Respondents	-0.154	-0.279			
Fisher's *z*	0.564	0.335			
4. Science Degree					
Population	-0.145	0.055	0.120		
Respondents	-0.044	-0.101	0.145		
Fisher's *z*	-0.452	0.661	-0.111		
5. Doctoral alumni					
Population	-0.009	0.061	-0.181	0.022	
Respondents	-0.014	0.114	-0.450*	0.107	
Fisher's *z*	0.022	-0.226	1.317	-0.378	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6-3 Enrolment Variables Correlations across Factorial Groups (Cont'd)

Predictors (July Mailing – Letter Reminder)	1	2	3	4	5
1. Women alumni					
Population					
Respondents					
Fisher's *z*					
2. Student with visa status					
Population	-0.155				
Respondents	-0.151				
Fisher's *z*	-0.019				
3. Deviation from avrg program duration (mths)					
Population	0.123	-0.083			
Respondents	0.137	0.070			
Fisher's *z*	-0.067	-0.725			
4. Science Degree					
Population	-0.284**	0.111	-0.051		
Respondents	-0.295	0.283	-0.037		
Fisher's *z*	0.055	-0.817	-0.064		
5. Doctoral alumni					
Population	0.026	0.093	-0.187†	0.047	
Respondents	0.134	0.346*	0.378*	-0.025	
Fisher's *z*	-0.514	-1.265	-2.774**	0.328	

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

their studies (i.e. length of stay), were more likely to cooperate once prodded to do so by the Dean of Graduate Studies' reminder letter. It is not expected that any of these chaotic findings would have a systematic impact on nonresponse error.

6.3.2 Relative Differences in Self-Reported Respondent Characteristics

To further pursue our understanding of which population group is over or underrepresented in a given factorial group, comparative group count and means analysis will be performed on the remaining variables including objective and subjective discrepancies. This kind of analysis is different from sample representativeness. We are no longer measuring the bias in reference to the alumni population; instead, we are measuring the relative difference between factorial groups. The actual representativeness of the sample on these variables remains unknown. Frame variables, already analyzed, are reintroduced to assess their relative representational bias. This will provide an additional test as to whether absolute biases found on particular factorial groups make a unique contribution to between-factorial group representational bias.

Again, as pointed out earlier, we would expect to see most of the topic saliency effects occur in the control group, July-Phone. The remaining groups, because of the shift in the mailing date and reminder letter, should moderate topic saliency effects. The results, shown in Table 6-4, point to only two significant changes across a battery of socio-demographic and program-related variables, both located in the objective discrepancies variable set. Conference expenses partially or not funded, and Canadian alumni whose previous degree was not granted from the university of Waterloo, showed significant between-group variations. Both, however, behaved unexpectedly: the former exhibited bias

but as an interaction between mailing and reminder format; the latter, most susceptible to relocation problems – thus survey topic not top-of-mind – managed to answer in greater numbers in the control group.

Table 6-4 Between-Factorial Group Representativeness

	September		July		
	Phone	Letter	Phone	Letter	
Degree Program Variables	21.7%	32.9%	36.1%	41.0%	*p**
Subjective Discrepancies					
Program completion longer than expected	41.2%	56.0%	27.6%	35.3%	0.183
Not satisfied with superv. on authorship issues	0.0%	0.0%	16.7%	7.4%	0.113
Financial support distribution not fair/equitable	21.4%	35.3%	41.2%	30.8%	0.691
Resources not adequately available for TA duties	22.2%	21.4%	25.0%	8.7%	0.570
Coop program expectations before enrolment	22.2%	11.1%	20.0%	11.8%	0.608
Research expert. expectations before enrolment	5.6%	7.4%	3.3%	5.9%	0.926
University expectations before enrolment	72.2%	74.1%	63.3%	82.4%	0.396
Objective Discrepancies					
Conference expenses partially/not funded	57.1%	11.1%	9.1%	61.5%	0.012
Reported loans at graduation	33.3%	33.3%	33.3%	26.5%	0.916
Reported unemployment at convocation	29.4%	12.5%	25.0%	17.2%	0.517
• Deviation from avrg program duration (mths)	10.5	10.5	13.8	10.3	0.780
• Student with visa status	11.1%	8.3%	18.5%	23.5%	0.417
Not an UW alumni before enrolment	47.1%	58.3%	82.8%	62.5%	0.073
Conditionners					
• Women alumni	44.4%	46.2%	53.3%	38.2%	0.687
Visible minority status	33.3%	29.2%	23.3%	26.5%	0.892
Married during degree program	33.3%	46.2%	37.9%	41.2%	0.847
With dependents during degree program	16.7%	23.1%	20.7%	23.5%	0.944
• Doctoral alumni	11.1%	19.2%	13.8%	23.5%	0.643
• Science degree	72.2%	73.9%	62.1%	64.5%	0.771
Financial support (K $)	26.0	32.9	65.9	37.3	0.767
Frequent visits to the Graduate House	33.3%	42.3%	33.3%	55.9%	0.247
Some research or teaching involvement	0.0%	11.1%	6.7%	17.6%	0.205
No research or teaching involvement	22.2%	14.8%	23.3%	11.8%	0.593
Item Nonresponse					
Missingness on unemployment	5.6%	11.1%	20.0%	14.7%	0.534
Missingness on program characteristics	22.2%	33.3%	33.3%	17.6%	0.408

* The *p* statistic was derived from chisquare tests. Significance for both deviation from avg. program duration and financial support were derived using the Anova test.

Moreover, lack of between-group variables amongst variables such as involvement and subjective discrepancies is counter-intuitive. If involvement or discrepancies were an essential component of topic saliency, we would have expected significant between-group variations. Finally, at 25 tests, the investigationwise problem rears its head: a Bonferronni correction would suggest looking at significance tests 0.004 or below. Taken together, this lack of overall change across factorial groups would seem to point to an essentially homogenous respondent sample.

6.4 Nonresponse Error across Factorial Groups

Previous analyses on the absolute and relative bias in sample representativeness have pointed to four specific population groups: women alumni, visa-registered alumni, Canadian alumni whose previous degree was not granted from the university of Waterloo, and alumni whose conferences expenses were either partially or not funded at all. High investigationwise error rate casts doubt as to whether these four population groups are indeed truly different in their respective samples. There is also doubt as to whether these differences are an expression of cooperation behaviour or the product of contactibility issues. This section will nevertheless proceed to analyse these four population groups for the presence of nonresponse error. Should nonresponse error be found, it will weaken the case for serendipity, and point to further inquiries about cooperation behaviour. This section will begin with an analysis of the impact of factorial groups' nonresponse on satisfaction scores; it will be followed by the impact on predictors of satisfaction scores.

6.4.1 Impact of Factorial Nonresponse on Satisfaction Scores

A glance at the univariate distribution of satisfaction clusters reveals an overall alumni population that are overwhelmingly satisfied to very satisfied with their degree program: 72.5% and 18.5% respectively for global satisfaction clusters, 51.4% and 36.7% respectively for supervisory experience, 66.1% and 21.1% respectively for departmental experiences (see Appendix C). This level of skewness towards the satisfied end of the scale may indicate the presence of a homogeneous set of respondents – potentially very different from nonrespondents. This skewness may explain the lack of findings between factorial group nonresponse. It also poses challenges to the study of nonresponse error. Low counts on the very dissatisfied clusters will limit our analyses exclusively to the impact of nonresponse on the very satisfied clusters.

To begin our analyses, we must ascertain whether there are significant differences in the probability of being very satisfied between factorial groups. Since we have posited that the control group (July-Phone) is likely to align itself with alumni who are most available and predisposed to respond, and since being very satisfied is likely to be generative of topic intensity, we should see higher probabilities amongst the control group. Topic saliency effects should be highest in the control group, and potentially differ between July and September mailings. At the very least, the factorial experiment should point to significant between-group differences. The Anova test with random effects will provide a measure of the average probability of being very satisfied, the variation of these probabilities between factorial groups, and an overall reliability index. Table 6-5 shows the estimates for the fixed and random effects.

Table 6-5 Between-Group Variance for Very Satisfied Experience Clusters

Dependent Variables	Fixed Effects		Random Effects	
	Unit Coeff.	Model s.e.	Variance	Reliability
Overall graduate experience				
Cluster 3 (very satisfied) n=20	-1.493	0.247***	0.000	0.000
Experience with research supervisor				
Cluster 1 (very satisfied) n=56	0.061	0.242	0.083‡	0.355
Experience with the department				
Cluster 3 (very satisfied) n=23	-1.319	0.235***	0.000	0.000

‡ $p < 0.15$, † $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

The average probability of being very satisfied (as opposed to all other clusters) on overall graduate experience and departmental experience is 18.3% and 21.1% respectively. There are no between-factorial group variances around this average probability. The reliability is a reflection of the near zero factorial group variance. Experience with supervisor, however, did indicate a (very) marginal level of variance. The average probability of being very satisfied with one's research supervisor is 51.5%. The average was not significant. Between-group variations around the average log-odds (0.061) are estimated at 0.083 with a significance level set at 0.106. It is expected that 95% of factorial group probabilities would fall between 37.7% and 65.2%. Baysian residuals (not shown) around the average fixed effect would estimate the probabilities for the July-Phone group at 44.5%, July-Letter at 55.8%, September-Phone at 54.2%, and September-Letter at 51.7%.

Two points of discussion can be drawn from these results. First, it is remarkable to register so little amount of variation across factorial groups for overall graduate and research supervisor. This would indicate that each factorial group regardless of survey protocol, and regardless of response rate ranging from 22% to 41%, would have the same probability of

being very satisfied. That may point, once again, to an homogenous set of respondents; it may also indicate that topic saliency effects are not discernable for small response rate variations. Supervisory experience, however, would seem to be most salient, and susceptible to survey protocol variations. Second, July mailing with phone message reminder for supervisory experience has the lowest probability; other groups are rather similar to one another. That would certainly corroborate the notion that the control group is dictinct to the extent that it may draw from the most available and predisposed of alumni. Strangely, however, the probabilities are opposite to expectations. If the control group is most aligned with topic saliency, and being very satisfied is analogous to topic intensity, we should have seen higher levels of probabilities in the control group, certainly not less. There may be issues of contactibility lurking underneath these results; they may also indicate that being very satisfied is not synonymous with topic intensity.

The question remains whether variations in population groups' representativeness are also related to variations in the probability of being very satisfied. Said differently, we need to know if, controlling for population group differences, alumni will exhibit significant between-factorial group variations. That is the basis of the contextual model elaborated in chapter 5. Unfortunately, the entry of all four predictors in the contextual logistic regression model created numerical problems in the September-Phone group. Low counts (n=18) in this particular group produced zero counts for some cross-tabulations, multicollinearity between predictors, and extreme values on some predictors. The recovering strategy was to assess contextual effects for each and every predictor independently. The costs are that multivariate analyses would have controlled for indirect effects, and thus would have permitted an

assessment of the direct impact of each predictor on the probability of being very satisfied with supervisory experience.

The first step in this strategy was to assess whether any of the population groups are related to being very satisfied with the research supervisor. This would immediately provide evidence as to whether nonresponse is to be considered missing completely at random or not. All four population groups were tested individually along with the random portion of the intercept. Only visa-registered alumni showed a significant relationship with being very satisfied with the research supervisor; the remaining three showed no significant effect. We can only conclude that women alumni, Canadian alumni whose previous degree was not granted from the University of Waterloo and alumni whose conferences expenses were either partially or not funded at all are missing completely at random. Response rate variations on these groups will not produce nonresponse error on satisfaction scores. A contextual model was elaborated for visa-registered students. The random portion of the intercept was modelled with the proportion of visa students in each factorial group. Results of the contextual analysis are presented in Table 6-6.

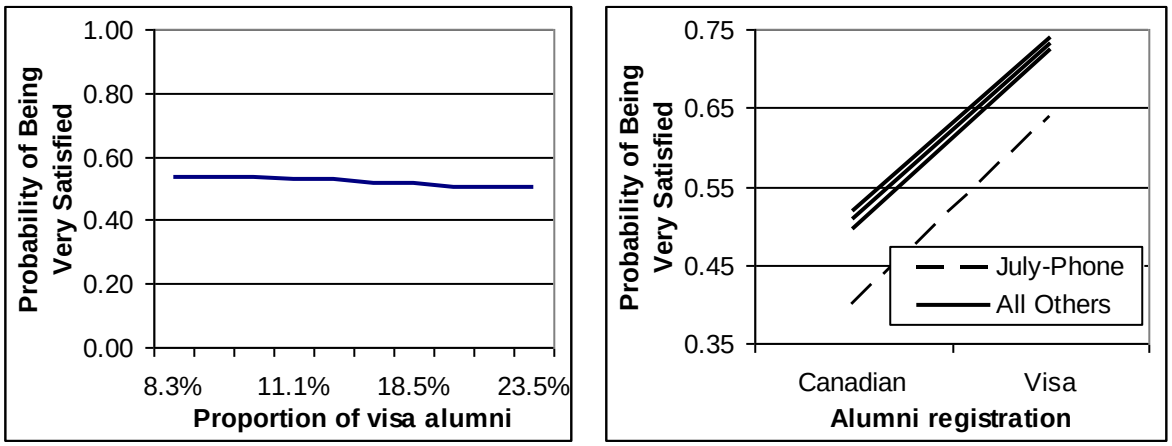
Table 6-6 Contextual Model for Very Satisfied Supervisor Experience

	Fixed Effects		Random Effects	
Predictors	Unit Coeff.	Robust s.e.	Variance	Relia-bility
Model for intercept				
Intercept	0.209	0.387	0.091*	0.370
Proportion of visa alumni	-0.855	3.158		
Model for visa status				
Intercept	0.992	0.291***		
Full Model: χ^2 = 3.148 df = 2, Tau = 0.091*				

† p < 0.10, * p < 0.05, ** p < 0.01, *** p < 0.001

The model estimates that, for visa-registered students, the log-odds of being very satisfied with one's research supervisor will decrease, on average, by 0.855 for one-percentage point difference between factorial groups. Said differently, probabilities of being very satisfied are expected to vary from 53.4% for visa proportions of 8.3% (Sept-letter) to 50.2% for visa proportions 0f 23.5% (July-letter). The left-hand portion of Figure 6-1 illustrates the predicted change in probabilities for changes in the proportion of visa students. The graph makes it clear that under or over representation in any factorial groups will have a very small impact on the probability of being very satisfied; indeed, the model tells us that the proportion of visa students is non-significant. While there will be significant variations between factorial groups, these are not caused by nonresponse (relatively speaking, given response rate ranging from 21.7% to 41%).

Figure 6-1 Contextual Logistic Results for Supervisor Experience (Very Satisfied)



If we re-arrange the predicted effect of visa-registered alumni on being very satisfied by factorial groups (see right-hand portion of Figure 6-1), we can immediately see that the control group still exerts a downward pressure on satisfaction scores. In fact, this control

group is responsible for the findings of significant variation across factorial groups. A dummy code was constructed to reflect this dichotomy between July-phone and all others, and subsequently was entered in the contextual model. Unfortunately, the model did not converge. We are left with uncertainty as to the meaning of the difference found between the control group and all others, particularly since it would reverse our hypothesized understanding of the topic intensity effects (through survey protocol) on nonresponse error.

6.4.2 Impact of Factorial Nonresponse on Predictors of Satisfaction

The previous section reported no impact of sample representational bias on satisfaction scores. This section will assess whether representational bias will moderate the relationship between the four populations groups and the probability of being very satisfied with one's research supervisor. Additional tests will be performed on secondary variables for which the original four population groups may have an indirect effect. A glance at the correlation table (see Table E-5 in Appendix E) indicates that resources not adequately available for TA duties might be impacted by bias in women alumni, and Table 6-3 would point to deviation from average program duration, science degree, and doctoral alumni.

We would expect the impact of these predictors on satisfaction scores to be larger in the July-Phone control group, and to be moderated by the combination of time and reminder protocol changes. September mailing should prove to be less prone to topic intensity effects, and reminder letters should incorporate into the response pool a wider range of satisfaction scores. These two factors should swell the probability of not being very satisfied. Moreover, it is expected that this trend will affect mostly the underrepresented category of these

predictors. Consequently, the odds of not being very satisfied for male alumni, Canadian-registered alumni, alumni who received their previous degrees at Waterloo, and alumni whose conference expenses were fully funded, will increase. The net result is a general decrease in the odds ratio of being very satisfied.

Table 6-7 Moderator Zero-Order Model for Very Satisfied Supervisor Experience

	Predictor Fixed Effects		Predictor Random Effects	
Predictors	Coeff.	Robust s.e.	Variance	Relia-bility
Intercept	__a	__a	__a	__a
Subjective Discrepancies				
Resources not adequately available for TA duties	-1.712	0.697†	0.556	0.163
Objective Discrepancies				
Conference expenses partially/not funded	-0.098	0.627	0.453	0.146
● Deviation from avrg program duration (mths)	-0.007	0.020	0.001	0.358
● Student with visa status	1.014	0.277†	0.032	0.023
Not an UW alumni before enrolment	0.216	0.333	0.067	0.081
Conditionners				
● Women alumni	-0.292	0.428	0.131	0.175
● Doctoral alumni	1.108	0.271*	0.012	0.010
● Science degree	0.066	0.243	0.001	0.001

† $p < 0.10$, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
[a.] Not entered, differs for each computation

Table 6-7 reports on the estimated between-group variation in the predictor's bivariate logistic regression coefficient on being very satisfied with supervisory experience. None of the predictors varied significantly across factorial groups. Two summations can be drawn from these results. First, female alumni, not UW alumni before enrolment, conferences expenses partially/not funded predictors are not significant within factorial groups, nor do they vary across factorial groups; missingness on these predictors can be considered missing completely at random. Relative over or under representation has no

impact on satisfaction scores. Second, all four predictors that did show over representation did not impact indirectly the bivariate relationship of doctoral, science, deviation from average program duration, and resources not adequately available for TA duties predictors on the probability of being very satisfied. Considering these results, and relative and absolute analyses in the previous sections, it could be surmised that the form resistant hypothesis holds across all four factorial groups.

6.5 Conclusion

Using the factorial experiment conducted on alumni during the spring/summer convocation of 2002, we have compared a battery of socio-demographic, program-related and discrepancies responses among four factorial groups. Both assessments of nonresponse and nonresponse error were conducted. Changes in survey design did increase response rate, but with only minor consequences for sample representativeness. Women alumni and visa-registered alumni were both overrepresented in the July mailing with phone message factorial group, but this overrepresentation did not register as significant when between-group variations were analyzed. The variables conference expenses partially/funded and Canadian students who received their prior degrees outside of Waterloo showed significant between-group variations. More importantly, while contextual analyses demonstrated no significant impact of representational bias on satisfaction scores, the July-Phone factorial group was noticeably different from the other groups. The probability of being very satisfied was below the others, but remained non-significant. Finally, the odds ratio on all predictors showing representational bias did not vary significantly across factorial groups. Admittedly, the factorial experiment was limited by low sample counts, and the inability to distinguish

between non-contact and refusals. In sum, respondents are likely to be homogenous in relation to their experience at Waterloo. Increasing response rates through survey protocol changes did not provide evidence of nonignorability.

Chapter 7
Cognitive Influences on Nonresponse Error

7.1 Introduction

In previous chapters, nonignorable nonresponse was studied using various descriptive and quantitative analyses at the student and contextual levels. For each quantitative analysis, factors were specifically chosen for their potential impact on nonresponse error; each has been shown to be, or may have proven to be a salient feature of the alumni's experience. The conjunction between the saliency of university experiences, coupled with the survey topic on university experiences, was hypothesized to be generative of nonignorable nonresponse. What have not been addressed, however, are the cognitive influences on nonresponse error. This chapter will focus on two aspects in particular. First, as Figure 2-1 illustrated, we need to understand the nature of topic saliency: that is to say, what is the basic mechanism that translates factors salient to the alumni into a decision to cooperate with a survey request. Second, as Figure 2-7 showed, the detection of nonignorability or ignorability may have been contaminated by the presence of measurement error in the form of response variance and/or response bias. Both of these cognitive aspects must be addressed to get a better understanding of the impact of topic saliency on nonresponse error.

The cognitive interview was set up precisely to better understand what goes on in people's minds as they are receiving and answering the survey. Through a "think-aloud" technique, respondents verbalize what they are thinking as they are answering the survey.

The objective in this chapter is to analyze these thoughts in order to assess the saliency of the survey and detect the presence of measurement error. At points, GSES data will be presented to contrast and compare with participants' verbalized impressions of the survey. The chapter begins with a specification of the methodology broadly defined in chapter three. It is followed by an analysis of topic saliency derived from the interaction of participants with the cover letter. The chapter concludes with an analysis of measurement error using verbal data on key questionnaire items.

7.2 Methodological Complications

During the data collection process, two unanticipated methodological problems were encountered and demanded revisions to the original methodological design. The reasons for these difficulties and the recovering strategies are discussed below. This section concludes with a description of participants to the cognitive interviews.

7.2.1 Sampling and Cooperation Rates

The original sampling design for the cognitive interviews was based on the assumption that recently graduated students would most closely resemble those alumni who had already received the GSES in 2000 to 2002. Alumni, unlike currently enrolled graduate students, have been through the entire program sequence (course, supervisor, thesis committee, thesis defence, etc). Since the survey deals with all of these aspects of their program, it was deemed necessary to locate recently graduated alumni and not enrolled graduate students. The consequence of this decision was to recreate, unwittingly, the same problems of nonresponse faced by the GSES survey, and to dismiss too quickly a readily-

available pool of currently available students. The sampling criterion based on maximum graduate experience "validity" may have been too high a price to pay.

The available pool of alumni was restricted to Kitchener-Waterloo residents for logistical reasons. Since such alumni would be geographically close to Waterloo, it would not be difficult for them to come in for an interview. To draw students to the University of Waterloo, we added a $20 monetary incentive. One hundred and thirty alumni of the 2003 convocation, having Kitchener-Waterloo as their primary home address, were each sent a cover letter soliciting their participation. A phone number and e-mail address were provided. Eight students out of 130 sampled alumni replied to the interview request; seven interviews were conducted; one respondent subsequently dropped out due to pressing travel arrangements.

Needless to say, a 6% response rate was not anticipated. Upon review of the recruitment strategy, it was felt to be regrettable that a reminder letter was never included in the original design. Such a letter, following Dillman (2000) would have helped in engaging alumni to participate in greater numbers. However, upon closer inspection, the recruitment strategy may have been operating under two false assumptions, both of which might have neutralized the effects of further reminder letters. First, a large (unknown) proportion of the pool alumni may not in fact have been geographically close enough to the University of Waterloo to participate in our study. Second, the monetary incentive may not have been sufficient to compensate anyone coming to Waterloo. It was felt that $20 would entice at least 15% of the alumni pool (20 interviews) living in Waterloo to participate in our study. It remains puzzling that such a high monetary incentive did not produce a stronger response, if

only in terms of inquiries about the study itself. We can only conclude that alumni may not have been geographically present in the first place. As a result, incentives and reminders might have gone unanswered.

In addition to mailing alumni, a snowball sampling method was devised whereby each interviewee were asked if he or she knew any other alumni who might be interested in participating in this study and if he or she could pass along my e-mail if any names came to mind. Unfortunately, a sampling design such as this relies on a very important premise that may not have been true for these participants. A snowball strategy presupposes the presence of a network of friends and acquaintances to pass along our participation request. Alumni may not be as connected to one another as they may have been during their graduate studies. As a result, the snowball strategy may have assumed too much of the alumni's ability to be an informant. In addition, our study on nonignorability may not have been very salient in the participants' mind once payment was made. In either case, the snowball strategy essentially failed and provided only one additional interview.

It became clear that a new recruitment strategy was required. The Graduate Studies Office (GSO), which provided the mailing addresses of the 2003 alumni, could not be taxed with any additional informal requests on our part as other equally important data requested several months earlier were still not forthcoming. The GSO, because of our mutual involvement in the GSES 2000-2002, was understood as the most logical interested party in our study; regrettably, on hindsight, it was also our only connection to the University of Waterloo bureaucracy. Creating new interfaces with the bureaucracy proved to be more difficult than anticipated. The idea of granting us access to student records or mailing

information, despite reassurances that it would be used for research purposes only, and despite ethics approval, was understandably met with suspicion. In the absence of any clear bureaucratic procedures to grant a graduate student such access, we were not able to secure the information needed to launch another set of recent alumni interviews. With time running short and resources running thin, it was decided to use whatever channels were already available to us. Attempts at creating new bureaucratic linkages were abandoned.

Two changes were made to the recruitment strategy. First, the alumni population was widened to include ABD soon-to-be-graduated students still registered with the University of Waterloo. As with their alumni counterparts, there were difficulties in tracking them down. Three options were available. The first was to approach departmental secretaries and other currently enrolled graduate students within the social science departments to collect any names of soon-to-be-graduated students. The second, suggested by the Director of Ethics, was to use graduate officers in each department as liaisons to keep us informed of new upcoming theses defences. The third was to utilize thesis announcements published in the University of Waterloo's The Gazette. The latter option was selected because of practicality and accessibility to campus-wide students. Following the announcement of a thesis defence in The Gazette, a cover letter was sent to the student's home department. Twenty letters were sent, three letters were returned unopened and two students replied. Two interviews were subsequently conducted. During the recruitment process, there were noticeable gaps in the advertisement of thesis defences. While such announcements were regularly posted in the newspaper version of The Gazette (now defunct, web only), a visit to the Gazette office

revealed that such advertisements are posted only sporadically and space permitting. As a result, an unknown quantity of graduate students was missed.

Concurrent to the above, a second strategy was devised. The pool of 2003 Kitchener-Waterloo alumni were sent a second letter inviting them once more to participate in our study. The rationale was two fold: 1) to reuse a readily available list of alumni; 2) to ascertain what is remembered of their experience at Waterloo two years after graduation. One hundred and twenty two letters were sent. Forty-three were returned unopened marked "not at this address". Only one alumnus replied through e-mail from British Columbia. The alumnus' brother, who received our letter, recognized its provenance and our name. It was he who notified the alumnus of our second attempt. The interview did not proceed for obvious logistical reasons. No further replies came from this sampling strategy.

It is perhaps one of life's ironies that a study on nonignorable nonresponse would yield such a dismally low cooperation rate. Without a doubt, there could have been better design strategies such as forging a formal arrangement with the alumni office, instead of relying on the GSO for example. It is also clear that slightly more interviews could have been collected if a reminder letter was sent during the original sampling strategy. Beyond these obvious issues, however, the on-going recruitment strategies sum up to a valuable set of evidence. First, there are serious doubts as to the validity (student rental address not parent's home), accuracy (incorrectly entered addresses) and stability (moving student) of the so-called "home address" provided by the student upon graduation. The number of return-to-sender letters suggests letters were sent either to incorrect addresses or to a gatekeeper-censored address. This fact does not bolster the assumption that home addresses used in the

GSES implementation were themselves valid, accurate or stable. If gatekeepers kept our letters from reaching their intended destination, then GSES reminder strategies may have been a failure. Second, if there was anything beyond non-contact (meaning letters were indeed received by the addressee), non-salience seems to prevail even with the offer of $20 compensation. Nevertheless, any further research on exit surveys may have to forgo stringent validity concerns and accept participants still involved in their program – with the added danger of students tunnelling their entire program satisfaction on the experience with their thesis defence.

7.2.2 Cognitive Interview Protocol

The cognitive interview was to proceed in three distinct phases. First, the participants were presented with an envelope containing a cover letter and the Graduate Student Exit Survey questionnaire. A combination of observational notes and the answers to standardized concurrent probe questions were to be collected to assess topic saliency, memory recall and contextual errors present in the survey material. Second, the participants were to think aloud while answering each question on the survey. Their verbalized thinking in conjunction with concurrent probes would allow the assessment of contextual errors as well as the presence of cognitive burden and question threat. The third phase would finalize the interview with a series of retrospective probes to collect overall impressions and any perceived mismatched expectations following the completion of the questionnaire (see appendix C for cognitive interview protocol)

The first phase did proceed along the general guidelines mentioned above but followed a less rigid and more fluid interview format than anticipated. Standardized concurrent probe technique maximizes the probability that each participant will provide comparable answers when exposed to the same stimuli. In practice, however, an overly structured interview would have made it impossible to establish any kind of convivial rapport with the participants, rapport that was deemed necessary in obtaining in-depth answers about their experience at Waterloo. Furthermore, the cognitive interview itself imposed a demand characteristic on the participants. Participants were too self-aware. Instead of simply voicing aloud what they were thinking, they were thinking about what they were thinking while answering the survey (being meta-cognitive). They knew they were being interviewed, focused too much on answering probe questions, and curtailed their answers to a narrow range. The interest in knowing about cover letter comprehension, cognitive biases and memory recall was still primary on the agenda, all of which, however, was better collected through an unstructured convivial interview format. The remainder of the interview was conducted through the think-aloud strategy as planned.

7.2.3 Cognitive Interview Participants

A total of nine participants[65] were recruited from the original 2003 K-W list of alumni, and The Gazette thesis announcements. Two of the K-W 2003 alumni participants revealed during the interview that they were still ABD students. The final list of 2003 alumni

[65] Cognitive interview sample sizes are notoriously low. Our sample size, albeit on the lower end of the spectrum, is not unique: such as Low (1999) with a sample size of nine participants. Moreover, it should be noted that each participant was interviewed for an average period of two hours. As a result, interview data provided depth if not breadth of alumni experiences. The analysis based on these nine cases remains however, speculative that is to say, generative of future hypotheses.

was undoubtedly revised after we received our copy from the GSO. There were a total of five alumni students and four ABD students. This mixture of alumni and soon-to-be graduated is not expected to affect the validity of the results obtained. The bulk of the survey questions pertains two both groups equally, with only a subset of questions relevant to alumni. While questions pertaining to the GSO may not apply fully for example, questions related to supervisory process, departmental experience, and university services would. As for the saliency of the items: to the extent that ABD thesis defence experience will not be markedly different from that of alumni, and to the extent that power relationships between ABD student, supervisor and thesis committee members do not play a major role in one's satisfaction level, the two populations should be equivalent.

Table 7-1 Cognitive Interview Participant Profile

Alumnus	Sex	Degree Program	Faculty	Registration	Program Duration (months)	Would return to Waterloo?
1	Male	MMath	Mathematics	Full-time only	20	Yes
2	Female	MSc	Science	Full-time only	35	No
3	Female	MASc	Engineering	Full-time only	36	No
4	Female	MASc	Engineering	Some part-time	37	Yes
5	Female	MMath	Mathematics	Full-time only	16	Yes
6	Male	MEng	Engineering	Full-time only	11	Yes
7	Male	PhD	Mathematics	Full-time only	46	No
8*	Female	MSc	Science	Full-time only	45	Yes
9*	Male	PhD	Arts	Full-time only	32	Yes

* ABD students

Table 7-1 provides a summary of the participants' program information. The participants were recruited almost exclusively from the Sciences, Mathematics and Engineering Faculties. Only two students were enrolled in a PhD program. Only two thirds

have expressed a desire to return to Waterloo, and only slightly more than half would have returned to the same program of study. Three participants admitted to being active in the Graduate Student Association. From their answers on the questionnaire (not shown here), more than half presented a paper at a conference, and two had publications. As a result, we would surmise that these participants were definitely engaged in their studies, and could be grouped as fully to somewhat involved in relation to their degree program experience at Waterloo. More participants may have provided a wider range of alumni, and yielded more information on the saliency of the survey topic.

It should also be noted that the recruitment of participants for cognitive interviews opens up the prospect of severe biases. While no attempt was made to achieve representativeness, rather diversity of experience was the intent, it remains a distinct possibility that only certain kinds of alumni may have been recruited. The issue is not so much that raised by Rosenthal and Rosnow (1975), in *the Volunteer Subject*; rather, it is the motivational basis for attending the interview and its impact on the determination of saliency. Tuckel, Leppo and Kaplan (1993) have shown that participants in a focus group motivated exclusively by monetary incentives were less likely to be psychologically motivated to take part in the research process, less interested in the research subject matter, and more likely to find the experience burdenful, all of which are indicators used to assess topic saliency and cognitive burden. Thus, our study must acknowledge the possibility that some of the participants will appear uninterested not because the topic is not salient but because participation was based solely on the receipt of a stipend.

7.3 Topic Saliency and the Decision to Cooperate

The question of whether topic saliency played a role in the decision to cooperate or not with the alumni exit survey grew as results trickled in from the summer convocation of 2000. Key questionnaire items, such as "would you return to Waterloo" showed unexpectedly low percentages (71.2%), and survey response rates were low as well (35%). From these results, there was a suspicion that respondents may not have been representative of the alumni population. Yet, when a random sample of nonrespondents was selected for answering machine/phone reminders, the nonresponse problem seemed much more one of non-contact and negligence. In fact, of those phone numbers called, very few nonrespondents were contacted directly. Instead, the answering machine or the parents were the most likely contact. When nonrespondents were contacted, most either declined to answer the survey or indicated various non-committal reasons for not sending it. Table 7-2 presents a selection of parents (P) and alumni (A) comments.

A quick glance at the parental responses demonstrates the extent of the specific problem of non-contact due to the action of gatekeepers. GSES questionnaires were mailed at the home address, presumably the parental address. The survey might have been, dependent upon the individual parent, forwarded to the alumnus' current address. Alternatively, parents may have informed the alumnus that an enveloped arrived from Waterloo. As Dillman (2000) demonstrates, the point of sending a survey package with a questionnaire and a cover letter, however, is to engage with the alumnus in the hopes that he or she will reciprocate by honouring the survey request. Reminders are meant to further this exchange relationship in the hopes of persuading the alumnus to send their surveys. Parents, acting as go-betweens,

break this exchange between some alumni and the GSO. The receiving date of the survey package, if it was sent at all, remains unknown. In addition, since the survey is anonymous, the GSO cannot definitively ascertain who did not respond to the survey. Under these conditions, one cannot tell the extent of nonresponse due to non-contact, much less whether non-contact was a source of nonresponse error[66].

Table 7-2 Selection of Comments from Final Reminder

A	Very Busy. Mature student. Didn't feel it applied to me.
A	Very Busy. Vacation, Work. Wait a couple months; we get tired with school work. Online version?
A	Was sent already, but didn't send postcard
P	Has not received questionnaire - parent's home, doesn't live there anymore
A	Too tired, would be better if sent later, but still fresh in my mind. [two months maybe?] yeah
P	Changed address twice. Waterloo, then Montreal. Does not know new address
A	Been too busy, forgot
A	Lost it, (kids?). Would like to have another
A	Don't remember getting one. Send another one
P	Works in Idaho
A	Ah yeah, yeah, I forgot ...
P	Went back to Indonesia
A	Just had a baby a few weeks ago, declined
A	Moved, will pickup mail at old address, and send it over
P	Moved. Don't know address
A	It's in the pile of mail. I'll look it up and send it to you
P	Moved to Northern Ontario
A	It's in the mail
A	Got lost in my mail pile. Send another one and I'll make sure it stays on top of pile
P	She's now in Germany
P	He just arrived here as of today
P	Off traveling won't be back till September

Source: GSES User's Guide (2002)

Leaving non-contact issues aside, the problem for survey researchers when faced with such generic responses as "too busy" or "it's in the mail" is how to interpret their meaning. One may consider these responses at face value and interpret "I'm too busy" as an instance of

[66] Schiltz (1988), for example, hypothesized that alumni who are hostile to their host institutions are unlikely to forward their change of address to the graduate studies office; as a result, non-contact may harbour a disatisfaction bias.

the opportunity-cost hypothesis: the alumnus calculated the costs and benefits of participating based on their available time and chose to decline. One may consider these responses as another way of saying "go away" (Brehm, 1993: 67) or as indicative of cognitive scripts such as "I never answer surveys" (Groves & Couper, 1998: 233). In either case, topic saliency or interest seems hardly the first consideration. If this was the case, one would expect some inquiries along the lines of "What survey?" or "What was it about?", following the telephone introduction: "to the best of our knowledge, we have not yet received your completed questionnaire" (see Appendix A).

Nevertheless, in the case of such refusals, one cannot know for certain, short of being present while the survey is received and answered, whether this apparent lack of interest could have nonignorable implications. We cannot say if topic salience or non-salience is related to one's satisfaction level: refusals may be based, as was theoretically demonstrated in chapter 2, on satisfaction. Individuals generally uninvolved but satisfied with their degree program would not find a satisfaction survey sufficiently salient to act on the basis of their satisfaction level. Moreover, these replies to our phone message might be the outcome of a cognitive script error whereby the intent of the survey was misunderstood as that of a "complaint" survey. Thus, while we may infer non-salience, these replies do very little in determining why they chose to delay answering, if they answered at all, the GSES survey.

This section will focus on the participants' behavioural and cognitive responses upon receipt of the survey material. The intent is to extract verbal responses while they are processing the survey material and more specifically while they are reading the cover letter. The analysis proceeds in two distinctive steps. First, an assessment will be made as to how a

survey topic is formed in the participants' mind, and how, if at all, their satisfaction levels may alter their reading of the message contained in the cover letter. Second, an assessment will be made whether the rationale for answering the GSES survey might be based on one's satisfaction levels.

7.3.1 En Route to Persuasion: Understanding the Survey Request

Surveys do not answer themselves. They are, at base, a request, an intrusion perhaps, to collect information on people's socio-economic backgrounds, attitudes, and experiences on a wide variety of topics. This request, when made through a self-administered questionnaire, will generally be introduced by a cover letter. The letter, when properly constructed, should provide reasons why and what kinds of information is requested, for what purpose and by whom. It is hoped that the sampled individual, upon reading the cover letter, will be persuaded to respond to our request. The persuasiveness to cooperate may come from the arguments of the cover letter and/or from the peripheral aspects of the survey material such as monetary incentives.

Petty and Cacioppo (1986)'s persuasion model provides the conceptual background that will explain the factors that influence the persuasiveness of the cover letter. Indeed, for a cover letter to be persuasive, it assumes that alumni are first motivated to read the letter, and second show the ability to understand its content. Both assumptions, however essential for topic saliency to operate, are in themselves insufficient to cause nonignorability. For topic saliency to operate as a biasing element in the decision to cooperate with a satisfaction survey request, one presupposes that alumni bias their cooperation on their pre-existing

satisfaction levels recalled during the appeal to answer the survey. Two such biases will be analyzed: the forewarning bias in relation to access to the questionnaire and schemata biases related to crystallized attitudes about graduate experiences. This section will begin with the first two assumptions, to be followed by an analysis of the two biases aforementioned.

7.3.1.1 The Motivation to Understand the Survey Topic

A cover letter will be useless if sampled individuals are unwilling to expend any cognitive effort to understand why the information was requested from them. If topic saliency is to have any impact on cooperation behaviour, it presumes that the cover letter was read to the extent that a survey topic was formed in the individual's mind. In other words, the degree of personal relevance to be drawn from a survey request depends on how the cover letter conveys the topic of the survey, and, in turn, on how motivated the sampled individuals are to read it. If the motivation for whatever reason is weak, topic saliency is likely to be weak as well.

The interview began by asking participants to pretend they were at home. They were informed that they were about to receive this white envelope in the mail. They were asked not to do anything more than what they would normally do at home. They were handed the survey envelope containing all survey material: cover letter, questionnaire, return envelope, and return mail postcard. Upon completion of reading the cover letter, participants were asked to articulate in their own words the content of the cover letter. A representative sample is shown in Table 7-3.

Table 7-3 Cover Letter Appraisal Verbal Responses

In your own words, could tell me what the cover letter is about?	
Alumnus 1	The cover letter was basically introducing the purpose of the survey and stating that ah hum I can ah hum hum oh now I don't now that 's funny heh I read it I processed it and now I don't remember most of the details. I remember some similar things to the consent form in the sense that any answers will be confidential, there's no number no number no tags associated with it so it won't be tracked as being my thing but that's … and I remember being contact information on there to which I can refer to later if I misplaced the form and I also recognize that the signature wasn't the signature on there but was a scanned signature printed in color … I mean I graduated from computer graphics so I notice these sorts of things maybe not others
Alumnus 3	The first thing I noticed there is a postage paid envelope so I mean if I had to pay like forty two-five cents or whatever. Even if not all that money I would probably say screw it. I wouldn't like actually spend some of my money to pay for it.—the survey. And then I would have to go and find a stamp …It's a fancy looking brochure. It doesn't almost look like a survey. Looks like you know a magazine or something else … It's kind of lengthy, the letter. It doesn't say too much about what they're sort of using it for. Sort of like they want to know our opinions or whatever.
Alumnus 4	They want me to speak some experience about the academic study for graduate student. I guess maybe they want find something. They need to improve in the future; if there are any problems in that school. I thought the comprehensive rank for Waterloo has fallen with the years. I'm just guessing.
Alumnus 5	They ask me to participate in the survey – relate my experience at Waterloo and hopefully that information could be used in the future to provide a better… to make them understand more about the graduate student life maybe? I'm really not sure but I think is something they need to evaluate their program or improve or stuff like that.

It is of interest to note how little cognitive effort participants showed in processing the cover letter. We might have expected, immediately upon reading the cover letter, a cogent summary of the cover letter's basic points; at the very least, an understanding of the purpose of the survey. And yet, three seconds after reading the cover letter, the majority of the participants (except alumni 9) simply could not provide, beyond basic generalities, the content of the cover letter. At best, when the most basic gist of the letter was retained, the formation of the survey topic seems to be intimately related to the purpose of the communication. That is to say, the motivation to understand what the cover is about is related

to the reasons why alumni are being asked to participate in a survey. Both alumni four and five, shown above, are grabbing at straws as to what the graduate studies office wants from them. Following Dillman (2000), the cover letter seems to have failed to provide compelling, persuasive reasons as to why their experiences needed to be collected through a satisfaction survey. If we are correct on this, topic saliency will not have a chance to operate because alumni self-select themselves based on the worthiness of the survey request as opposed to a particular satisfaction level.

For some participants, the cognitive focus, what they seem most keen to talk about when asked to convey in their own words the content of the cover letter, is precisely those secondary features of the survey that have very little to do with the message of the cover letter itself. Alumnus three was particularly adamant about the return envelope, the questionnaire design, and the wordiness of the letter, and yet could not provide any of its content. Alumnus one found particularly salient the stratagem used by the Graduate Studies Office to mimic a real signature, yet, couldn't tell who was the actual sponsor of the survey. When the interviewer revealed the sponsor's name, the participant, despite looking at the signature in the cover letter, replied: "It doesn't surprise me but I didn't notice". In fact, only half of the participants correctly identified the Graduate Studies Office as the sponsor; the other half either did not notice or did not connect the cover letter signature with sponsorship: "I just noted that the signature of being of the Graduate Studies but I don't know" (alumni 5). Consequently, neither the message nor its peripheral cues (appeal to authority) are motivating enough to understand the survey request.

7.3.1.2 The Ability to Understand the Survey Topic

Perhaps this lack of motivation stems from an inability to understand some key words of the cover letter that highlight what is sought. The issue is, following Petty and Cacioppo (1986), message complexity and comprehensibility stemming from the inappropriate use of ambiguous words that do not convey specifically what is being asked. In the case of the GSES cover letter, we found that the word "experience" in the sentence "[…] we want to learn more about graduate students' experiences with University's faculty members, staff and services", may have confused alumni as to what was being requested, and as a result, imposed a cognitive burden to expend cognitive efforts to seek out a meaning. When asked, for example, to explain in their own words the meaning of the word "experience" from the cover letter, participants were generally vague and limited to interpersonal relationships between their classmates, and quality of life at university. Only alumnus one, who could not remember the gist of the cover letter, was able to define the term experience using some of the terminology found in the cover letter.

The problem with these reflections on the term "experience" is that they are essentially at odds with the explicit definition of the word experience in the cover letter: "we want to learn more about graduate students' experiences with University's faculty members, staff and services. As well we seek more insight into students' progress through the various stages of their graduate program" (see appendix A). They are also at odds with the questionnaire focus on resources and supervisory themes. Consequently, there is a definite gulf between students and the University over the meaning of graduate experience. The saliency of the survey topic is foiled by interpretations that are essentially at cross-purposes.

If these reflections demonstrate what is salient in the alumni's minds upon reading the letter, then, short of asking questions solely on inter-personal relationships, it is unclear if any surveys about graduate experiences would to become salient in the alumnus' mind.

Table 7-4 Cover Letter Verbal Responses on the Meaning of Experience at Waterloo

In your own words, could tell me what the word "experience" means?	
Alumnus 1	My interaction with faculty, my interaction with staff ensuring in having certain things done or having me to complete certain form. My experience presumably with other students and also in the classroom and maybe extra curricular interests as well the kind of thing that appeals to grad students and what would make a graduate student come here
Alumnus 2	They're wondering if I enjoyed myself. If it tells us that or if it was worth my one two three year, worth my money, worth my time, worth my effort, and whether I came out feeling good about it or if I had a crappy time which I never came in the first place.
Alumnus 7	I think maybe it means how I study here and how's my life at the university. Maybe it means how I feel about the relationship between me and my supervisor and my department

7.3.2 Forewarning Bias: Access to the Survey Material

If motivation and ability to read the cover letter are weak, the alumni can still be persuaded to cooperate with a survey request if they can have prior access to the questionnaire. Groves and Couper (1998) anticipated that self-administered questionnaires might be particularly vulnerable to nonignorable nonresponse because all material is at the disposal of the alumni for their perusing. The alumni, upon perusing the questions in the questionnaire, decide not to answer the survey because the "survey topic", surmised according to their own interpretation of the material, is of little or no interest to them. Conversely, alumni may decide to cooperate with a satisfaction survey request precisely because the "survey topic" is of interest to them. In either case, access to the survey material

allows the alumni to focus on the survey questions, and potentially use them to base their decision to participate on their overall satisfaction with their degree program. The self-administered questionnaire, unable to counteract non-salience with a tailored motivational rebuttal, as would be the case for a telephone or a face-to-face interview, finds itself at a serious disadvantage. As a result, topic saliency effects might be more pronounced in self-administered surveys such as the GSES.

Upon submission of the GSES survey package, three participants immediately disregarded all other material and zeroed in on the questionnaire. One of these participants admitted to looking at the questionnaire to assess how long it would take (alumni 2). The remaining six participants took the cover letter and read it as if it was part of a scripted behaviour: 1) open envelope 2), look for the cover letter 3), read letter. As one participant suggested, referring to the cover letter, "I guess this is where you start" (alumni 9). In all cases, once the cover letter was read, they immediately proceeded to answer the survey as if operating under a scripted behaviour (see chapter 2), perhaps originating from the familiarity of course evaluations and other forms to be filled out over the course of a student's career. Such habitual behaviour proved resistant to my attempts at redirecting their attention to the cover letter. In fact, in some cases, the survey material had to be turned over and the questionnaire closed, in order to get their attention focused on my questions.

The behavioural component of the cognitive interview does support the theoretical possibility that students will look at the questionnaire prior to reading the cover letter. Consequently, students may have their own impressions about the survey prior to being persuaded with arguments found in the cover letter. It is unclear, however, if such behaviour

will necessarily be guided by a desire to assess the topic, or to assess the burden of the task in terms of survey length. The participants were pondering the survey task itself. In practice, it is more likely, as the results have shown, that students will bring forth two habitual behaviours well learned through the several years within an academic institution: 1) attend to the instructions first, 2), answer multiple-choice questions. Therefore, having the survey material may not be as problematic as suggested by Groves. Students are likely to attend to the cover letter prior to looking at the survey; they will gather the topic of the survey from the cover letter and not by perusing the questionnaire. In that regard, mail surveys are thus no different from in-person interviews or telephone surveys: the interviewer can frame the survey topic and offer persuasive arguments to honour the survey request.

7.3.3 Schemata Bias: Saliency of Graduate Experiences

Another factor that may bias the ability to be persuaded by the cover letter's request to answer the survey comes from the alumni's prior knowledge, or "attitudinal crystallization" (Schuman & Presser, 1981) of their graduate experience. Schemata bias occurs when the cover letter is read, or the questionnaire is perused, with the biased lens of satisfaction or dissatisfaction with one's graduate experience. One possible avenue of such bias, as stipulated in chapter two, is the attenuating effect of satisfaction on cooperation behaviour. General satisfaction, in this case, reduces the alumni receptivity to the argument that graduate experiences are indeed needed to "find new and better ways of doing things at the department and Graduate Studies Office levels" (see Appendix A). Satisfaction provides the argument that everything was fine as it was, and consequently the alumnus would not be convinced that his or her voice is required. Conversely, a highly dissatisfied alumnus, may

agree wholeheartedly that the GSO does indeed need to be reformed, and will be persuaded to answer the survey. In this case, it could be suggested that the argument provided in the cover letter flows, or rather is congruent with, a highly dissatisfied attitude towards graduate experiences.

Cognitive interviews are limited in their ability to ascertain the degree of schemata bias on the decision to cooperate. However, the depth, breadth and intensity of recall associated with graduate experiences might provide some clues as to whether these recalled experiences would play a role in how a satisfaction survey request might be perceived. A sample of verbal reports, presented in Table 7-5, indicates that, in most cases, participants were satisfied with their experiences. Alumni three, and to some extent alumni seven recollected some negative experiences associated with their degree program. In both cases, however, they were able to easily recollect their experiences at Waterloo, and for the most part, to verbalize them eloquently. The potential existed for a biased reading of the cover letter in terms of their satisfaction levels, but as we have seen in the previous sections, they didn't seem to connect their recollections with the purpose of the cover letter. As to the possibility that dissatisfaction may play a role in the interpretation of the cover, the case of alumni three seems to disconfirm this possibility. The participant, while depicting a rather negative experience, did not seem to be persuaded by the message of the cover letter. Instead, peripheral aspects such postage stamps and lack of incentives were noticed.

Table 7-5 Recalled Graduate Experiences at Waterloo

Alumnus 4	Good. [what comes to mind, something particularly special about it]. It's a very specific program at Waterloo compared to other universities in Canada - such as the core program.
Alumnus 5	Quite positive. I liked it. That's why I'm doing my PhD here. [what were your expectations?] My expectations were to know about statistics at that time because I was not from that background before. I was doing in my working and my post-graduate study I found that it was a very important tools. And I just kind like it so I just wanted you know get a degree there and hopefully maybe even do my research in that area.
Alumnus 6	I think Waterloo is a kind of.. I first came here I was impressed at the engineering the background you know. From this building, or if you're going to this building you will saw, introduced to every area, research area. When you go to some of the classroom it's totally different, different than I was in my previous country. I think kind of a different culture you know. Different education, different programming. People give you the right to acquire experience. [when you came to Waterloo, what did you want from Waterloo? What were your expectations?] Oh, actually when I decided to continue my education [inaudible] university which I can't get in, and I choose Waterloo because it is famous throughout with my undergraduate major which is mechanical engineering. I think they have the famous university in the North American and when I came here I think that is true [laugh] that's what I was thinking. For me I hope when I get through this program I hope I update my you know the professional skills and update my knowledge. For example you can learn the latest software, how to get ideas new methodology.
Alumnus 7	I think it was a very good experience. from this university compared to when I studied in (country's name) But I also have some regret I think in some aspects at least in my department did not do well… for .. especially for international students [Is there a particular aspect, an event a person that comes to mind when you think about your experience at Waterloo?] Yeah, I learned how to do research from my supervisor. I had two supervisors. Yeah .. They were very nice. I think hum one encouraged me to find [inaudible], and another one taught me how to do research. So at least that …

Table 7-5 Recalled Graduate Experiences at Waterloo (continued)

Alumnus 1	Overall, it was a positive experience. I enjoyed working with my supervisor. My lab mates were generally congenial and friendly. Many are friends now some are just acquaintances and will probably remain so. I was the grad student rep for computing science for a year. I'm happy that I did that because I got to know a bit more how the university works at the grad level. I was involved as an undergrad. Got to know some of the things at the undergrad level but now I see things from a different perspective. I can see why certain things are done the way they are. So it's been positive in that way in terms of extra-curricular activities. I was captain of a sports team as well. I was able to meet more graduate students outside of my program that way. Overall I enjoyed the courses I took. I thought they were very useful. Two of them were directly useful to my research and they've been indirectly useful in trying to finding work. I have had a contract recently and I have been just interviewed for a full-time position at another company. Overall the experience was a good one. There were hiccups. There were some bad points but incredibly good points as well. Actually another, which I appreciate for being able to be in the master's program, is that I have been able to apply for scholarships and to take part in certain things. I left for a month at some point….
Alumnus 2	Well, I'm back so obviously I didn't hate it. It was good. There were some good parts and some bad parts but overall I'm pretty happy with the place. […] I'm just trying to isolate my master's from my undergrad that's why it's taking me a second. The people. I really enjoyed my department. I enjoyed the people. I liked most of the professors in the department, the social activities. I enjoyed the research and stuff. There's a lot of great researchers in my department but I think it for the most part it's the people in general. The first thing that I really remember
Alumnus 3	I think it really depends on the kind of department and stuff you're in. I have to say intellectually I really didn't learn lot. I'd have to say, at least for my department, the focus is on undergraduate studies and so personally I don't think it was the best place necessarily for me to be. I think most of what I learned was in personal skills more than anything else. Yeah I had some issues. [did it meet your expectations?] No! I think the big thing was with funding. I was unfunded after my second year. I had lots of difficulties with my supervisor, as well as with the grad officers within my department. They seem, they're not very supportive. I don't find it was a very supportive working environment. In some ways, it seemed a little bit more of a business than anything else. I think overall, which is unlike other departments I suppose; I didn't meet anyone in my department. I know some departments have some stuff like Christmas parties or orientation for their grad student. My department didn't do any of that. The grad chair, although now it's changed, the impression was that he was just in that position because no one else wanted to do it. TA ships were a problem. There's too many students for that many TA's. The process of who gets chosen is not very clear. I was never given one because of my relationship with my supervisor and my department. There some issues there. So yeah lots of things. There doesn't seem to be as much support, I mean everyone likes grad students – it's like a nice cause but when it comes down to actually support them it's another issue

7.3.4 Saliency of the Survey Topic

Following the cognitive analysis of the cover letter, participants were asked whether they would be inclined to answer the survey. The objective was not to record whether they would in fact answer the survey. Hypothetical questions within the framework of cognitive interviews are known to have little validity (Willis, 1999). Rather, the point was to get the participants thinking about their decision to participate in the survey, and to verbalize how they arrived at their decision. Their rationale for choosing to participate in the survey or not should provide some clue whether topic saliency factors in their decision. According to Groves (2000)'s leverage-saliency theory, participants will base their decision to cooperate on various features of the survey, positive and negative, that stand out the most. For topic saliency to have nonignorable consequences, participants must base their decision on how satisfied or dissatisfied they are with their graduate experience. The affinity between the survey topic, alumni satisfaction, and their recalled experiences should be the overriding factors, above and beyond peripheral features of the survey such as questionnaire length, or lack of incentives. The analytical objective is to record the various salient features of the survey that come into play during their assessment of whether or not to participate.

A sample of participants' rationales, shown in Table 7-6, shows that they flow from the analyses presented earlier. The motivation to read the cover was low; the ability to understand the meaning of the word "experience" was equally low. It should be no surprise that the inclination to cooperate with the survey was generally moderate. The rationales drew from several sources: alumni eight drew from the "scarcity heuristics" (Groves & Couper,

1998: 33); alumni seven drew the norm of reciprocation; alumni five and two drew their rationale from the opportunity-costs perspective of social exchange.

Table 7-6 Rationale for Survey Cooperation Verbal Responses

Alumnus 8	I generally appreciate being asked for feedback cause I'm often frustrated if I do something and I feel that I have something that I like to comment on that I don't know necessarily where to direct it. So I personally would definitely take the time to fill this out if I received it.
Alumnus 7	Yeah, I think hum, I think yeah I have studied four years so I think maybe it's necessary for me to give an answer and [inaudible] department of Graduate Studies help to improve their service
Alumnus 5	Hum, I really it really depends if I'm in a good mood or not right? If I'm very busy I probably ignore it but if I don't have much to do…I mean… Well.. Why not?
Alumnus 2	[Grabbed the questionnaire] I'm flipping through to see how long it is. I'm taking a look at some of the questions to see what they would want to me to answer. Basically how much time it's gonna take me although that would change if I had nothing to do I would probably sit down and do it anyways but if I just came home from school and I had to run off and do a million things it would probably be shove in a pile somewhere and I would probably never find it or do it. Right now I would probably do it if I had nothing else to do but that would be the determining factor is what I was doing at the moment that I came at my door
Alumnus 3	I think the base problem with the survey is that there's really no incentive to fill it out. So it's like if people don't – I don't know – like see I would probably be more inclined to fill it out if I had a bad experience, if I had a bad experience I would be more likely to fill it out in order to complain. But other people might be like "I'm so glad to be out of that place. I don't want to read through any more UW stuff or do anything for them" [if you had a good experience would you fill it out?] Hum, I don't know if it was exceptionally good. If it was just about average probably would not. I think that's the big thing the less work they can make it for people or they can give something for that for people, a draw or something like that. Just some incentive for filling it out.
Alumnus 1	Yes! I imagine that most of my opinions would be voiced by others just in case something would be missed I would have to. So I don't think they would miss them but in the oft chance that they would I would fill out the survey. It would through written opinions. For the non-written opinions, just for checked boxes it's useful data I think these are useful data. I would submit it anyway.

Unexpectedly, however, alumni three, who expressed dissatisfaction earlier, wished for incentives. Following the exchange perspective, it could be hypothesized that

dissatisfaction costs. If the university wishes to collect information from dissatisfied alumni, the latter seem to expect compensation. Alumni one, who recollected very positive experiences, might be drawing from the organizational citizenship behaviour; the rationale, distinctively framed in terms of the organization's interest, "I think these data are useful", points to reciprocation based as much on helpfulness, as to identification with the university. Alumni one and three, taken together, seem to corroborate our findings in the previous chapters: GSES, which does not offer any incentives, may be prone to a positive satisfaction bias.

7.3.4.1 Saliency of Questionnaire Subtopics

It is common for survey researchers to judiciously sort survey sections and questions in such a way as to place the most salient, the most appealing and the less threatening topics first. This practice exploits the allure of topic saliency to entice the sampled individual to respond to a survey request. The question, for this thesis, is whether the alumni who may have had glanced at the questionnaire, would have altered their decision to cooperate to the survey. In an attempt to revisit the issue of topic salience and questionnaire access, participants were asked to quickly leaf through the questionnaire, scan the section titles, avoid inspecting the individual questions, and indicate which two sections are the most important and relevant to their graduate experience at Waterloo. If the survey (or portions of it) is indeed salient to their experience, participants should have no problem in ranking, themselves, the sections of the questionnaire.

Figure 7-1 Saliency of Graduate Student Survey Sections

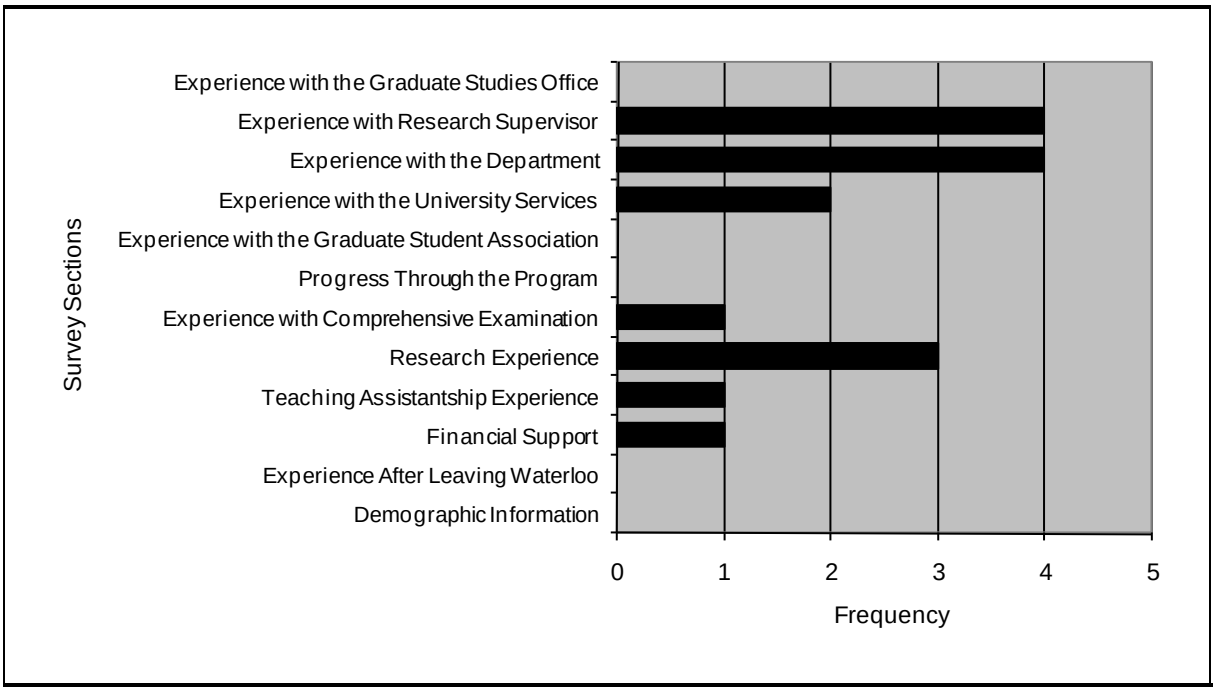


Figure 7-1 reports on the frequency of each section. Clearly, perhaps not surprisingly, experience with research supervisor and experience with departments are the two sections most cited. These two sections undoubtedly are salient in participants' mind because these two aspects of the graduate experience constitute the bulk of students' interactions with the university. Other sections of importance are research experience and teaching assistantships. The choice, as alumni eight pointed out, was based on degree program relevance: "[…] I tend to feel more strongly about filling things out that have to do with my degree as opposed to whether or not I use the library". The remaining sections, such as socio-demographics, graduate studies office, or experience after leaving Waterloo were perceived as helpful but not relevant to them specifically. Thus, the questionnaire is not specifically focused on the alumni's immediate concerns; its "topic" covers several dimensions of the graduate

experience that are not salient to the alumni. There is no reason to believe, unless one is specifically biased towards the department and the supervisor, that glancing at the questionnaire will necessarily cause nonignorability.

7.4 Topic Saliency and Measurement Error

In the previous section, we have explored the cognitive aspects of answering a survey, as participants were introduced to the survey material generally and to the cover letter in particular. This exploration did not address what might occur if alumni were subjected to a concerted effort, on the part of the GSO, to cooperate with the survey request. As indicated in chapter two, efforts aimed at increasing response rates from a population for whom the survey is generally non-salient yielded answers largely contaminated by measurement error. Answers, taken cumulatively, were either infected by a large dose of response bias such as anchoring and acquiescence bias, or infected by response variance due to satisficing cognitive strategies. In both cases, the issue, as was shown in Figure 2-7, is the inability to detect the presence of nonresponse bias in the survey variables of interest, when the latter are themselves biased by measurement error.

This section will examine the participants' cognitive responses as they interacted with the survey questionnaire. After they were instructed to read the cover letter, recall their graduate experience and rank survey topical sections, the second phase of the cognitive interview began by instructing the participants to begin filling out the questionnaire. While they were answering each question, they were instructed to think aloud and to note any questions that might be difficult to understand or too threatening to answer. The verbal

responses and the highlighted questions will be used to assess for the presence of measurement error in terms of response bias and response variance. We will begin with an assessment of measurement error related to response rate, to be followed by measurement error due to randomness.

7.4.1 Nonresponse and Measurement Error Model

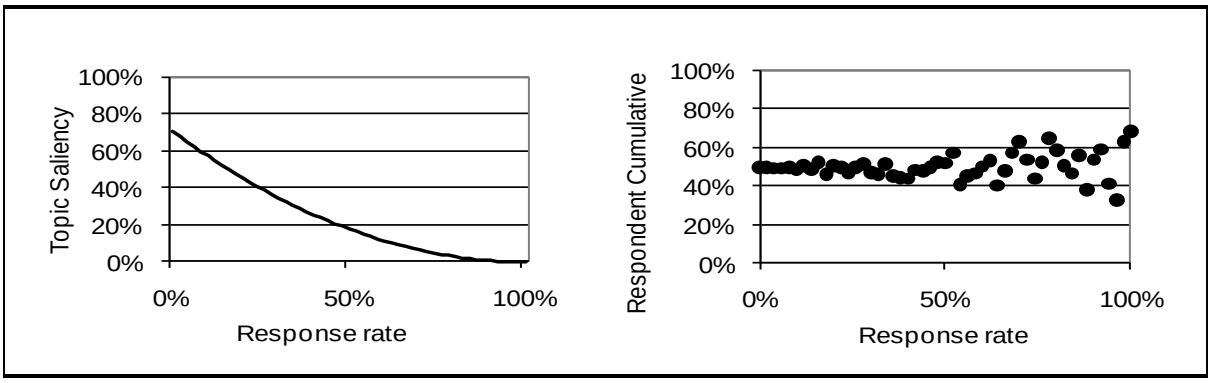
The nonresponse-measurement error[67] model provides an explanation for the presence (or absence) of nonignorable bias caused by measurement error as opposed to the product of true changes in the variable of interest. It posits that nonignorability is likely to occur if an increase in response rate is also accompanied by an increase in measurement error. Two types of relationships will be examined. First, we will examine whether an increase in response rate is associated with an increase in response variance; second, we will examine whether an increase in response rate is associated with response bias.

7.4.1.1 Response Variance and Topic Saliency

The introduction of response variance associated with an increase in response rate is likely to occur if the cause of nonresponse was originally based on topic non-saliency. As reluctant sampled individuals are incorporated in the response pool through various forms of incentives, they are likely to answer using various satisficing heuristics, or to offer quick top-of-head opinions, or non-attitudes (Converse, 1970). The resultant answers, taken cumulatively, will harbour ever more levels of response variance; measurement errors are likely to follow a heteroscedastic pattern (see Figure 7-2).

[67] See measurement error type I on page 27, and Groves (2006b) nonresponse-measurement error model

Figure 7-2 Anticipated Heteroscedastic Noise Level with Increasing Response Rates



One mechanism through which response variance may be linked to response rate is the relationship between topic saliency and question generality. Verbal reports of cognitive interview participants for the general question "satisfaction with methods of student evaluation" (see Table 7-7) revealed that the capacity to evaluate was contingent upon topic saliency. Participants who were able to recall a vivid experience from the stimulus of the question were able to confer meaning to the question; alumni six, eight and nine were capable of attributing meaning to the question, recalling a vivid element of their degree program, and mapping their answers onto the satisfaction scale provided. Conversely, participants who could not recall any experiences were unable to affix a stable meaning to the question itself. Alumnus #1, while grasping at straws to affix meaning to the question, nevertheless answered correctly in selecting the category "no opinion" -- preventing the introduction of non-attitudes. Alumnus #3, also fishing for meaning, selected dissatisfied. Alumnus #2 bypassed the need for meaning altogether; satisfaction was derived by the outcomes of evaluation methods. From these responses, we may submit the hypothesis that general questions amplify the distinction between saliency and non-saliency: those who find the topic salient will experience no difficulty in inferring meaning from general questions;

those who find the topic non-salient will quickly reach for heuristics to make sense of, or otherwise answer, a general question.

Table 7-7 Verbal Responses to "Methods of Student Evaluations"

Alumnus 1	I don't know whether that refers to. I guess it refers to a combination of things: performance as a TA if one is TA. I've never seen an RA evaluation form but there could be such this. Evaluations of student progress through a thesis program. Now computer science has a termly report which must be completed. I didn't have to do a termly report. I had to do one report at the end of the first year, and then after the fifth term I would have to do a termly report but I don't have to that know – or I didn't have to do that because I finished after five terms. […] So I because don't know what the questions refers to I have to say no opinion.
Alumnus 2	I'm satisfied. They passed me so. I'm pretty happy with it.
Alumnus 3	Yikes! I don't even know how they do that. I know we have to fill out like a progress report or something every so often but like we submit it and then nothing really happens. I don't know. I never had got any … you only get any feedback if it's negative like where they're like you've been here too long we need you out. Yeah and I don't really know … they supposedly use these progress reports to come up with internal scholarships, although no ones really quite sure how they decide […] I don't like that question because I don't really know … we have a progress report but it's like not really a method of evaluation because you're not really …
Alumnus 6	For this what I know is (pause) what I'm getting clearly is the method. For me, you know. I think when I did my master degree I thought my mark was quite well […] I also applied for some of the scholarship because I'm committed, I'm good. But they were no results after that. I'm sure maybe they had students better than me. But for me I don't know exactly the method for evaluation for the student performance. It depends on their mark? or depends on their research, the area?, what they've done? What have they've used for knowledge? Didn't know this method clearly for me.
Alumnus 8	I would say dissatisfied because the only feedback you get is marks and courses. We don't actually even have to fill out evaluation sheets; or what do they called, progress reports? Because I know that in some departments its every term you fill out progress reports and sometimes you get feedback from the department. But there's absolutely none of this. So I think that could be a problem. I think that sometimes things kind of (pause) slip a bit when there's nobody keeping tabs. When the students (pause) when they're only paid attention when you're over the deadline, which is a bit late I think.
Alumnus 9	There are the yearly review which it depends on the division of psychology you're in that will determine actually if you're getting feedback or not -- and that is just that "we're satisfied", or "you're progress is being satisfactory or acceptable". No one really gets any feedback. So I'll say dissatisfied with that.

7.4.1.2 Response Bias and the Zone of Tolerance

A second issue, introduced in 7.4.1, is the introduction of response bias for an increase in response rate; as we have already alluded, strategies aimed to increase response rates amongst nonrespondents who have found the topic non-salient were linked with the presence of acquiescence bias. As noted earlier, the literature shows that as more and more sampled individuals are brought into the respondent pool, respondents tend to agree more readily with agreement questions, and to report being "satisfied" to satisfaction questions. As a result, an increase in response rate is likely to peak the distribution around the "agree" or "satisfied" end of the scale. If respondents who found the topic salient also reported satisfaction, no bias would be found.

The detection of nonignorable bias in satisfaction scores is rendered inherently difficult due to the qualitative nature of the variable. To be "satisfied" or "dissatisfied" depends ultimately on how the question is formulated, and how the respondents confer meaning to the categories (Foddy, 1993). On very salient questionnaire items, shown in Table 7-8, participants did not experience any difficulty in expressing what each satisfaction score meant for them. They were able to provide a cogent rationale as to why their choice rested on a particular rating as opposed to another. The rationale provided the evaluative background in the selection of a satisfaction score. What is of note, however, is the distinctive clustering and spacing between each satisfaction score. Indeed, upon closer examination of the verbal response of these two participants, it became clear that being "very dissatisfied" is synonymous with sheer and utter incompetence, whereas being "very satisfied" signifies above expectations. The category "dissatisfied" seems to harbour

boundaries that are more diffuse: “I would hate it” or “a few key moments in there with no problem”.

Table 7-8 Verbal Responses to “Finance Office” and “Admissions Procedures”

Alumnus 2	I found the finance office people the most frustrating and arrogant people in the world to deal with because they treat you like crap and they don’t give a break and they don’t consider your time important at all. And so I would put that I was … I don’t know that I would put very dissatisfied because I guess there were some terms that things went OK. So I think I would just put dissatisfied. [what would very dissatisfied mean?] Very dissatisfied to me would mean that nothing ever went right and if I had to bounce a check because of them which actually almost was going to do but didn’t. I had a few good occurrences with them but my overall opinion was most of the time, when I walk in there, I was never treated like I mattered and that bothered me. So I put I was dissatisfied. I would put very dissatisfied if everything always went wrong. Dissatisfied means there were a few key moments in there with no problem.
Alumnus 5	I don’t know how to fill this one out. I wouldn’t say I was satisfied but I wouldn’t say I was dissatisfied either. It’s in between. It’s not good but it’s not like so bad that I would say I’m dissatisfied. [what does dissatisfied means to you?] Dissatisfied means that I was totally (pause) I’d hate it. I’d hate admission procedure. Satisfied would be I was very glad. Everything went so smoothly. Very satisfied would be extremely good. Very dissatisfied means it’s a complete disaster.

It is precisely this diffusiveness of meaning that provides room for the existence of a zone of tolerance, elaborated in chapter two. Within this zone, the evaluation of satisfaction remains unchanged for variations in the quality of given service or program. What the verbal responses reveal is that the zone of tolerance is created out of a leniency bias. As Table 7-9 indicates, when asked whether their supervisor returned their work in a timely fashion, participants moderated their evaluation with the context of the situation, or the personality of the supervisor. Although participants may not have agreed completely with the statement, they have biased their answers towards the “agree” end of the scale. This finding is consequential in light of the acquiescent bias found in the literature. If respondents, despite

finding the topic salient, exercise leniency or "even-handedness" (Foddy, 1993) to the extent that it biases their answers towards the "satisfied" or "agree" end of scale, their responses will be indistinguishable from reluctant nonrespondents who base their answers on acquiescent bias.

Table 7-9 Verbal Responses to "Returned my Work in a Timely Manner"

Alumnus 2	I guess I agree cause he knew if it was important he would get it back to me when he… when it was needed. But he wasn't the quickest bunny in the forest. Sometimes things took a little longer than I think they should of.
Alumnus 5	Ah that one I would not say agree – between agree and disagree. Sometimes it's very good, sometimes when it's busy, it's very late. Again I don't know where to put it. I probably put it as agree. It's understandable, I put agree.
Alumnus 3	That question is a little bit vague some people take courses with their supervisor so if I was going to say returned my work in a timely manner during you know when I'm in a course, that could mean one thing or returned my work in a timely manner referring to my thesis or publication drafts it could mean something else. Although I would have to say overall my supervisor is generally pretty good with that stuff so I'd say I'd agree

If alumni attributing saliency and non-saliency to a question respond in a similar fashion, albeit for different reasons, there was one reported instance where saliency did in fact introduce measurement bias. The challenge with general questions, where saliency is concerned, is to distil an array of clearly recollected experiences into one single answer. As was shown in the examples above, under some circumstances, salient experiences can be summed together to answer meaningfully a single general question; it is equally probable, however, that the respondent may not be able or willing to do so. Alumnus #1, in response to the question on "student housing", recalled two very salient aspects that could not be reconciled with the generality of the question. A substantial degree of cognitive burden resulted because the participant could not map his answer to the Likert scale provided;

answering satisfied or dissatisfied would have represented one aspect of the student housing experience to the detriment of the other. Unable (or unwilling) to reconcile the conflict, alumnus #1 chose to answer "no opinion" despite having a clear opinion about student housing. In this case, the generality of the question introduced a measurement bias: a salient attitude went unreported.

Figure 7-3 Verbal Responses to "Student Housing"

SECTION 4: EXPERIENCE WITH THE UNIVERSITY SERVICES

❶ **In summarizing your experience at Waterloo, please indicate your opinion of each of the following using the scale shown below.**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
b. Student housing	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
GSES 2000-02 Responses:	7.1%	21.8%	9.6%	4.1%	38.1%

Cognitive Interview Responses	
Alumnus 1	The question (…) is not detailed enough. There are many aspects to students housing which aren't addressed here. So it's hard. I'd put something between satisfied and dissatisfied if there were something there -- because my experience in residence with the other persons in residence has been mostly positive but with respect to the administration it has been largely negative. So it's hard to say. I can't say no opinion. That's just one that I would mark. This is one for which I would actually cross out the other services and put a comment in for part b [referring to 4.1b] because there so much I could say about it. So I don't know what I should even answer here. I guess no opinion although I have a really strong opinion which doesn't fit.

A closer look at the responses provided by alumni during the convocation 2000-02 (Figure 7-4), illustrates the difficulty in assessing this particular kind of bias for the study of nonignorability. If, following the example provided above, the category "no opinion", registering 38.1% of the respondents, is contaminated with alumni who had an opinion, one can only presume that those who offered an opinion, did so from the summation of

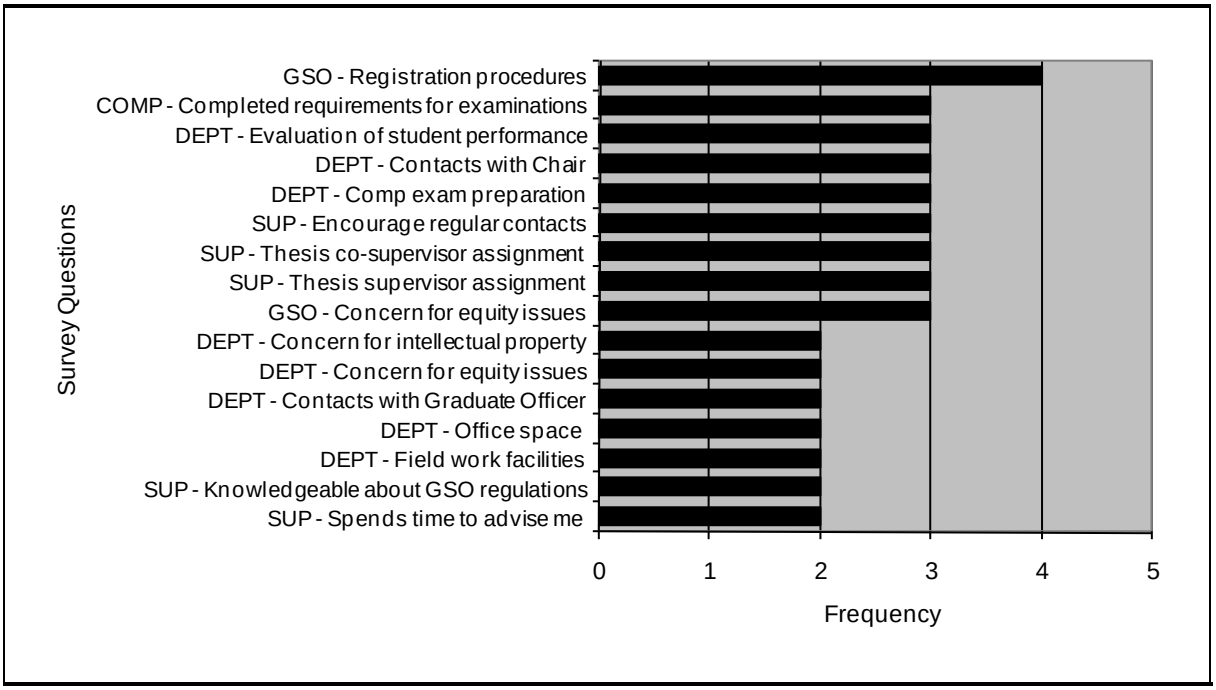
experiences that led to a satisfaction or dissatisfaction score. If the resistance of alumni #1 stemmed from an unwillingness to let go of the less savoury aspects of the student housing experience, then we might speculate that a "no opinion" bias is tantamount to not disclosing one's dissatisfaction, and consequently results in a satisfaction bias. Of course, this speculative hypothesis presupposes that respondents did offer a genuine attitude to an admittedly all-encompassing general question about student housing. Moreover, it also assumes that those who commissioned the survey, alumni who answered this question, and those who will interpret the answers are all on the same page regarding the meaning of that particular question. Issues of measurement bias and measurement variance often occur simultaneously.

7.4.2 Nonresponse Error Dilution Model

Regardless of the response rate, or whether the topic is of interest to potential respondents, if the variables of interests are badly worded, ambiguous, or otherwise too general, the ability to detect nonignorability will be thwarted before any survey request is made. Indeed, if respondents, despite their best efforts, cannot comprehend, recall or evaluate the question at hand, they are likely to misconstrue the meaning of the question or to infer the wrong context/topic. In any case, they are unlikely to devote the necessary cognitive energy to unpack these questions, will utilize various satisficing strategies (Krosnick, 1991), and as a result will introduce response errors. If these response errors are sufficiently large, they will submerge or rather attenuate valid answers on the variable of interest by increasing the response variance. This increased variance, or random noise, will mitigate against the possibility of detecting nonignorable bias. It will no longer be possible to compare valid

answers across response rates; the likely scenario, under increased response variance, is the finding of no bias, not because nonignorability does not exist, but rather because it can no longer be detected.

Figure 7-4 Frequency of Ambiguous Questions Signaled by Participants



During their interaction with the questionnaire, participants highlighted a number of ambiguous questions that stand the risk of introducing response variance errors. A glance at Figure 7-4 reveals that a good number of these questions, important for the university, simply have not been understood. This section will examine why these questions were perceived as ambiguous, and how they may have an impact on the detection of nonignorability. The section begins with an examination of question ambiguity in terms of word/meaning comprehension; this is followed by a discussion of the presence of ambiguity in general

questions; the section concludes with an examination of ambiguity due to poor content validity.

7.4.2.1 Ambiguous Organizational Terminology

A classic example of a question deemed important to an organization but posed to the alumni with all the comprehensibility of bureaucratese is the concern for equity issues. In bureaucratic parlance, the organization wishes to know whether the alumni, through personal experience, peers, or familiarity with organizational policies, perceived a genuine concern from the university for fairness and inclusivity towards minority groups writ large -- concern understood, presumably, in terms of organizational policy scope and application. Negative perceptions of fairness and inclusivity amongst the alumni population would undoubtedly be understood by the organization as a signal warranting further inquiries. Negative perceptions are also likely to be salient to the alumni, if only in terms of discrimination mentioned in chapter two, and as a result may be a source of motivated cooperation behaviour. Indeed equity issues could easily be one of those questions where we would expect, if discrimination did occur, nonignorable bias.

And yet, the majority of participants in our cognitive interviews singled out this question as ambiguous. In fact, they had no idea to what it referred, nor did they have a precise idea of its meaning. This perhaps explains the particularly mystifying GSES 2000-02 survey results, shown in Figure 7-5. To both questions, 36% of the respondents seemed to have been “satisfied” with equity issues; some, 12% to 14%, were “very satisfied”, and fewer still, 3.1% to 6.2%, “dissatisfied” to “very dissatisfied”.

Figure 7-5 GSES 2000-02 Response Distribution on "Equity Issues"

SECTION 1: EXPERIENCE WITH THE GRADUATE STUDIES OFFICE (GSO)

❶ **In summarizing your experience at Waterloo, please indicate the extent to which you are satisfied or dissatisfied with each of the following statements as they refer to your experience with the Graduate Studies Office (GSO):**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
g. Concern for equity issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
GSES 2000-02 Responses:	14.5%	36.1%	2.6%	0.5%	32.3%

SECTION 3: EXPERIENCE WITH THE DEPARTMENT

❶ **In summarizing your experience at Waterloo, please indicate your opinion of each of the following using the scale shown below.**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
t. Concern for equity issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
GSES 2000-02 Responses:	12.3%	36.0%	4.2%	2.0%	33.1%

* Not Stated/Skipped percentages not indicated: 14% for question 1.1g, and 12.4% for question 3.1t

The reported answers to "equity issues", salient to the organization, are likely to be muddled. Four concomitant response processes might be at work: 1.) respondents understood the question, had real attitudes and responded accurately; 2.) respondents had no idea what the question meant, and answered accurately by using the "no opinion" category; 3.) respondents had no idea what the question meant, inferred a meaning, and produced an opinion despite the uncertainty; 4.) respondents had no idea what the question meant, could not be bothered to infer a meaning, and selected their answer based on contextual clues or by choosing the "center" of the scale, that is to say "satisfied". The third response pattern introduces a response variance; the fourth response pattern introduces a response bias towards "satisfied". Consequently, the first response pattern, the real answers sought by the

survey, are likely to be attenuated by noisy "non-attitudes" on the one hand, and likely to be incorrectly peaked towards "satisfied" due to response bias, on the other.

7.4.2.2 Ambiguous Referents

The question that posed the most problems, and was signalled as the most confusing, contained the terminology "registration procedures". Again, from the point of view of the organization this question should not have posed any real difficulties. Alumni had to register through the registrar's office when they first enrolled in their degree program, and when they selected their classes for each term; both types of enrolment embodied a set of registration procedures that should have been recalled by the alumni. The generality of the question was undoubtedly deliberate to get the gist of the experience. And yet, the participants, as Figure 7-6 indicates, had no idea to what the terminology referred. Alumnus #1 found an aspect of registration procedures salient, but most did not.

The issue is not ambiguity in terms of incomprehension of the terminology as was the case with the words "equity issues"; the issue is ambiguation due to multiple meanings housed under one expression. If we were to return once again to the responses provided by the alumni GSES 2000-02 survey, the majority (57.5%) seemed to have found these procedures satisfactory. If this particular question posed as much difficulty as the participants in the cognitive interview seem to have suggested, the correct answer should have been "no opinion"; only 3.1% of the alumni population availed themselves of that option. Either the participants of the cognitive interview unduly focused on the question format and, as a result, missed the question's basic meaning, or the GSES respondents answered satisfied without

really knowing what the question meant. The lack of a clear referent that would have disambiguated the question is likely to have produced a satisficing heuristic: respondents either based their answers on the previous answers (contextual bias), on the likert scale, on the need to be consistent, or simply by forging a quick opinion (non-attitude). In all cases, the response outcome will be unreliable answers, and consequently increased variance in satisfaction scores.

Figure 7-6 GSES and Participant Responses on "Registration Procedures"

SECTION 1: EXPERIENCE WITH THE GRADUATE STUDIES OFFICE (GSO)

❶ **In summarizing your experience at Waterloo, please indicate the extent to which you are satisfied or dissatisfied with each of the following statements as they refer to your experience with the Graduate Studies Office (GSO):**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
f. Registration procedures	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
GSES 2000-02 Responses:	18.6%	57.5%	14.0%	4.6%	3.1%

Cognitive Interview Responses	
Alumnus 1	I don't like Quest. Does the registration procedure deals with course registration or term registration because there is a bit of a difference? [I don't know] Ok I'll assume for the sake of this that the registration procedures are for mainly course registration and Quest is unhelpful in some way.
Alumnus 5	Registration procedures? I found that confusing. I don't what that means. Does it have anything to do with the GSO at all?
Alumnus 8	I'm not sure. I'm assuming they're talking about every term? When you're going to make sure you're registered as a student? Or is this when you're registering for your program? It's one thing I'm not sure what there asking me there.
Alumnus 9	I don't know what they are. I'll put no opinion because I have no idea what registration procedure are.

* Not Stated/Skipped percentages not indicated: 2.2% for question 1.1f

A classic method to ascertain the degree of reliability of a given indicator/question is to pose an identical question using a different format or a different scale. Following Sudman, Bradburn and Shwarz (1996), if the question topic is salient, answers should be consistent despite changes in question form. General, but salient questions, should elicit (through recall) a stable set of referents from which questions are evaluated and ultimately answered. One such question was available on the GSES questionnaire, and was used to ascertain the degree of reliability of answers given to the supervisory process. The question on "quality of research supervision" posed in the department section of the questionnaire was cross-tabulated with agreement clusters for supervisory experience. The two question sets are of a different format, placed in different sections of the questionnaire, and under different scales. Participants in the cognitive interview, as Figure 7-7 indicates, clearly understood the "quality of research supervision" as pertaining to the supervisor. Alumni #1 says it succinctly, "I have already have gone at length about that".

Respondents to the GSES 2000-02 survey, however, seemed to have had different ideas. Again, as shown in Figure 7-7, 83.2% of the respondents reported being very satisfied to satisfied with their research supervision. When these responses are cross-tabulated with agreement clusters however, one can clearly distinguish instability amongst satisfaction scores. Of the respondents who were in general agreement with their supervisor's role performance, the majority (53.7%) replied being "satisfied"; some, however, were "very satisfied" (36.9%) and even dissatisfied (8.3%). The largest instability is to be found amongst those who were in general disagreement with their supervisor's role performance. Paradoxically, 31.0% of these individuals were essentially satisfied with their research

supervision. Worse still, are those who didn't have a supervisor (missing/not applicable), and yet, 51.6% of these individuals managed to report a satisfaction score.

Figure 7-7 GSES and Participant Responses on "Quality of Research Supervision"

SECTION 3: EXPERIENCE WITH THE DEPARTMENT

❶ **In summarizing your experience at Waterloo, please indicate your opinion of each of the following using the scale shown below.**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
e. Quality of research supervision	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
GSES 2000-02 Responses:	41.0%	42.2%	9.9%	3.6%	3.2%
GSES 2000-02 Satisfaction with Supervisor Clusters					
Agree	36.9%	53.7%	8.3%	0.6%	0.6%
Strongly Disagree/Disagree	0.0%	31.0%	40.2%	26.4%	2.3%
Strongly Agree	66.5%	30.0%	1.7%	0.0%	1.7%
Missing/Not applicable	10.3%	37.9%	3.4%	0.0%	48.3%

Cognitive Interview Responses	
Alumnus 1	Very satisfied. I have already have gone at length about that
Alumnus 2	I was satisfied with that. I wasn't very satisfied because it would have been nice to have been a bit more supervised but overall he (supervisor) was good to me
Alumnus 3	I would say very dissatisfied mainly because quite honestly… in terms of providing resources, my supervisor was excellent […] in terms of contributing intellectually to my thesis, it was pretty minimal so that why I'm saying pretty dissatisfied.
Alumnus 8	I guess I would have to talk about my main supervisor here and say very dissatisfied
Alumnus 9	I already just talked about that! In section 2, question 2, I would have thought. So I'll go on with satisfied.

* Not Stated/Skipped percentages not indicated: 12.4% for question 3.1t

Admittedly the variance found on the strongly disagree cluster may have been related to non-mapping attitudes between agreement and satisfaction scales. It is conceivable, for example, that respondents may disagree with the supervisor's basic handling of the supervisory process without necessarily judging their experience of that process as

unsatisfactory. That may account for some of the variance towards the satisfactory end of the scale on the research supervision question, but certainly not all. In sum, general questions, despite eliciting salient experiences for the alumni, did not yield stable attitudes; respondents were sensitive to the question format. Moreover, the more negative the experience, the greater the sensitivity to either the question format or the generality of the question. There may be a combined effect: respondents, faced with an ill-defined general question, and having recalled a negative experience, may not be able to contextualize the meaning of the question properly, and, as a result, answer unreliably. Said differently, it is easier to be satisfied or to be in general agreement with a question posed in general terms, then being dissatisfied – unless, of course, one is thoroughly and utterly dissatisfied. If this argument is correct, general questions are biased towards the production of satisfaction, and dissatisfaction responses are attenuated by randomness.

7.4.2.3 Ambiguous Content Validity

As Foddy (1993) is at pains to point out, the "making of an intelligible request" is inherently a two-way communication process. It presupposes that initiators of the survey request can formulate their questions clearly. In turn, respondents infer meaning from survey questions and confer meaning to their answers using the scales provided. Finally, the survey analyst will infer meaning from these answers. If all goes well, initiators, respondents and analysts achieve intersubjectivity as to the meaning of both the questions and answers. We have covered, thus far, problems in achieving intersubjectivity because the initiators of the survey request used bureaucratese and/or made use of general questions to communicate their request. GSES 2000-02 respondents and cognitive participants revealed a third threat to

Figure 7-8 GSES and Participant Responses on Reasons for Not Returning to Waterloo

SECTION 6: PROGRESS THROUGH PROGRAM

❾ If you were to start your program again, would you select Waterloo?

☐ Yes

☐ No ⟶ Please comment:

GSES 2000-02 Responses

Category	N	Pct	Example
Prog/Teaching Quality	9	22.0%	Irritating admin, subjective grades…
Financial Support	6	14.6%	Lack of program funding, organization, clarity of protocol
Broaden Experience	5	12.2%	I have two degrees from Waterloo already…
General	5	12.2%	I did not enjoy my experience at Waterloo
Climate/Quality of Life	4	9.8%	I did not enjoy the atmosphere in the department…
Changes Interest	4	9.8%	I think I would want to take a different path with my studies
Course Selection	3	7.3%	I am highly disappointed by the limited selection…
Advisor/Supervision	2	4.9%	Found that there was no help when relation with sup. soured
Location	2	4.9%	The city did not really have much to offer in terms of rec.
Not Stated	1	2.4%	
Total	41	100.0%	

Cognitive Interview Responses

Alumnus 1	I still would do it here. If I knew that I could do this thing called geometric algebra I'd do it here because nobody in Canada does it within computer science - almost nobody else in the world actually within computer science does this work. I would actually, (…) one of the main reasons I'd want to stay in Waterloo to do a PhD as well.
Alumnus 3	Ah … hum … I mean in the end like you can't go back and do things but you know. I don't know I did learn things from another perspective but probably not. I wouldn't because the problems of funding, the problems of department, and supervisor support, supervisor environment. I think the problems of grad studies in general that graduate students are often seen as second class citizens in undergrad institutions
Alumnus 9	Yes, it worked out very well for me. So no regrets whatsoever
Alumnus 2	Right now? I would say no but the only reason is because now that I'm doing a PhD I don't like the idea that I did all three degrees at the same place. I would never change my engineering degree from here (…) and I would never change my PhD from here because the guy I'm working with is fantastic. So I probably would have liked to have gone to a different school in the middle

achieving intersubjectivity. A more insidious form of ambiguity stems from requests that appear intelligible to all parties involved but are ultimately without meaning because a single question, seemingly designed to measure a single attitude, may in fact measure two or several different attitudes at once.

This was the case for another question, salient to the university, pertaining to the alumni's overall impression of their degree program. Ultimately, despite any difficulties encountered during their degree program, the university may wish to know whether the services and programs provided were, from the alumni's perspective, of quality. One of the ways to get at this information is to ask whether the alumni would return to the same university. The GSES questionnaire taps into this dimension by asking plainly: "if you were to start your program again, would you select Waterloo again?". The answer is to be supplied using dichotomous "Yes" or "No" categories.

The responses collected from the GSES 2000-02 survey seemed encouraging: 77% of the respondents indicated their desire to return to Waterloo, 18% would not, and 5% declined to answer. For those who chose not to return to Waterloo, the GSES survey provided them with the opportunity to clarify. The answers from these respondents can be seen in Figure 7-8. The answers were grouped into themes and tabulated. Examples of answers are given in the right-most column. It can be observed that program and teaching quality is the greatest reason why they would not return to Waterloo. This evidence is supported by (Thomas & Galambos, 2004) findings that "faculty came to class well prepared" is the dominant reason for students' "satisfaction with college in general". Other factors are financial support and the desire to broaden one's school experience. Participants in the cognitive interview were

equally diverse: some would return to Waterloo because of the program or research expertise, while others would not return to Waterloo because of various difficulties encountered during their degree program or because of a desire for a change of scenery.

There lies the rub. Both respondents and participants encountered no difficulty in interpreting this question. Whether they would return to Waterloo or not, both were able to provide a rationale for their choices. For those who found the question salient, answers were given in terms of the service provided; those who did not find the question particularly salient expressed peripheral reasons for choosing or not choosing Waterloo. Whether salient or not, answering "yes" amounts to the same interpretation: the services provided were adequate or are still sought after. However, answering "no" harboured two very different explanations: "no, I detested my experience at Waterloo", or "no, I would have preferred another city to live in". The problem is immediately apparent if, instead of 77% of returning alumni, we imagine the analyst compiling 63% or 55%. How can these proportions be interpreted?

Two consequences flow from this fact. First, if one interprets this question as one of service quality, the danger is to misconstrue a "no" as indication of poor service quality tout court, when in fact, the "no, get lost" category, undoubtedly indicating irateness with the university, will be submerged and attenuated with the "no, thank you" category. Thus, while it may be possible to assert that 77% of alumni were, on average, happy with the services and programs provided by Waterloo, it can't be said that 18% were unhappy with Waterloo. The question flows in one direction only. Second, any findings of co-variations between response rate and the desire not to return to Waterloo must be interpreted carefully. The fact that, for

example, one finds an increase in the proportion of the "no" category with an increase in response rate, should not be seen as unexpected if the proportion of irate alumni is smaller than those who would have preferred another locale. The former, likely to find the topic salient, are also likely to have answered first; the latter, likely to find the survey less salient, are also likely to answer last. In sum, conceptually ambiguous questions lead to interpretative problems as to the meaning and significance if bias is found; they also lead to response variance in one or both categories.

7.5 Conclusion

The cognitive interviews, despite their relatively low numbers, reveal some troubling evidence against the continuance of organizational satisfaction surveys. First, even when expressly asked to read the cover letter, the results suggest that participants only superficially attended to its content. They recall only one or two events of their graduate experience. Their interest centered much more on interpersonal relationships and quality of life issues than role performance. Their experience at Waterloo is salient to them, but is not properly tapped by the questionnaire. The section on question order and topic saliency reveals that the GSES survey is perceived as an omnibus survey. It contains some sections that are relevant to the alumnus graduate experience at Waterloo such as supervisor and research experience, but it also contains many other sections that are not salient, seen as unappealing and even irrelevant. It was made quite clear on more than one occasion that one should not ask questions that are already known or accessible through other means. Finally, the generality of the questions throughout, most likely shaped to appeal to a wide audience, resulted in heuristic (satisficing) coping mechanisms that thwart the ability to detect nonignorability by

inducing response variance for some questions, and biasing answers towards the satisfied score for others.

Chapter 8
Ignorable Student Satisfaction Surveys

8.1 Introduction

This dissertation began with a concern for the accuracy of survey estimates. Declining response rates across a variety of different survey modes and topics, coupled with an increased reliance on surveys as a policy-making tool within organizations, has increased concerns as to whether survey results adequately represent their population. Following Rubin (1987)'s distinction between ignorable and nonignorable nonresponse, survey research has focused on the causes of nonresponse that are also systematically related to the variables of interest. Topic saliency was the concept of choice of researchers because of its strong linkages between survey cooperation and the variables of interest, the latter understood here as the survey topic.

The goal of this dissertation was to assess ignorability for a particular class of surveys, namely alumni satisfaction surveys. The impetus of our research was the intuition that satisfaction, being an often highly charged sentiment, might be conducive to generating nonignorable survey cooperation behaviour. Alumni satisfaction surveys are increasingly becoming, as much as graduation ceremonies, a perennial feature across campuses. In light of the declining response rates, the capacity of surveys to adequately represent the sentiment of a given alumni population is uncertain. Empirical findings have to date remained mixed: some studies have detected a dissatisfaction bias, others a satisfaction bias and, others still,

nothing at all. A multi-method approach was utilized to apprehend both the quantitative and qualitative aspects of ignorability, pertaining to both nonresponse error and measurement error. It was hypothesized that the alumni's organizational relationship, constructed by his or her graduate experience, will serve as the basis for deciding whether or not to cooperate with the alumni survey request.

This final chapter will begin with a summary of significant findings. It is followed by a discussion of the general hypotheses elaborated in chapter two in light of the thesis findings. Lastly the chapter concludes with the methodological and organizational implications of the findings for survey practitioners and public administrators, and recommendations for future research.

8.2 Summary of Results

The findings of this dissertation do not lend themselves to easy conclusions. Results were accompanied with confounding explanations that greatly weakened internal validity. Anonymity concerns during the implementation of the GSES foiled any possibility of differentiating non-contact from refusals. Since topic saliency is a cooperation-based behaviour, the effects of topic intensity could not be ascertained directly. In addition, restrictions in the public version of the NGS opened up unbridgeable problems of coverage, not to mention social desirability effects, mode effects and time-shifting effects. Results were also compromised by a weakened statistical validity mostly due to the quasi-exploratory aspect of the dissertation. Statistical validity was at times weakened by significance tests performed under low power conditions which prevented us from detecting nonresponse error,

and at other times, weakened by the number of significance tests that increased the investigationwise error rates, and as a consequence, introduced doubt as to the veracity of nonresponse errors that were found. Other methodological problems, mostly the product of analyses conducted on secondary data sets, limited our reach in terms of the types of analyses to be conducted. We were unable to access nonrespondents, which would have allowed more precise assessments of nonresponse bias; we were unable to conduct early-late analyses, which would have made use of all survey variables in the assessment of nonresponse bias. In sum, the results of this study are limited and tentative.

Nevertheless, the use of a multi-methodological approach did compensate for the limitations of one method with the strengths of another; it has permitted triangulation of results from studies of nonresponse conducted at different levels of analysis. In doing so, it has bolstered our ability to detect overriding trends in nonresponse behaviour, and how nonresponse has impacted nonresponse error. A summary of the thesis results will begin with a synthesis of trends surrounding nonresponse and nonresponse error. It will be followed by a determination of the ignorability status of nonresponse error.

8.2.1 Nonresponse and Nonresponse Error

The essential but insufficient condition for the emergence of nonresponse error is the observation of non-random systematic nonresponse patterns. These patterns, following Groves' model elaborated in chapter two, are a mixture of properties of the alumni, his or her context, the survey design, the survey administration, and the eventual relationship with the survey material in question. The second condition for the emergence of nonresponse error is

the observation of a systematic relationship between the variables of interest and variables that are either the direct cause of nonresponse, or that are indirectly affected by variables that are cause for nonresponse. Several variables that showed overrepresentation in the response pool were also related with alumni's satisfaction with their degree program[68]. Our dissertation focused on alumni characteristics, departmental characteristics, and survey protocol changes. A review of the evidence from preceding chapters indicates that nonresponse cannot be classified, by definition, as missing completely at random.

Women alumni, visa-registered alumni, and deviation from normal program duration were all overrepresented in the response pool, and were also related with satisfaction scores. Overrepresentation is not, however, consistent throughout convocation years, nor is it consistent across survey protocols (e.g. factorial groups). Women alumni seem to be consistently overrepresented across survey protocols, but not convocation years. The reverse is true for visa-registered alumni and deviation from normal program duration. Correlations amongst these frame variables reveal that visa male and Canadian female are overrepresented once all convocations are combined. This chaotic pattern may have its source in the mixture of contactibility and refusals. Science-related departments and degree program were both representative of the combined convocation population.

The impact on nonresponse error is equally complex. By and large, these three frame variables played only a minor role in predicting satisfaction scores. Visa-registered alumni

[68] It should be reiterated that over or under representation across population groups is not synonymous with causing nonresponse. However, when these population groups are entered into a multivariate model, the direct and indirect impact of their representational bias can be ascertained. That link between sample representation and the variables of interest is sufficient to meet the second condition aforementioned.

was positively but marginally related to satisfaction with supervisor. Deviation from normal program duration was positively related to dissatisfaction with supervisor, albeit substantively small in its impact. Women alumni were significantly and positively related with dissatisfaction with overall graduate experience and with supervisory experience. Taken together, the estimated adjusted satisfaction score will generally increase both very satisfied and very dissatisfied scores by one percentage point for logistic models based on all convocation periods (including random models), and by three percentages points across factorial groups (visa-registered alumni only). All these adjusted scores fell well within sampling error.

Contextual analyses, however, showed a disturbing trend that points to the presence of nonresponse error. Starting with comparisons between the NGS and GSES surveys, the former would seem to suggest that being very satisfied and being very dissatisfied are under represented in the GSES – a reversal of the original hypothesis of sample truncation. If we suspend the confounding issue of coverage, the comparison suggests a "satisfied" bias in the GSES[69].

[69] The significance of this finding is difficult to ascertain in light of an intractable coverage error. Comparisons between GSES and NGS on university student field of study codes (see NGS question A06FOSP) reveal several areas in which the databases differ significantly from each other: GSES has seven times the proportion of engineering and mathematics/physical sciences alumni than NGS, and twice the amount of social sciences alumni; GSES has little, if any, alumni with education, commerce, business administration, and technologies and trades degree programs. High proportions of engineering and physical sciences alumni also imply well-funded degree programs through the Natural Sciences and Engineering Research Scholarship (NSERC) amongst others. Neither the direct impact of financial support on satisfaction, nor the indirect impact of financial support on involvement and conference expenses funding are accounted for in the GSES-NGS model. Moreover, the model does not account for differences in overall reputation of the University of Waterloo, the overrepresentation of technologies and trades degree programs in the NGS population, self-selection with respect to returning alumni students, and in coop program enrolments. All these unaccounted differences will reveal themselves in the intercept of the NGS and GSES models. Thus, any significant differences between the

In addition, departmental response rates are negatively related to the probability of being very dissatisfied with department and overall graduate experience; and positively related to the probability of being very satisfied with supervisory experience. The effect persists despite controlling for departmental and alumni characteristics. Thus, for two alumni with identical values on socio-demographics and program-related, objective and subjective discrepancies, the one originating from a department with higher response rates will have a greater probability of being very satisfied with supervisory experience than the other alumni originating from a department with a lower response rate. Whether this relationship is spurious remains unclear. What is clear, however, is that an unknown contextual variable is co-related with the propensity of an alumnus to respond to a survey request, and with the probability of being very satisfied.

Equally troubling is the finding that some contextual variables known to be related with survey cooperation were also related with satisfaction clusters. Peer influence was significantly related to satisfaction with overall graduate experience; average grade point average to being very satisfied with overall graduate experience; and proportion of doctoral alumni to being very dissatisfied with overall graduate experience. Departmental engagement was significantly related with being very satisfied, and with supervisory and departmental experiences. While over or under-representation could not be ascertained on these contextual variables, the potential for nonresponse bias remains.

---

model intercepts cannot exclude the possibility of being the product of coverage error, and are not necessarily an indication of nonignorability.

Finally, variations in survey protocols in terms of mailing dates and reminder formats did not show any nonresponse bias despite between-group overrepresentation in women alumni, visa-registered alumni, Canadian alumni whose previous degree was not granted from Waterloo and alumni whose conference expenses were either partially or not funded at all. Results did show, albeit statistically not significant, that the July mailing with phone message (control group) was distinctively different than all other factorial groups. The probability of being very satisfied with supervisory experience was, against expectations, lower in the control group. This would suggest a potential source of nonresponse error associated with a specific combination of mailing date and reminder format.

8.2.2 Nonresponse Error: Ignorable or Nonignorable?

There is no doubt that respondents and nonrespondents are different. A definite relationship was observed between satisfaction clusters and response rates. It cannot be said that nonresponse is missing completely at random. The question is to ascertain if this observed relationship is spurious, that is to say missing at random, and thus ignorable, or whether this relationship is in fact nonignorable. The key issue centers on whether the dataset can accurately represent the alumni population if the relationship between variables remains constant throughout response rates. This is the basis of the form resistant correlation hypothesis: nonresponse will alter the marginals of a given variable but the relationship between a given predictor and a variable of interest will remain constant despite change in the composition of one or both variables. The results point towards ignorability (MAR), but with some noted reservations.

Nonresponse is essentially ignorable. First, women alumni, visa registration and deviation from normal duration despite significant nonresponse did not significantly impact the variables of interest. Second, the impact of subjective and objective discrepancies on satisfaction clusters did not vary significantly when compared with the NGS dataset, across departmental response rates, or between-factorial groups. Third, item missingness (missingness on socio-demographics), a major confounding factor in the determination of nonignorability, is positively related to being very dissatisfied with supervisory experience; item missingness (skipped departmental identification) moderates the variations of Canadian alumni whose previous degree was not from Waterloo on being very dissatisfied with supervisory experience. Fourth, flowing from the first three, logistic models drawn from these predictors poorly explained the initial "variance" present in satisfaction clusters; McFadden measures (R-square analog) varied from 12% to 38%, and the cluster classification measure, lambda, remained non-significant except for the model of dissatisfaction with overall graduate experience. Admittedly, low values in McFadden and lambda, particularly for satisfaction models, are likely to be the results of predictors singularly focused on discrepancies. Nevertheless, if the models were truly harnessing topic saliency, we should have seen stronger coefficients. Fifth, results from cognitive interviews have demonstrated that participants found the survey topic to be not salient; interviews have also revealed that participants' answers contained numerous sources of measurement errors that would foil any distinction between respondents and nonrespondents.

However, the results also revealed an isolated case of nonignorability. Comparisons with the NGS dataset showed that visible minority alumni who completed their degree

program later than normally expected scored higher in the GSES dataset than the NGS dataset. It would seem to signify that the odds of being very satisfied with the overall program are greater at lower response rates for visible minority alumni who completed their studies later than expected. At the same time, GSES visible minority alumni are also related to being very dissatisfied with their degree program; that relation becomes non-significant in the NGS. Unable as we are to disentangle all the confounding issues related to the NGS dataset, the status of this particular finding remains uncertain. The impact of visible minority on any of the satisfaction clusters did not vary across departments; it did not vary across factorial-groups either. This particular variable is important because it might relate to racial inequity issues within an organization, and thus must be subjected to further analysis.

Beyond this singular case of nonignorability, it should be remembered that the question of ignorability or nonignorability cannot avoid the omnipresent epistemological issue associated with the drawing of conclusions from individuals who, by definition, have not responded. Despite the various methodological stratagems utilized to infer characteristics from nonrespondents, they are, epistemologically speaking, unknowable. This unsettling fact signifies that one can never be certain that the response pool is truly representative of the population under study – unless one achieves 100% response rates. Failing that, there may still lurk an unmeasured predictor of nonresponse bias that will rear its head at the most inopportune moment – perhaps in the same vein as for the Literary Digest. In the absence of such variables, one is forced to conclude that the seemingly spurious relationship between response rate and satisfaction clusters is biased irretrievably, and thus to be considered

potentially nonignorable. The form resistant correlation may hold, but descriptive statistics cannot be corrected or adjusted.

8.3 Theoretical Explanations of Ignorable Behaviour

A second goal of this thesis was to provide an explanation of survey cooperation behaviour that might distinguish between ignorable and nonignorable nonresponse. Lacking in many of the empirical studies reviewed, and topic saliency research in particular, was a rationale that would bind survey cooperation behaviour to satisfaction with their alumni's experience at a given university. A review of the literature on the nature of satisfaction and topic saliency revealed that satisfaction-based response behaviour is intimately linked with topic intensity. To paraphrase Warren (1996), topic saliency occurs at the margins of satisfaction: we had anticipated that alumni who were either "very satisfied" or "very dissatisfied" with their graduate experience were most likely to find the survey topic salient, and as a result respond more readily; alumni who were simply "satisfied" were unlikely to find the survey topic salient, and would respond least. The net effect was a severe truncation of satisfaction scores in the middle of the distribution. This truncation would affect both descriptive and analytical statistics, that is to say both the marginals and the form resistant correlation. In essence, survey cooperation behaviour motivated by topic intensity was hypothesized as nonignorable.

The rationale for this satisfaction-based behaviour was derived from several theoretical outlooks on the modern societal relationship entertained by citizens to their organizations and institutions. Those individuals who entertain a meaningful relationship

with their organization are more likely to respond out of organizational citizen behaviour or out of organizational loyalty. The vast remainder, those who preferred an instrumental relationship with their organization, are likely to trust the judgement of the organization particularly upon receiving a satisfactory outcome; they are unlikely to supersede the judgement of public administrators with their own opinions and evaluations. In cases where dissatisfaction does occur, they will respond on the basis that the organization is trustworthy. In cases where dissatisfaction is understood as systemic, disaffection and alienation are likely to ensue; dissatisfied individuals are likely to become terminal nonrespondents. In sum, the bulk of the missingness comes from those individuals who essentially trust institutions[70].

However, the results of our study, contrary to our hypotheses, point to ignorability. Trust-based survey cooperation behaviours, anticipated to reduce participation, and thus truncate the satisfaction distribution in the response pool, did not materialize. In fact, if we are to believe the GSES and NGS comparison, the reverse seems to have occurred. A marginal increase in the probability of being very satisfied and being very dissatisfied was observed in the NGS, whereas it should, according to our hypothesis, have been observed in the GSES. Indeed, the same situation occurred in the July mailing with phone message. That particular factorial group, the control group, which should have been the most conducive to generating topic saliency effects, produced the same contrary behaviours: the probability of being very satisfied with supervisory experiences was lower in the control group. At best,

[70] It should be reiterated that trust is not always synonymous with approval. The concept of trust, as we have elaborated in previous chapters, is akin to Luhmann (1979)'s reduction of complexity. Although we may not always believe organizations to be deserving of our trust, we nevertheless do so because thinking otherwise would imply getting deeply involved in all aspects of organizational life – organizations as central to social life as the purveyors of food, health, education, clothing, etc.

there seems to have been a reversal of the hypothesis, but more likely, organizational-based survey cooperation behaviour, if it occurred at all, doesn't seem to have an impact on nonresponse error. Admittedly, we were never in a position to determine the exact nature of cooperation behaviour since we were not able to remove contactibility as a confounding factor. Nevertheless, it was hoped that a trend would emerge using four different methodological approaches.

With these new pieces of evidence, considering the methodological limitations aforementioned, we would submit five explanations for the finding of ignorability:

1. The conceptual background stipulating survey cooperation behaviour is incorrectly specified. The theoretical outlook assumed that the alumni had or have a "relationship" with the organization. Graduate students, whether enrolled in a master's or doctoral\ program, can easily spend a minimum of two years to five or more years within institutional walls. During that time, it was expected that a very satisfied experience would also translate into a meaningful set of relationships – relationships that would induce organizational citizen behaviour. According to Cropanzano & Mitchell (2005: 884)'s review of social exchange concepts, the emergence of OCB is susceptible to the "target of exchange"; while a meaningful relationship may have developed with a supervisor, it may not translate into a commitment towards the organization itself. As a result, OCB –based cooperation behaviours will never materialize. At the opposite end of the spectrum, the reporting of dissatisfaction might be more costly than originally anticipated. Surveys, such as the GSES, are, following Brehm (1993), meant to subsidize participation and equalize voices, to make it easier to report one's experience, to circumvent the burden placed upon the client/customer to escalate

complaints from one hierarchical organizational ladder rung to the next (Heskett et al., 1997). In theory, participation should have been effortless. Yet, the one dissatisfied cognitive interview participant estimated that compensation was in order; being helpful under conditions of dissatisfaction has a price. Beyond compensation, the reason why individuals do not complain is varied: they may not believe that the "organization will do anything if they do complain", they may fear that a "complaint will get a service provider […] into trouble" (Heskett et al., 1997: 180). More research is needed to identify reasons for a dissatisfactory experience that transcend the grids of bureaucratic processes.

2. The topic is simply not salient. The crux of the argument behind the theoretical outlook is the propensity to respond on the basis of topic intensity. It was hypothesized that organizational citizenship behaviour, loyalty, and to some extent complaint behaviours would be good candidates for the explanation of nonignorability. Moreover, as was alluded in the previous point, graduate alumni are topically involved. Despite graduation, one does not cease to be a "student" of Waterloo. To paraphrase Hirshmann (1970: 101) on public goods, there is "no escape from consuming them unless one were to leave the community by which they are provided". Yet, these basic arguments for the presence of topic saliency may have been foiled by the overriding bureaucratic interests that framed the GSES topic; what Benjamin (1994: 219-221) dubbed an "insider institutional approach". It was instructive to observe how little interest cognitive interview participants showed for the survey. They tended to ignore the cover letter; focused upon peripheral aspects of the surveys such as mailing stamps, questionnaire cover and length; and could only recall vague impressions of

their graduate experience. The evidence points to the severance of topic intensity, if present at all, from the motivation to respond.

3. The population was homogeneous with respect to the variables of interest. It is interesting to note that aspiration theory, what one wanted compared to what one had, is the strongest reason for alumni to be "very satisfied" with their respective degree program experience. In particular, the meeting of university expectations (e.g. reputation, location, etc) as reasons for choosing Waterloo was a significant predictor of being very satisfied with their overall program and with their supervisor. This relation did not vary across departmental responses rates or survey protocols. Enrolling, attending, and graduating from Waterloo seem to virtually guarantee being very satisfied – regardless of whether one is a respondent or not. The same could be said with respect to very dissatisfied alumni. Equity issues, such as authorship conflicts and income fairness, regardless of how others fared (social comparison), virtually guaranteed to being very dissatisfied with all aspects of their degree program. Whether one is a respondent or not, equity issues are sources of dissatisfaction. This homogeneity would explain why the form resistant correlation held across response rate variations.

4. Satisfaction scores were rendered homogenous through measurement error. One of the interesting findings from cognitive interviews is the propensity to utilize the "satisfied" score despite having expressed being dissatisfied for a particular questionnaire item. Participants, in recalling their experience, would moderate their dissatisfaction with contextual and/or character-based references. While it may be considered good news that alumni are not committing the familiar attribution error, the outcome is a leniency bias toward satisfaction.

In addition, it was observed that the generality of a given question was also related to the inscription of being "satisfied". General questions will bring forth multiple, contradicting recalled experiences that are difficult to reconcile under a single likert scale. The eventual summation of these recalled experiences seems to invariably point to being satisfied. If we add the propensity, as elaborated in the literature, to answer satisfied (e.g. "say yes to anything") when the question is not particularly salient, the net effect is a distorted distribution towards the satisfied score. This lack of variability in the distribution of satisfaction would explain, once again, why the form resistant correlation prevailed.

5. Construct validity issues. Finally, it must be acknowledged that concepts such as organizational citizenship behaviour, loyalty or complaint behaviour are difficult to define let alone operationalize. OCB, in particular, was operationalized as the degree of involvement with one's degree program. It was anticipated that those alumni, who were fully involved in all aspects of their program, including conference attendance and paper/article submission would, relatively speaking, be more satisfied than other alumni who were not involved. Involvement was underwritten by the understanding that an intense participatory relationship was forged with a supervisor, the home department and, to some extent, with other offices across the campus. Yet, logistic models reveal that involvement rarely figures as a predictor of being very satisfied. In fact, most models, but especially those focusing on very satisfied clusters, explained very little of the original "variance". Ironically, it would seem that the GSES was ill-equipped to harness what constitutes a "very satisfied" and a "very dissatisfied" graduate experience. This inability to properly harness the original variance

meant that few predictors are strongly related to the variables of interest, and thus introduce only a marginal level of nonresponse error.

8.4 Implications: Methodological and Organizational

The impetus for researching nonignorability originally grew out of our experience with the Graduate Student Exit Survey. For three consecutive years, the response rate never hovered beyond 40%. Every year, once all the responses had finally trickled in, the same anxieties resurfaced. As a novice survey practionner, we faced the responsibility of producing summary reports knowing that respondents may not have been representative of the alumni population. This issue of representativeness unwittingly fused together methodological concerns centered on proper survey design with organizational concerns focused on the weight these data can legitimately have in relation to public administration. Indeed, the same two questions haunted the production of a report: should greater resources be put into place to increase response rates, and what impact does nonresponse have on our understanding of alumni satisfaction with their experience at Waterloo? These two questions became the underlying focus of this dissertation.

Unfortunately, the results of our study have little external validity. They cannot be used to infer about the survey response propensity of alumni at large, nor can they speak to satisfaction surveys in general. As was discussed in the introduction and the methodology chapters, the dataset focused on three consecutive convocation years, and was drawn from a mid-size regional university whose reputation is likely to create satisfaction ceiling effects. However, despite the numerous methodological limitations, our findings can illuminate these

questions without answering them fully. They can provide the focus needed to address the more pertinent issues surrounding survey design and data reporting. This section will attempt to draw the implications of our findings to these practical matters by addressing methodological concerns first, then tackling organizational concerns second.

8.4.1 Should response rates be increased?

From the point of view of the survey practitioner, response rates hovering around thirty-five percent, such as the GSES, will undoubtedly spark some anxiety as to the representativeness of the survey results. One may ponder whether additional efforts should be put into place in order to decrease nonresponse error. The key questions are whether such efforts are needed, and at which group they should be targeted. Our results point towards ignorability. There is little evidence that increases in response rates compromised the form resistant hypothesis on a battery of subjective, objective and socio-demographic characteristics; there is sufficient evidence, however, that a spurious relationship exists between response rates and satisfaction clusters. Nonresponse is missing at random, and remains so despite the controlling of predictors of topic saliency at the alumni and departmental levels. As a result, the question of increasing response rates is contingent upon the types of statistics one wishes to use. If the GSES dataset is intended for modelling purposes only, response rates need not be increased. However, if the dataset is intended to be used for the reporting of absolute percentages, response rates must be increased. It should be said that, at 35% response rates, decreasing the size of the sampling error might become the overriding concern.

More importantly, the question as to whether response rates should be increased or not cannot be a substitute for a population-appropriate survey design. Indeed, the GSES may have been lucky. It is reasonable to deduce that the finding of ignorability may have been based upon a series of methodological alignments that worked together to homogenize the respondent pool. Although there are no specific pieces of evidence to support this contention, three cautions can be derived from our results. First, the inability to properly discern non-contact from refusals may have unwittingly left us in the dark about the potential for bias from the most satisfied of alumni. The inability to detect a link between full involvement and being very satisfied may have its source within involvement itself. Students that are highly involved during their degree program are likely to also be involved after graduation[71]. As a result, there is a possibility that contactibility is related to the probability of being very satisfied. This potential bias is compounded by a survey protocol that unduly limits the type and frequency of reminders – as was the case for most convocations (e.g. control group protocol). The consequences of poor reminder strategies are well documented in Dillman (2000). The net result is a potential truncation of the "very satisfied" from the response pool.

Second, as suggested above, dissatisfaction may have been more costly to report than originally anticipated. Dissatisfied alumni may require, in return for their helpfulness, a form of reciprocation. Survey designs that specifically avoid the use of incentives may unwittingly discourage dissatisfied alumni from answering. Again, the net result is a potential truncation of dissatisfaction from the response pool. Third, nonresponse bias may have been eradicated

[71] Schiltz (1988) makes a similar point with the most dissatisfied of alumni. It is likely, according to the author, that dissatisfied individuals simply disappear and do not forward to the university their current address. Once again, the issue of contactibility may have a relation with the probability of being dissatisfied.

by purging specific bias-prone groups from the population frame. This was undoubtedly the unwitting outcome following the decision to eliminate students who failed to register or withdrew from their respective programs. These individuals, although not technically alumni, would certainly hold different satisfaction levels than graduating alumni; at the very least, certain questions may be more pertinent and salient to their experience. Removing these students from the population frame may have predisposed the sample towards ignorability.

8.4.2 Does nonresponse make any difference to the understanding of alumni satisfaction?

Public administrators who commission alumni surveys are undoubtedly also placed in a bind if survey estimates are drawn from only 35% of the alumni population for a given year. While it is understood that satisfaction exit surveys are only one amongst many other alumni statistics to be used as performance indicators, they nonetheless remain, rightly or wrongly, important indicators to government agencies, and other consumers of educational data[72]. Our results point to three general recommendations:

First, the understanding of alumni satisfaction must rest on a better understanding of the difference between precision and accuracy. Our findings suggest that nonresponse error on satisfaction clusters remains small. At lower response rates, with small populations such as is the case for the GSES or for graduate student surveys generally, the confidence interval of the estimate (precision) is likely to exceed the bias caused by nonresponse error (accuracy). Under these specific conditions, and strictly for the survey results at hand (i.e. not collated with other convention), one may simply fall back on the largest of the two errors,

[72] See for example (Cheng & Tam, 1997; D. Smith, 2000)

namely the precision. However, for larger samples such as the polling of an entire undergraduate alumni population ranging in the thousands, a response rate of 35% will yield much smaller confidence intervals; these intervals are likely to become equivalent if not smaller than the bias caused by nonresponse error. Unless the data is accompanied by nonresponse bias analyses that confirm ignorability, under large polling conditions, public administrators must remain extremely sceptical of the representativeness of survey results.

Second, the understanding of alumni satisfaction, traditionally conveyed through descriptive statistics, must give way to a better understanding of the impact of nonresponse on the reporting of absolute percentages. Indeed, the reporting of percentages on individual variables must be read with scepticism even with ignorability confirmed. As we have demonstrated in chapter two, when predictors of satisfaction are also related to survey cooperation behaviour, absolute percentages (marginals) are likely to vary under varying response rate conditions. The magnitude of the nonresponse error will depend on how strongly predictors are correlated to satisfaction scores. Our findings suggest that common enrolment variables such as women alumni, visa-registered, or time-to-completion are weakly related to satisfaction and produced low levels of nonresponse error. Nevertheless, contextual analyses did report changes in the probabilities of being very satisfied and very dissatisfied, particularly within departmental response rates. In addition, there are still doubts about the more difficult to obtain variables such as departmental involvement; it has been shown to have an impact on satisfaction, and was observed in the literature as having a relation to survey cooperation behaviour. All of this would cause bias in absolute percentages. The recommendation proposed by (Stouffer, Suchman, DeVinney, Star, &

Williams, 1949: 105-229) is to report the variables of interest as comparisons between groups, as relative percentage points, or more abstractly as odds ratios. If nonresponse is ignorable, relative percentages will not change for a variation in response rate. Our results seem to uphold the form resistant correlation hypothesis.

Third, understanding of alumni satisfaction must be based on questions that are salient to survey participants. Perhaps the most important finding of this dissertation centers on the survey questions themselves. The ability to detect the impact of nonresponse on the understanding of alumni experience is only as good as the measurement instrument will allow. In that regard, it should be said that the GSES questionnaire was plagued by overly general, and at times irrelevant and tedious survey questions. Upon closer scrutiny, cognitive interviews provided evidence of leniency bias which homogenized answers towards the satisfaction scale, and evidence of contextual bias which introduced randomness in the satisfaction scale. These two biases might partly explain why multivariate models did so poorly in accounting for the variance within each satisfaction cluster. They may account for the inability to detect potentially real nonignorable response patterns. The end result is a distorted, highly skewed distribution towards the satisfied end of the scale – a distribution that masks and reassures. Therefore, the findings of this thesis should prompt public administrators to carefully craft survey questions that define the topic clearly, that are relevant to the student experience, that are succinct in number, and if at all possible, have their meaning affixed to theoretically-driven concepts (as opposed to bureaucratic processes).

Fourth, it should be noted that public administrators who commission organizational surveys, and survey practitioners who implement them, should not altogether dismiss

nonresponse in enrolment variables (such as gender, visible minority, visa-registered, etc), despite the fact that nonresponse may be deemed empirically ignorable. We would contend there remains a responsibility to ascertain, as a separate study perhaps, why these groups are not responding, and which aspects of their experience may be associated with their decision not to cooperate with the survey request. This proactive methodological approach has the advantage of highlighting and preventing potential sources of nonignorability, if and when any new experiential dimensions are included in any future request, a real possibility for anyone genuinely concerned with making satisfaction surveys topically salient. This proactive approach also attends to what Berinsky (2004) calls "exclusion bias" with respect to the concept of representation; self-excluded voices still threaten the validity of satisfaction surveys as representative of alumni experiences. If, as we have suggested in the introduction, the latent functions of satisfaction surveys are to open lines of communication with alumni, and inform management of alumni experiences, satisfaction surveys should strive for equal opportunity in representing the individual's authentic graduate experience. This would imply going beyond questions based on organizational procedures, and towards questions salient to students themselves.

8.5 Recommendations for Future Research

Having summarized the research findings and assessed their implications for survey practitioners and public administrators, we can further conclude that the limitations encountered throughout our study on ignorability suggest three future lines of inquiry:

First, ignorability might be endemic to satisfaction surveys. It was hoped that satisfaction surveys would be, in the minds of some alumni, sufficiently salient to generate distinguishable satisfaction scores between respondents and nonrespondents. It was hypothesized that topic intensity was at the root of nonignorable response behaviour. However, cognitive interviews revealed that satisfaction is not analogous to attitude crystallization, but rather, is the product of a long chain of cognitive processes. Satisfaction, as a cognitive outcome, is a complex interaction between the understandings of the survey request, the recalling of vague experiences, the often-conflicting evaluations of multiple disjointed events, the mapping of complex judgements into simple scales, and the editing of answers according to whom the question is directed to. This cognitive distance between the survey request and the eventual emergence of satisfaction suggest that topic saliency is largely a muddled construct, not susceptible to be generative of intense emotional involvement. And yet, there is ample evidence amongst the customer satisfaction literature to recognize that dissatisfaction often remains untapped by satisfaction surveys and is channelled into word-of-mouth behaviours. We need to better understand what are the cognitive processes that lead to topic intensity, and thus create the motivational force to answer a survey request.

Second, ignorability might be endemic to organizational experience. Universities, like many other large organizations, are heavily bureaucratized and subject to several regulatory layers. This organizational culture may have played a large role in homogenizing the experiences of graduate students during their degree program. Given this streamlining process, it is perhaps not surprising that responding alumni did not seem to be radically

different from nonresponding alumni. Even in terms of degree program involvement, alumni who embodied prescribed research and teaching assistantship roles were not particularly more inclined to be very satisfied with their degree program than others who took an instrumental stance to their degree program. The survey says precious little about what constitutes a meaningful, engaging experience; and what constitutes an alienating organizational experience. We are not in a position to explain why, for example, those who exceeded time-to-completion deadlines were most apt to be very satisfied, when compared to the NGS dataset. It is conceivable that the highs and lows of organizational life reside at the margins of the mainstream organizational experience. To better understand the consequences of nonresponse to student satisfaction surveys, we need to understand how organizational involvement relates to topic involvement.

Third, ignorability might be endemic to surveys constructed around organizational self-interests. While geared to understanding customer/client/citizen experiences, it is not uncommon for organizations to understand these experiences through the grids of their own organizational procedures and outcomes, and not through the individual's own understanding of what their experience should have been. As Miller and Bender (2005) explain, the college experience is steep with the socialization of standards of behaviour, and performance. And yet, little energy is put forth to understand what kind of expectations graduate student may have regarding their degree program. This oversight may have been at the root of the poor logistic model fit, and may also have been the one of the reasons lurking behind the nonignorable impact of visible minority on being very satisfied and being very dissatisfied. The costs are that we have no idea why alumni, beyond the trite bureaucratic milestones,

have reported satisfaction levels the way they did. We have little idea why visible minority alumni nonresponse was nonignorable – with all the implications of social inequalities this brings forth. Finally, because the survey showed such a poor understanding of the university experience, we are not in a position to understand which aspects of the graduate experience are likely to be generative of ignorability, and which are generative of nonignorable nonresponse. Thus, to understand the consequences of nonresponse to student satisfaction surveys presupposes a good understanding of student expectations themselves.

Appendix A

Graduate Student Exit Survey Questionnaire

The first version of the Graduate Student Exit Survey was mailed two weeks after the spring and fall convocation of the year 2000. The initial questionnaire was derived from Queen's University's survey of graduates**. U**pon inspection of the data quality from the responses received, it became evident that students were having difficulty managing the twelve page questionnaire. They were unable to ascertain which sections were relevant to them; they were unable to properly navigate through the various skip patterns; and they were dropping too many questionnaire items. Furthermore, the Queen's survey missed some elements important to the Graduate Studies Office which needed to be included in the questionnaire. Since the questionnaire was originally intended to be consistent with those of other institutions, very few structural changes were permitted. Nevertheless, a serious re-organization of the questionnaire items was completed prior to the convocation of 2001. A further modification to section 6, question 9 of the questionnaire was added prior to the convocation of 2002. The final version of the questionnaire is presented in this appendix. The data set was reformatted to reflect the new version of the questionnaire. As a result not all questions are available to all convocation years. A complete list of changes can be reviewed in the GSES User's Guide. Three main structural differences are noted below.

Table A-1 Questionnaire Modifications

Sections	Modifications
Supervisor Agreement Scale	Section 2, question 2 • Replaced agree/disagree scale with conventional four point Likert agreement scale. • Replaced "Not applicable" with "No opinion". The category "not applicable" is strictly reserved for populations for which the question does not apply.
University Services	Section 4 University services, question 1a ▪ Added "Finance Office" ▪ Added "Athletic Facilities"
Progress through Program: Would You Return to Waterloo	Section 6, Question 9 ▪ Added a comment box following "No, would not select Waterloo again"

See Graduate Student Exit Survey User's Guide 2002

In addition to the final version of the GSES questionnaire and all other materials related to the implementation of the survey were also included in this appendix: a copy of the anonymous postage-paid postcard sent with the questionnaire envelope, a copy of the reminder post card sent two weeks after the initial mailing of the questionnaire to gently remind the alumni to send his or her survey (mailing was based on the original mailing list

minus those alumni who were removed upon receipt of the post card above), and a copy of the cover letter sent with the questionnaire envelope. The name of the alumni was automatically printed on the cover letter after the word “Mr.” and “Dear Mr.”. The signature was originally done manually for the 2000 convocation but reverted to a signature image for the subsequent convocations.

UNIVERSITY OF WATERLOO GRADUATE STUDENT SURVEY
GETTING STARTED

University of Waterloo

We would like to ask you a series of questions about various aspects of your program. These questions are designed to be consistent with those in exit surveys at other institutions so that more general information on the quality of graduate programs in Ontario can be compiled.

SECTION 1
EXPERIENCE WITH THE GRADUATE STUDIES OFFICE (GSO)

❶ **In summarizing your experience at Waterloo, please indicate the extent to which you are satisfied or dissatisfied with each of the following statements as they refer to your experience with the Graduate Studies Office (GSO).**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
a. Dissertation/thesis submission	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
b. Accessibility and helpfulness of advice/information in general	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
c. Advice on scholarships, awards and other funding issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
d. Admissions procedures in general	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
e. Responsiveness and helpfulness of the admissions and records staff	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
f. Registration procedures	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
g. Concern for equity issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
h. Other*	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0

↳ *Please specify ☐

❷ **Do you have other comments concerning the Graduate Studies Office (GSO) (e.g. admission, registration, tuition fees/deadlines, scholarships, etc)?**

SECTION 2
EXPERIENCE WITH RESEARCH SUPERVISOR

❶ **After how many terms in your program were you assigned a thesis supervisor?**

- ☐ Assigned at the beginning
- ☐ Assigned during the program → ☐ terms
- ☐ Not applicable

Were you assigned a thesis co-supervisor?

- ☐ Assigned at the beginning
- ☐ Assigned during the program → ☐ terms
- ☐ Not applicable

❷ **Using the following scale, please indicate the extent to which you agree or disagree with each of the following statements as they refer to your research supervisor?**

	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. Was willing to spend the time necessary to advise me on academic/creative matters.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
b. Encouraged regular contacts to discuss my progress and/or academic work.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
c. Encouraged submission of work in progress for publication, presentation or exhibition.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
d. Was knowledgeable about the field.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
e. Could be relied upon to give me constructive criticism on my work.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
f. Returned my work in a timely manner.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
g. Was knowledgeable about Graduate Studies regulations.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
h. My overall relationship with my supervisor was good.	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0

❸ **Were you satisfied with discussions and decision-making with your supervisor with respect to authorship issues on publications?**

- ☐ Yes
- ☐ No → Please comment:

SECTION 3
EXPERIENCE WITH THE DEPARTMENT

❶ **In summarizing your experience at Waterloo, please indicate your opinion of each of the following using the scale shown below.**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
a. Appropriateness of course offerings	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
b. Program advising/guidance	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
c. Advice during orientation	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
d. Quality of teaching	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
e. Quality of research supervision	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
f. Appropriateness of degree requirements	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
g. Preparation for comprehensive exams	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
h. Relationship with other members of your supervisory committee	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
i. General "climate" for graduate study	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
j. Laboratory/studies facilities & equipment	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
k. Field work facilities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
l. Computing facilities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
m. Office space	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
n. Efforts to assist you finding employment upon graduation	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
o. Contacts with Graduate Officer	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
p. Contacts with Chair	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
q. Contacts with office staff	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
r. Oral defense process	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
s. Methods of evaluation of student performance	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
t. Concern for equity issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
u. Concern for intellectual property issues	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0

❷ **Do you have other comments concerning your department (e.g. participation in departmental governance, adequacy of space, keeping students informed administrative staff, etc)?**

SECTION 4
EXPERIENCE WITH THE UNIVERSITY SERVICES

❶ **In summarizing your experience at Waterloo, please indicate your opinion of each of the following using the scale shown below.**

	Very Satisfied	Satisfied	Dissatisfied	Very Dissatisfied	No Opinion
a. Finance Office	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
b. Student Housing	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
c. Day care	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
d. Safety on Campus	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
e. University Library	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
f. Athletic Facilities	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
g. Other services*	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0

↳ *Please specify _____

SECTION 5
EXPERIENCE WITH THE GRADUATE STUDENT ASSOCIATION (GSA)

❶ **On average, how often did you visit the Graduate House during your degree program?**

- ☐ Never
- ☐ Less than once per term
- ☐ A few times per term
- ☐ At least once per month
- ☐ Can't say/can't remember

❷ **On average, how frequently did you attend a GSA social event during your degree program?**

- ☐ Never
- ☐ Less than once per term
- ☐ A few times per term
- ☐ At least once per month
- ☐ Can't say/can't remember

❸ **How satisfied or dissatisfied were you with the manner in which the GSA addressed Graduate Student concerns?**

- ☐ Very satisfied
- ☐ Satisfied
- ☐ Dissatisfied
- ☐ Very dissatisfied
- ☐ No opinion

SECTION 5 (CONTINUED)
EXPERIENCE WITH THE GRADUATE STUDENT ASSOCIATION (GSA)

❹ **In your opinion, are there any services that the GSA does not offer but should?**

SECTION 6
PROGRESS THROUGH THE PROGRAM

❶ **Degree program in which you were enrolled most recently (MA, MAcc, MAES, MASc, MEng, MES, MFA, MMath, MSc, MTax, PhD)?**

Degree Program

❷ **Department/School:**
(Note: if you changed department/schools during your program, record the most recent department you were in.)

Department/School

❸ **When did you first enroll in this program?**

Month Year

❹ **If you withdrew from the program, please indicate in which term.**

Month Year

Please indicate the reason why you withdrew from your program:

❺ **Please indicate the number of terms in which you registered in each of the following categories:**

Full-time Part-time Inactive

❻ **Date of completion of your program: (date when you submitted your thesis to the Graduate Studies Office (hardcopy or electronically), or the date on which all formal requirements were complete.)**

Month Year

☐ Did not complete program

❼ **The length of time spent in your graduate program was**

☐ Shorter than I expected
☐ As I expected
☐ Longer than I expected

SECTION 6 (CONTINUED)
PROGRESS THROUGH THE PROGRAM

❽ **Why did you decide to study at Waterloo?**

❾ **If you were to start your program again, would you select Waterloo?**

☐ Yes

☐ No → Please comment:

Would you select the same program?

☐ Yes

☐ No

SECTION 7 (PhD Only)
EXPERIENCE WITH COMPREHENSIVE EXAMINATION

IF COMPLETED DEGREE WAS A MASTER'S DEGREE SKIP TO SECTION 8

❶ **After how many terms in your program did you complete the requirements for your comprehensive examinations?**

☐ Terms

❷ **Were the objectives of the comprehensive examinations clear to you?**

☐ Yes

☐ No

❸ **Using the following scale, please indicate the extent to which you agree or disagree with each of the following statements as they refer to your experience with the comprehensive examinations.**

	Strongly Agree	Agree	Disagree	Strongly Disagree	No Opinion
a. It was an important part of my overall learning for the degree	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
b. It was important in preparing me to undertake doctoral level research	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
c. Its primary role is to serve as a gatekeeper to the doctoral degree	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0
d. The time taken to prepare for the exam was too great for the value it produces	☐ 1	☐ 2	☐ 3	☐ 4	☐ 0

SECTION 8
RESEARCH EXPERIENCE

❶ **Did your degree involve a research thesis?**

☐ Yes
☐ No → **SKIP TO SECTION 9**

❷ **After how many terms in your program did you begin work on your thesis proposal?**

☐ Terms

❸ **After how many terms in your program was your thesis proposal approved?**

☐ Terms

❹ **Did you attend any Graduate Student Research Conferences at Waterloo?**

☐ Yes → if yes, how often? ☐
☐ No

❺ **Were you a presenter at such meeting?**

☐ Yes → if yes, how often? ☐
☐ No

❻ **Did you ever present your research at a seminar, colloquium, conferences (other than Graduate Student Research Conference) or group exhibition at Waterloo?**

☐ Yes
☐ No

❼ **Did you attend academic conferences outside Waterloo?**

☐ Yes → if yes, how often? ☐
☐ No

❽ **Did you deliver any papers/exhibit at such meetings?**

☐ Yes → if yes, how often? ☐
☐ No

❾ **Were your expenses funded?**

☐ Yes →
☐ Partially →
☐ No

If yes/partially, by whom?

❿ **Have you had a paper(s) from your research in this program published in a refereed journal?**

☐ Yes → if yes, how many? ☐
☐ No

SECTION 9
TEACHING ASSISTANTSHIP EXPERIENCE

❶ **Did you hold a teaching assistantship at any time during your program?**

- ☐ Yes
- ☐ No → **SKIP TO SECTION 10**

❷ **Did you participate in any organized programs to prepare for teaching assistantships or future university teaching offered by the following and, if yes, did you find them helpful?**

a. Department or faculty seminar
- ☐ Yes ——→ Was it helpful? ☐ Yes ☐ No
- ☐ No

b. Teaching Resources and Continuing Education (TRACE)
- ☐ Yes ——→ Was it helpful? ☐ Yes ☐ No
- ☐ No

c. Certificate in University Teaching
- ☐ Yes ——→ Was it helpful? ☐ Yes ☐ No
- ☐ No

❸ **Did you complete the certificate in University Teaching?**

- ☐ Yes
- ☐ No

❹ **Did you feel that resources were available to adequately prepare you for your role as a teaching assistant?**

- ☐ Yes
- ☐ No
- ☐ No opinion

❺ **How satisfied or dissatisfied were you with the course instructors' supervision of your teaching activities?**

- ☐ Very satisfied
- ☐ Satisfied
- ☐ Dissatisfied
- ☐ Very dissatisfied
- ☐ No opinion

❻ **Did you regard your time as a teaching assistant as a positive experience?**

- ☐ Yes
- ☐ No
- ☐ No opinion

SECTION 9 (CONTINUED)
TEACHING ASSISTANTSHIP EXPERIENCE

❼ **In what ways do you think that the TA experience could be improved? List in order of importance.**

SECTION 10
FINANCIAL SUPPORT

❶ **How did you support yourself during your most recent graduate program? Please give us your best estimate of the amount of support you received in each category in each year of your program.**

	Income Year 1	Income Year 2	Income Year 3	Income Year 4	Income Year 5
a. Teaching Assistantship	$	$	$	$	$
b. Research Assistantship	$	$	$	$	$
c. Ontario Graduate Scholarship	$	$	$	$	$
d. OGSST	$	$	$	$	$
e. UW Scholarship	$	$	$	$	$
f. Graduate Student Bursary	$	$	$	$	$
g. International Graduate Student Bursary	$	$	$	$	$
h. NSERC/SSHRC/MRC	$	$	$	$	$
i. Sponsored (CIDA, CBIE, etc)	$	$	$	$	$
j. Other services*	$	$	$	$	$

↳ *Please specify

❷ **We would like to gain more information about the financial pressures associated with graduate study and request data on changes in your indebtedness during the period in your most recent graduate program. Please estimate your total indebtedness (OSAP, Canada Student Loans, loans, etc).**

$ At the beginning of the graduate program you have just completed.

$ At the end of the graduate program you have just completed.

SECTION 10 (CONTINUED)
FINANCIAL SUPPORT

❸ **Do you believe that Waterloo's financial support was distributed in a fair and equitable manner?**

- ☐ Yes
- ☐ No
- ☐ No opinion

❹ **At any time during your program were you registered as an international student (i.e. with a student visa)?**

- ☐ Yes
- ☐ No → **SKIP TO SECTION 11**

❺ **Did you receive a tuition of fee waiver as an international student?**

- ☐ Yes → If yes, for how many terms? ☐
- ☐ No

SECTION 11
EXPERIENCE AFTER LEAVING WATERLOO

❶ **On leaving Waterloo, <u>which one</u> of the following are you pursuing or do you intend to pursue?**

- ☐ Another degree → Which degree are you pursuing? ☐ PhD ☐ MA ☐ Other
- ☐ Post-Doctoral fellowship/traineeship
- ☐ Employment (including self-employment)
- ☐ Other, please specify: ☐
- ☐ Uncertain

❷ **Which best describes your first professional employment after leaving Waterloo?**

- ☐ Research associateship
- ☐ Faculty position → Tenure track? ☐ Yes ☐ No
- ☐ Research → For whom? ☐ Industry ☐ Government ☐ Other
- ☐ Professional → Please specify: ☐
- ☐ Other non-academic position → Please specify: ☐
- ☐ Seeking employment → **SKIP TO SECTION 12**

❸ **Where is your position located?**

- ☐ Ontario
- ☐ Another Canadian Province or territory: ☐
- ☐ Outside of Canada → Which country? ☐

ADDITIONAL COMMENTS

University of Waterloo

Dear Dr. Sivak:
I am mailing my completed survey to you under separate cover. You can therefore remove my name from the list maintained for follow-up contacts.

I am:

Name: .....

Address: .....

.....

.....

University of Waterloo

About two weeks ago a survey seeking your opinions about your experience as a graduate student at the University of Waterloo was mailed to you. If you have already completed and returned the survey, please accept our sincere thanks. If not, please consider doing so today.

This survey is routinely sent to all graduating graduate students at Waterloo. Your responses are very important to us. We are especially grateful for your help because it is only by asking former graduate students, such as yourself, to share your experiences that we can understand how to make improvements to the overall graduate studies experience at UW.

Sincerely,

Dr. Jacob G. Sivak

Dean of Graduate Studies
University of Waterloo
Telephone: 519-888-4567, ext. 3439

University of Waterloo

Dean of Graduate Studies

University of Waterloo
200 University Avenue West
Waterloo, Ontario, Canada
N2L 3G1

519-888-4567
Fax 519-746-3051
http://www.grad.uwaterloo.ca

June 17, 2002

Mr. [insert name here]
Graduate Student
Statistics & Actuarial Science
MATHEMATICS

Dear Mr. [insert name here]

We would like to provide this opportunity for you to reflect on the time you have spent as a graduate student at Waterloo. In particular, we want to learn more about students' experiences with the University's faculty members, staff and services. As well, we seek more insight into students' progress through the various stages of their graduate program.

The information that you provide to us is especially important in contributing to our on-going effort to find new and better ways of doing things at the department and Graduate Studies Office levels. Your voice, in concert with those of all other former graduate students, is valuable in helping us to improve the overall experience of current and future graduate students.

Your answers are completely anonymous. There is no name, or identifying code, attached to each questionnaire. In addition, the information the survey provides will be released in an aggregate form in which no individual answers or persons can be identified. Individual surveys will not be made available to any department, and any information associated with this survey cannot be included in your file. This survey has been reviewed by and received ethics clearance through the Office of Research Ethics. If you have any questions, comments or concerns about this survey please feel free to contact the Office of Research Ethics at 519-888-4567, ext. 6005, or you can write to us at the address on the letterhead.

Your participation in this survey is voluntary and you may decline to answer any question that you prefer not to answer. However, you can help us immensely by taking a few minutes of your time to share your graduate student experiences at Waterloo.

Thank you for taking the time to respond to this important survey.

Sincerely,

Jacob G. Sivak
Dean of Graduate Studies
Professor of Optometry

Telephone Reminder Messages

A second and final reminder message was delivered on the answering machine, to the parent or to the alumni directly. No multiple calling strategies were used. One phone call per nonresponding alumni.

1. Message left to an answering machine

This is a message for [*Student's name*]
My name is _____. I am a graduate student from the University of Waterloo. About four weeks ago, a Graduate Student survey was sent to all recent graduates. To the best of our knowledge, we have not yet received your completed questionnaire. Your response to this survey is voluntary. We would appreciate, however, your willingness to consider completing the questionnaire and send it to Waterloo. Hearing from everyone is crucial in helping us improve the overall experience of current and future graduate students. If you have any question or concern, you may contact Penny Pudifin at (519) 888-4567 ext 2845.

2. Message left to an occupant

Hi. May I speak to [*Student's name*]

[*if not available*]
May I leave a message?

My name is _____. I am a graduate student from Waterloo. A few weeks ago, the University mailed a survey to all its recent graduates. To the best of our knowledge, we have not yet received a questionnaire from [*student's name*]. If she/he hasn't had a chance to respond, we look forward to receive the completed questionnaire as soon as possible. If there is any question or concern, she/he may contact Penny Pudifin at (519) 888-4567 ext 2845.

3. Message left to the student

Hi. My name is _____. I am a graduate student at the University of Waterloo. About four weeks ago, a Graduate Student Survey was sent to all recent graduates. To the best of our knowledge, we have not yet received your completed questionnaire. Your response to this survey is voluntary. We would appreciate, however, if you could find the time to complete the questionnaire and send it to Waterloo. If you have any question or concern, please feel free to contact Penny Pudifin at (519) 888-4567 ext 2845

Appendix B

Cognitive Interview Protocol

The following questions were administered to all participants. The cognitive interview was broken into three distinct sections. First, participants were handed in a closed envelope containing cover letter, return envelope, questionnaire and mail-back postcard. They were instructed to behave normally as if they would receive this through the mail. Observational data was gathered.

Cognitive Interview Introduction

Thanks again for participating in this study. As I explained in my letter, this study will be testing a graduate student exit survey that was administered for three years here at Waterloo. Similar surveys are being administered at other Universities in Canada and the United States.

The interview has three parts and I'll give you a sense of where we are at as we move along the interview. At points, I'll be asking questions about your experience (as a recent alumnus) at Waterloo, and/or I'll ask you to think aloud while reading some sections of the survey material.

Because self-administered surveys are done at the student's home and that it contains only written material, survey researcher are not able to see what happens when the student receives the survey material, and are not able to know what goes on in there heads while interacting with the material. The basic goal of the interview is that have someone present with the recent alumni, and have him or her think aloud while you read the survey material.

Thinking aloud means saying out loud everything that pops into your head as you are trying to assess the material presented to you. You tell me what you are currently trying to do, what you are trying to understand, what is it that you are recalling as you are reading certain words.

All your answers to my questions, questionnaire, and field notes will be kept entirely confidential. Your name will never be linked to any part of my research thesis, or any subsequent publications. Your name will not be revealed.

It will be too difficult for me to write down all that you say, and want to make sure I get your exact remarks recorded. So is it all right if I tape record our discussion?

Could you please sign the various consent forms?

Please say what you really think. They are no wrong or right answer. It's not a test. You are free to answer or decline to answer any questions you wish. You're helping us learn how people will use the booklet and how we can improve upon it.

Any questions before we get started?

Phase I: Processing the Cover Letter

Please read the cover letter, highlight any words that are unclear, and let me know when you are ready to talk about it?

In your own words, could tell me what the cover letter is about?

Did you notice the section on "we want to learn more about students' experiences with the University's faculty members, staff, and services"?
No – Skip
Yes – Could you tell me in your own words what the cover letter meant by the experience?

Did you notice the section on "Your answers are completely anonymous"?
No – Skip
Yes – Did reading section make you more likely or less likely to trust the GSO?
Yes/No Why is that?

Did you notice who is sponsoring the survey?
No – Skip
Yes – Could you tell me who it is? How did you find out?

In general, how would you qualify your graduate experience at Waterloo?
Is there a particular event(s), person(s), or aspect(s) of your program that comes to mind when you think about your graduate experience at Waterloo?

Phase II: Sorting Questionnaire Sections

A) If you could glance at the questionnaire, you will see that they are twelve sections. Without going to deeply into reading the actual questions, could you order these sections according to what seemed the most relevant to your experience at Waterloo? From the most important (1) to the least important (12)

Section 1: Experience with Graduate Studies Office

Section 2: Experience with Supervisor

Section 3: Experience with Department

Section 4: Experience with University Services

Section 5: Experience with Graduate Student Association

Section 6: Progress through the Program

Section 7: Experience with Comprehensive Examination

Section 8: Research Experience

Section 9: Teaching Assistantship Experience

Section 10: Financial Support

Section 11: Experience after Leaving Waterloo

Section 12: Demographic Information

B) Without looking at any sections in detail, are there any that seem a little threatening to you?

Phase III: Processing the Questionnaire

A) Next, I would ask you to start answering the questionnaire. Please highlight any words or phrases that may seem unclear, confusing. Circle with the pen, questions that appear to touch upon a relevant aspect of your experience at Waterloo. Should there be any questions that makes you uneasy (perhaps to the point of re-evaluating your participation), please let me know. As you answering each questions, let me know what you are thinking about.

Debriefing

A) Looking back at the survey, was it what you initially expected when you read the cover letter?

B) Looking back at the survey was it generally difficult to answer (difficulty in remembering, too many unclear words, too long, too many open-ended questions)?

C) Looking back at the survey, was it intrusive? Do you feel comfortable in sending these answers to the GSO?

D) Looking back at the survey, did it seem relevant to your graduate student experiences?

Appendix C

Multiple Correspondence Analysis Cluster Loadings

Table C-1 GSES 2000-02 Global Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4	5	6	7
		Count	116	154	186	54	153	18	131
		Percent	14.3%	19.0%	22.9%	6.7%	18.8%	2.2%	16.1%
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1G	Preparation for comprehensive exams	SAT	14.76						
Q7.3A	Comprehensives part of overall learning	AGR	13.38						
Q7.3A	Comprehensives part of overall learning	S AGR	13.10						
Q7.3C	Comprehensives as gatekeeper	AGR	12.82						
Q7.3B	Comprehensives important for research	AGR	12.60						
Q7.3B	Comprehensives important for research	S AGR	11.89						
Q7.3D	Comprehensives as time consuming	DSG	11.75						
Q3.1G	Preparation for comprehensive exams	V SAT	11.07						
Q7.3C	Comprehensives as gatekeeper	DSG	10.11						
Q7.3D	Comprehensives as time consuming	AGR	10.06						
Q7.3B	Comprehensives important for research	DSG	8.97						
Q7.3D	Comprehensives as time consuming	S DSG	8.54						
Q7.3C	Comprehensives as gatekeeper	S AGR	8.34						
Q3.1R	Oral defence process	V SAT	7.81	3.18					
Q7.3D	Comprehensives as time consuming	S AGR	6.86						
Q9.6	Ta as positive experience	YES	5.77	3.53			3.66		
Q3.1E	Quality of research supervision	V SAT	5.02	11.30			3.27		
Q3.1H	Relationship with supervisory committee	V SAT	4.62	5.79					
Q3.1R	Oral defence process	SAT	3.95		2.80				
Q9.5	Satisfaction with instructors supervision	SAT	3.88		2.77				
Q3.1P	Contacts with chair	V SAT	3.78	3.16					
Q3.1H	Relationship with supervisory committee	SAT	3.64		4.17				
Q5.3	Satisfaction with GSA response	SAT	3.63				2.81		

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1B	Program advising/guidance	V SAT	3.54	4.71			3.27		
Q2.2C	Encouraged submission of work	AGR	3.43				9.68		
Q1.1A	Thesis submission	V SAT	3.41	7.90					
Q2.2G	Knowledgeable about GSO regulations	AGR	3.26		2.79		8.68		
Q2.2C	Encouraged submission of work	S AGR	3.21	14.77					
Q3.1U	Concern for intellectual property issues	SAT	3.20						
Q3.1S	Evaluation of student performance	V SAT	3.13	6.36					
Q2.2F	Returned work in a timely manner	AGR	3.12		2.63		11.94		
Q3.1M	Office space	V SAT	3.06	4.33					
Q2.2E	Provided constructive criticism of work	S AGR	3.04	16.05					
Q3.1P	Contacts with chair	SAT	2.92						
Q2.2B	Encouraged regular contacts	AGR	2.91				12.09		
Q2.2A	Spent time to advise me on academic matters	S AGR	2.91	17.82					
Q3.1N	Assistance in finding employment	SAT	2.87				3.18		
Q2.2G	Knowledgeable about GSO regulations	S AGR	2.85	11.39					
Q1.1G	Concern for equity issues	SAT	2.81						
Q3.1J	Laboratory/studies facilities & equipment	V SAT	2.80	3.11			2.93		
Q2.2B	Encouraged regular contacts	S AGR	2.67	16.16					
Q2.2H	Overall relationship was good	AGR	2.66		5.19		15.23		
Q2.2D	Knowledgeable about the field	S AGR	2.61	14.89	2.47				
Q3.1L	Computing facilities	V SAT	2.56	3.93			3.98		
Q9.5	Satisfaction with instructors supervision	NOP	2.54						
Q3.1M	Office space	SAT	2.36		3.17				
Q2.2H	Overall relationship was good	S AGR		17.85					
Q2.2F	Returned work in a timely manner	S AGR		14.68					
Q4.1C	Day care	NOP		7.91	11.23				
Q3.1F	Appropriateness of degree requirements	V SAT		7.45					

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1I	General "climate" for graduate study	V SAT		7.10					
Q7.3A	Comprehensives part of overall learning	NR		7.01	7.96		5.78		6.81
Q7.3B	Comprehensives important for research	NR		7.01	7.96		5.78		6.81
Q7.3C	Comprehensives as gatekeeper	NR		6.98	7.92		6.12		6.78
Q7.3D	Comprehensives as time consuming	NR		6.94	7.89		6.08		6.75
Q3.1Q	Contacts with office staff	V SAT		6.55			3.74		
Q3.1K	Field work facilities	NOP		6.17	8.29				
Q1.1B	Advice accessibility and helpfulness	V SAT		5.94			3.03		
Q3.1R	Oral defence process	NOP		5.44	8.37				
Q4.1B	Student housing	NOP		5.26	7.60				
Q6.9B	I would choose the same program	YES		5.24					
Q4.1E	University library	V SAT		5.08			2.94		
Q1.1G	Concern for equity issues	V SAT		5.02					
Q3.1G	Preparation for comprehensive exams	NR		4.88	5.97		5.47		6.16
Q3.1U	Concern for intellectual property issues	V SAT		4.68			2.42		
Q1.1E	Admissions staff responsiveness	V SAT		4.47					
Q6.9AA	I would choose Waterloo again	YES		4.44					
Q3.1O	Contacts with graduate officer	V SAT		4.40					
Q3.1T	Concern for equity issues	NOP		4.38	4.69				3.79
Q1.1F	Registration procedures	V SAT		4.14			4.19		
Q5.3	Satisfaction with GSA response	NOP		3.93	7.04				3.10
Q1.1D	Admissions procedures	V SAT		3.84			3.46		
Q3.1N	Assistance in finding employment	NOP		3.63	5.21				4.13
Q1.1G	Concern for equity issues	NOP		3.54	4.99				
Q3.1T	Concern for equity issues	V SAT		3.49			2.54		
Q3.1M	Office space	V DIS		3.19		3.39			
Q1.1C	Scholarships advice	V SAT		3.14			3.90		

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1H	Relationship with supervisory committee	NOP		3.14	7.67				
Q3.1C	Advice during orientation	V SAT		2.97			2.75		
Q3.1U	Concern for intellectual property issues	NOP		2.85	3.34				4.25
Q9.5	Satisfaction with instructors supervision	V SAT		2.80			3.48		
Q3.1K	Field work facilities	V SAT		2.63			2.66		
Q3.1D	Quality of teaching	V SAT		2.57					3.88
Q4.1B	Student housing	V SAT		2.47			2.81		
Q4.1D	Safety on campus	V SAT		2.43			3.94		
Q3.1A	Appropriateness of course offerings	V SAT		2.36			3.10		
Q3.1E	Quality of research supervision	SAT			8.80		2.79		
Q1.1A	Thesis submission	SAT			6.52				
Q2.2A	Spent time to advise me on academic matters	AGR			6.50		13.35		
Q2.2G	Knowledgeable about GSO regulations	NOP			6.45				
Q2.2C	Encouraged submission of work	DSG			5.91	4.37			
Q4.1D	Safety on campus	NOP			5.71				
Q3.1I	General "climate" for graduate study	SAT			5.66				
Q1.1C	Scholarships advice	NOP			5.63				
Q3.1P	Contacts with chair	NOP			5.57				7.04
Q2.2C	Encouraged submission of work	NOP			5.43				
Q2.2B	Encouraged regular contacts	DSG			5.36	5.87			
Q3.1C	Advice during orientation	NOP			5.32				
Q3.1Q	Contacts with office staff	SAT			5.32				
Q3.1J	Laboratory/studies facilities & equipment	NOP			4.70				2.60
Q3.1D	Quality of teaching	SAT			4.52				
Q3.1F	Appropriateness of degree requirements	SAT			4.47				2.52
Q2.2E	Provided constructive criticism of work	AGR			4.42		13.62		
Q3.1L	Computing facilities	NOP			4.29				

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1E	Quality of research supervision	DIS			4.21	3.89			
Q2.2F	Returned work in a timely manner	S DSG			4.21				
Q1.1F	Registration procedures	NOP			3.87				
Q3.1O	Contacts with graduate officer	NOP			3.74				7.08
Q3.1B	Program advising/guidance	SAT			3.56				
Q2.2B	Encouraged regular contacts	S DSG			3.40	2.42			
Q2.2F	Returned work in a timely manner	DSG			3.22	6.18			
Q1.1D	Admissions procedures	SAT			3.03				
Q3.1S	Evaluation of student performance	SAT			2.99				
Q1.1B	Advice accessibility and helpfulness	SAT			2.78				
Q3.1B	Program advising/guidance	NOP			2.75				
Q1.1E	Admissions staff responsiveness	SAT			2.64				
Q3.1K	Field work facilities	SAT			2.53				
Q2.2E	Provided constructive criticism of work	DSG			2.38	7.83			
Q4.1E	University library	NOP			2.34				3.37
Q3.1B	Program advising/guidance	V DIS				10.17			
Q3.1S	Evaluation of student performance	V DIS				9.41			
Q3.1E	Quality of research supervision	V DIS				8.94			
Q2.2H	Overall relationship was good	DSG				8.59			
Q2.2A	Spent time to advise me on academic matters	DSG				7.94			
Q3.1I	General "climate" for graduate study	V DIS				7.63			
Q6.9AA	I would choose Waterloo again	NO				7.28			
Q2.2D	Knowledgeable about the field	DSG				6.95			
Q3.1A	Appropriateness of course offerings	V DIS				6.47			
Q3.1C	Advice during orientation	V DIS				6.40			
Q3.1D	Quality of teaching	DIS				6.08			
Q2.2G	Knowledgeable about GSO regulations	DSG				5.72			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1N	Assistance in finding employment	V DIS				5.54			
Q3.1J	Laboratory/studies facilities & equipment	V DIS				4.96			
Q3.1U	Concern for intellectual property issues	V DIS				4.96			
Q6.9B	I would choose the same program	NO				4.91			
Q3.1O	Contacts with graduate officer	V DIS				4.87			
Q3.1I	General "climate" for graduate study	DIS				4.74			
Q3.1U	Concern for intellectual property issues	DIS				4.64			
Q3.1B	Program advising/guidance	DIS				4.43			
Q3.1O	Contacts with graduate officer	DIS				4.35			
Q3.1L	Computing facilities	V DIS				4.16			
Q3.1P	Contacts with chair	V DIS				4.07			
Q3.1T	Concern for equity issues	DIS				4.03			
Q9.6	Ta as positive experience	NO				4.03			
Q3.1Q	Contacts with office staff	DIS				3.91			
Q1.1C	Scholarships advice	V DIS				3.90			
Q5.3	Satisfaction with GSA response	DIS				3.68			
Q1.1B	Advice accessibility and helpfulness	DIS				3.56			
Q3.1P	Contacts with chair	DIS				3.55	2.74		
Q3.1F	Appropriateness of degree requirements	DIS				3.47			
Q2.2E	Provided constructive criticism of work	S DSG				3.31			
Q3.1F	Appropriateness of degree requirements	V DIS				3.21			
Q6.9B	I would choose the same program	NR				3.18			
Q3.1H	Relationship with supervisory committee	DIS				3.13			
Q9.5	Satisfaction with instructors supervision	DIS				2.97			
Q1.1G	Concern for equity issues	NR				2.95	8.41		
Q1.1F	Registration procedures	V DIS				2.89			2.36
Q3.1C	Advice during orientation	DIS				2.82			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

SECT.	ITEM LABEL	SCALE	SATISFACTION CLUSTERS 1	2	3	4	5	6	7
Q4.1C	Day care	NR				2.70	16.18		
Q4.1B	Student housing	V DIS				2.63			
Q1.1E	Admissions staff responsiveness	DIS				2.62			
Q1.1A	Thesis submission	DIS				2.36			
Q1.1G	Concern for equity issues	DIS				2.34			
Q2.2D	Knowledgeable about the field	AGR					14.54		
Q4.1B	Student housing	NR					10.61		
Q5.3	Satisfaction with GSA response	NR					10.50		
Q3.1K	Field work facilities	NR					9.34	6.17	7.86
Q3.1N	Assistance in finding employment	NR					7.79	8.27	
Q3.1U	Concern for intellectual property issues	NR					6.89	7.82	
Q1.1C	Scholarships advice	NR					6.36		
Q3.1T	Concern for equity issues	NR					5.86	8.47	
Q3.1P	Contacts with chair	NR					5.73	7.82	2.96
Q3.1R	Oral defence process	NR					5.63	6.57	15.05
Q4.1D	Safety on campus	NR					4.19	2.74	
Q3.1T	Concern for equity issues	SAT					3.60		
Q3.1H	Relationship with supervisory committee	NR					3.48	6.93	16.31
Q3.1O	Contacts with graduate officer	NR					3.38	9.17	
Q3.1O	Contacts with graduate officer	SAT					2.80		
Q3.1C	Advice during orientation	NR					2.74	9.39	
Q3.1L	Computing facilities	NR					2.50	10.21	
Q3.1A	Appropriateness of course offerings	NR						12.46	
Q3.1B	Program advising/guidance	NR						11.98	
Q3.1I	General "climate" for graduate study	NR						11.07	
Q3.1F	Appropriateness of degree requirements	NR						10.78	
Q3.1Q	Contacts with office staff	NR						10.75	2.45

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

SECT.	ITEM LABEL	SCALE	SATISFACTION CLUSTERS 1	2	3	4	5	6	7
Q3.1D	Quality of teaching	NR						10.53	
Q3.1S	Evaluation of student performance	NR						9.79	
Q3.1M	Office space	NR						8.56	5.22
Q3.1J	Laboratory/studies facilities & equipment	NR						7.95	7.88
Q3.1E	Quality of research supervision	NR						7.32	19.35
Q2.2H	Overall relationship was good	NR						6.75	23.58
Q2.2D	Knowledgeable about the field	NR						6.31	24.11
Q2.2E	Provided constructive criticism of work	NR						6.31	24.11
Q2.2A	Spent time to advise me on academic matters	NR						6.28	23.95
Q2.2B	Encouraged regular contacts	NR						6.28	23.95
Q2.2F	Returned work in a timely manner	NR						6.28	23.95
Q2.2G	Knowledgeable about GSO regulations	NR						6.23	23.72
Q2.2C	Encouraged submission of work	NR						6.08	23.10
Q1.1D	Admissions procedures	NR						4.17	
Q1.1F	Registration procedures	NR						4.10	
Q1.1E	Admissions staff responsiveness	NR						3.78	
Q4.1E	University library	NR						3.67	
Q1.1B	Advice accessibility and helpfulness	NR						3.22	2.54
Q9.6	Ta as positive experience	NR						2.81	13.45
Q9.5	Satisfaction with instructors supervision	NR						2.69	13.51
Q1.1A	Thesis submission	NR							18.85
Q3.1M	Office space	NOP							12.48
Q3.1Q	Contacts with office staff	NOP							5.70
Q3.1E	Quality of research supervision	NOP							4.92
Q1.1B	Advice accessibility and helpfulness	NOP							4.55
Q3.1S	Evaluation of student performance	NOP							4.14
Q1.1A	Thesis submission	NOP							3.66

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-1 GSES 2000-02 Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1C	Advice during orientation	SAT							3.65
Q4.1B	Student housing	SAT							3.54
Q4.1E	University library	SAT							3.35
Q4.1D	Safety on campus	SAT							3.00
Q3.1A	Appropriateness of course offerings	SAT							2.66
Q1.1F	Registration procedures	DIS							2.41

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-2 GSES 2000-02 Supervisor Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	342	90	230	150			
		Percent	42.1%	11.1%	28.3%	18.5%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q2.2A	Spent time to advise me on academic matters	AGR	23.19						
Q2.2E	Provided constructive criticism of work	AGR	22.02						
Q2.2H	Overall relationship was good	AGR	21.37						
Q2.2D	Knowledgeable about the field	AGR	19.11						
Q2.2B	Encouraged regular contacts	AGR	18.94						
Q2.2F	Returned work in a timely manner	AGR	18.31						
Q2.2C	Encouraged submission of work	AGR	14.38						
Q2.2G	Knowledgeable about GSO regulations	AGR	12.35						
Q2.2G	Knowledgeable about GSO regulations	DSG	3.23	4.84					
Q2.2G	Knowledgeable about GSO regulations	NOP	2.35						
Q2.2E	Provided constructive criticism of work	DSG		14.08					
Q2.2A	Spent time to advise me on academic matters	DSG		12.50					
Q2.2H	Overall relationship was good	DSG		11.75					
Q2.2D	Knowledgeable about the field	DSG		10.76					
Q2.2B	Encouraged regular contacts	DSG		10.57					
Q2.2C	Encouraged submission of work	DSG		8.93					
Q2.2F	Returned work in a timely manner	DSG		7.94					
Q2.2B	Encouraged regular contacts	S DSG		7.63					
Q2.2F	Returned work in a timely manner	S DSG		7.52					
Q2.2E	Provided constructive criticism of work	S DSG		6.73					
Q2.2A	Spent time to advise me on academic matters	S AGR			24.19				
Q2.2E	Provided constructive criticism of work	S AGR			22.89				
Q2.2H	Overall relationship was good	S AGR			22.28				

S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-2 GSES 2000-02 Supervisor Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q2.2B	Encouraged regular contacts	S AGR			21.60				
Q2.2F	Returned work in a timely manner	S AGR			20.95				
Q2.2D	Knowledgeable about the field	S AGR			19.90				
Q2.2C	Encouraged submission of work	S AGR			19.33				
Q2.2G	Knowledgeable about GSO regulations	S AGR			16.34				
Q2.2E	Provided constructive criticism of work	NR				26.73			
Q2.2D	Knowledgeable about the field	NR				26.73			
Q2.2A	Spent time to advise me on academic matters	NR				26.48			
Q2.2B	Encouraged regular contacts	NR				26.48			
Q2.2F	Returned work in a timely manner	NR				26.48			
Q2.2H	Overall relationship was good	NR				26.28			
Q2.2G	Knowledgeable about GSO regulations	NR				26.14			
Q2.2C	Encouraged submission of work	NR				24.67			

S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-3 GSES 2000-02 Department Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES					
			1	2	3	4	5	6
		Count	326	73	52	185	158	18
		Percent	40.2%	9.0%	6.4%	22.8%	19.5%	2.2%
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6
Q3.1U	Concern for intellectual property issues	SAT	13.46					
Q3.1P	Contacts with chair	SAT	12.05					
Q3.1T	Concern for equity issues	SAT	12.03					
Q3.1O	Contacts with graduate officer	SAT	11.99					
Q3.1S	Evaluation of student performance	SAT	9.06					
Q3.1Q	Contacts with office staff	SAT	8.37					
Q3.1B	Program advising/guidance	SAT	8.27					
Q3.1M	Office space	SAT	6.98					
Q3.1N	Assistance in finding employment	SAT	6.74					
Q3.1L	Computing facilities	SAT	6.58					
Q3.1I	General "climate" for graduate study	SAT	6.18			5.96		
Q3.1D	Quality of teaching	SAT	5.71					
Q3.1J	Laboratory/studies facilities & equipment	SAT	5.43					
Q3.1N	Assistance in finding employment	DIS	5.03					
Q3.1C	Advice during orientation	SAT	4.95					
Q3.1F	Appropriateness of degree requirements	SAT	4.80			3.52		
Q3.1M	Office space	DIS	3.44					
Q3.1J	Laboratory/studies facilities & equipment	DIS	3.41					
Q3.1P	Contacts with chair	DIS	3.24		3.66			
Q3.1A	Appropriateness of course offerings	DIS	2.75					
Q3.1C	Advice during orientation	DIS	2.63		3.65			
Q3.1O	Contacts with graduate officer	DIS	2.60		4.09			
Q3.1P	Contacts with chair	NR		14.16				7.82

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-3 GSES 2000-02 Department Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1U	Concern for intellectual property issues	NR		12.47				7.82	
Q3.1T	Concern for equity issues	NR		12.30				8.47	
Q3.1O	Contacts with graduate officer	NR		11.37				9.17	
Q3.1N	Assistance in finding employment	NR		11.02				8.27	
Q3.1M	Office space	NR		9.32				8.56	
Q3.1J	Laboratory/studies facilities & equipment	NR		6.84				7.95	
Q3.1L	Computing facilities	NR		4.93				10.21	
Q3.1Q	Contacts with office staff	NR		3.63				10.75	
Q3.1C	Advice during orientation	NR		3.36				9.39	
Q3.1S	Evaluation of student performance	NR		2.80				9.79	
Q3.1A	Appropriateness of course offerings	SAT		2.49		3.35			
Q3.1S	Evaluation of student performance	V DIS			10.40				
Q3.1B	Program advising/guidance	V DIS			9.42				
Q3.1A	Appropriateness of course offerings	V DIS			9.03				
Q3.1I	General "climate" for graduate study	V DIS			8.70				
Q3.1O	Contacts with graduate officer	V DIS			7.63				
Q3.1N	Assistance in finding employment	V DIS			7.60				
Q3.1P	Contacts with chair	V DIS			7.22				
Q3.1C	Advice during orientation	V DIS			6.96				
Q3.1I	General "climate" for graduate study	DIS			6.22				
Q3.1U	Concern for intellectual property issues	V DIS			5.55				
Q3.1D	Quality of teaching	DIS			5.52				
Q3.1Q	Contacts with office staff	DIS			5.42				
Q3.1L	Computing facilities	V DIS			5.12				
Q3.1J	Laboratory/studies facilities & equipment	V DIS			5.04				
Q3.1F	Appropriateness of degree requirements	DIS			4.34				
Q3.1U	Concern for intellectual property issues	DIS			4.25				

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-3 GSES 2000-02 Department Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1F	Appropriateness of degree requirements	V DIS			3.87				
Q3.1T	Concern for equity issues	DIS			3.64				
Q3.1B	Program advising/guidance	DIS			3.61				
Q3.1M	Office space	V DIS			3.06				
Q3.1L	Computing facilities	DIS			2.65				
Q3.1S	Evaluation of student performance	DIS			2.45				
Q3.1P	Contacts with chair	NOP				18.40			
Q3.1O	Contacts with graduate officer	NOP				17.09			
Q3.1M	Office space	NOP				15.44			
Q3.1T	Concern for equity issues	NOP				14.92			
Q3.1U	Concern for intellectual property issues	NOP				14.40			
Q3.1N	Assistance in finding employment	NOP				13.09			
Q3.1S	Evaluation of student performance	NOP				10.77			
Q3.1Q	Contacts with office staff	NOP				10.07			
Q3.1L	Computing facilities	NOP				9.48			
Q3.1C	Advice during orientation	NOP				9.40			
Q3.1J	Laboratory/studies facilities & equipment	NOP				8.12			
Q3.1B	Program advising/guidance	NOP				7.02			
Q3.1I	General "climate" for graduate study	V SAT					15.42		
Q3.1S	Evaluation of student performance	V SAT					14.33		
Q3.1U	Concern for intellectual property issues	V SAT					13.42		
Q3.1B	Program advising/guidance	V SAT					13.29		
Q3.1F	Appropriateness of degree requirements	V SAT					12.93		
Q3.1T	Concern for equity issues	V SAT					12.88		
Q3.1O	Contacts with graduate officer	V SAT					12.60		
Q3.1P	Contacts with chair	V SAT					12.52		
Q3.1Q	Contacts with office staff	V SAT					12.39		

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-3 GSES 2000-02 Department Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1L	Computing facilities	V SAT					11.39		
Q3.1C	Advice during orientation	V SAT					10.78		
Q3.1M	Office space	V SAT					10.38		
Q3.1A	Appropriateness of course offerings	V SAT					9.37		
Q3.1J	Laboratory/studies facilities & equipment	V SAT					8.87		
Q3.1D	Quality of teaching	V SAT					7.97		
Q3.1N	Assistance in finding employment	V SAT					5.44		
Q3.1A	Appropriateness of course offerings	NR						12.46	
Q3.1B	Program advising/guidance	NR						11.98	
Q3.1I	General "climate" for graduate study	NR						11.07	
Q3.1F	Appropriateness of degree requirements	NR						10.78	
Q3.1D	Quality of teaching	NR						10.53	

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-4 GSES 2000-02 Involvement Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	413	84	111	204			
		Percent	50.9%	10.3%	13.7%	25.1%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q9.2CA	Participation in teaching certificate	NO	25.60						
Q9.1	Assigned teaching assistantship duties	YES	21.79	7.82					
Q9.3	Teaching certificate completion	NO	21.73						
Q8.1	Research thesis degree	YES	18.65	5.74	7.93				
Q9.2BA	Participation in trace programs	NO	17.86						
Q8.7A	Conference attendance	YES	11.96	5.69					
Q8.10A	Publications in refereed journal	NO	10.67		5.62				
Q9.2AA	Participation in department seminar	NO	10.41						
Q8.8A	Conference paper presentation	YES	9.82	4.37					
Q9.2AA	Participation in department seminar	YES	9.70	7.30					
Q8.6	Other research presentation at Waterloo	YES	8.51	4.28					
Q8.6	Other research presentation at Waterloo	NO	6.24		5.85				
Q8.10A	Publications in refereed journal	YES	4.59	5.11					
Q8.8A	Conference paper presentation	NO	3.48						
Q8.7A	Conference attendance	NO	3.14		7.08				
Q12.6	Interested in representing Waterloo	NO	3.05						
Q9.2CA	Participation in teaching certificate	YES		21.34					
Q9.3	Teaching certificate completion	YES		11.78					
Q9.2BA	Participation in trace programs	YES		11.10					
Q12.6	Interested in representing Waterloo	YES		4.65					
Q9.2CA	Participation in teaching certificate	NR			15.98	13.73			
Q9.2AA	Participation in department seminar	NR			15.94	14.08			
Q9.2BA	Participation in trace programs	NR			15.89	13.82			

PT = Part-time; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-4 GSES 2000-02 Involvement Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q9.1	Assigned teaching assistantship duties	NO			11.55	14.08			
Q9.3	Teaching certificate completion	NR			11.55	14.08			
Q9.1	Assigned teaching assistantship duties	NR			5.93				
Q6.5	Term enrolment	PT			3.59	3.97			
Q12.6	Interested in representing Waterloo	NR			2.35				
Q8.7A	Conference attendance	NR				29.02			
Q8.10A	Publications in refereed journal	NR				28.29			
Q8.6	Other research presentation at Waterloo	NR				28.19			
Q8.1	Research thesis degree	NO				26.23			
Q8.8A	Conference paper presentation	NR				19.10			
Q8.1	Research thesis degree	NR				6.72			

PT = Part-time; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-5 GSES 2000-02 Reasons to Choose Waterloo Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	99	48	38	627			
		Percent	12.2%	5.9%	4.7%	77.2%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q6.8D	Chose Waterloo - location	NR	24.22						
Q6.8F	Chose Waterloo - research expertise	NR	24.22						
Q6.8H	Chose Waterloo - co- op	NR	24.22						
Q6.8C	Chose Waterloo - program/courses	NR	24.22						
Q6.8E	Chose Waterloo - previous degree	NR	24.22						
Q6.8I	Chose Waterloo - other	NR	24.22						
Q6.8B	Chose Waterloo - supervisor	NR	24.22						
Q6.8G	Chose Waterloo - financial support	NR	24.22						
Q6.8A	Chose Waterloo - reputation	NR	24.22						
Q6.8H	Chose Waterloo - co- op	YES		99.99					
Q6.8B	Chose Waterloo - supervisor	NO		4.94		9.47			
Q6.8E	Chose Waterloo - previous degree	NO		4.75	2.90	10.31			
Q6.8F	Chose Waterloo - research expertise	NO		3.71		22.71			
Q6.8A	Chose Waterloo - reputation	NO		3.69		4.77			
Q6.8C	Chose Waterloo - program/courses	YES		3.63		3.43			
Q6.8I	Chose Waterloo - other	NO		2.53	2.95	12.60			
Q6.8F	Chose Waterloo - research expertise	YES			99.99				
Q6.8D	Chose Waterloo - location	NO			2.90	11.77			
Q6.8H	Chose Waterloo - co- op	NO			2.66	24.00			
Q6.8G	Chose Waterloo - financial support	NO				14.22			
Q6.8C	Chose Waterloo - program/courses	NO				9.68			
Q6.8A	Chose Waterloo - reputation	YES				8.54			
Q6.8E	Chose Waterloo - previous degree	YES				6.02			

NR = Item nonresponse (no answer, skipped, etc)

Table C-5 GSES 2000-02 Reasons to Choose Waterloo Cluster Loadings (Cont'd)

SECT.	ITEM LABEL	SCALE	SATISFACTION CLUSTERS 1	2	3	4	5	6	7
Q6.8B	Chose Waterloo - supervisor	YES				5.18			
Q6.8I	Chose Waterloo - other	YES				3.59			
Q6.8D	Chose Waterloo - location	YES				3.24			

NR = Item nonresponse (no answer, skipped, etc)

Table C-6 GSES 2000-02 Item Nonresponse Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES		
			1	2	3
		Count	453	327	32
		Percent	55.8%	40.3%	3.9%
SECT.	ITEM LABEL	SCALE	1	2	3
Q9.4	Resources adequately available for ta duties	ANS	25.55		
Q10.1	Financial Support- All years All Sources	ANS	15.08		
Q10.3	Equitable financial support distribution	ANS	12.38		
Q8.9A	Conference presentation funded fully	ANS	11.82		
Q2.3A	Satisfaction with authorship issues	ANS	11.24		
Q10.4	International student registration	ANS	6.99		
Q12.4	Stated as visible minority	ANS	6.94	5.35	
Q11.2A	Seeking employement	NR	3.45		2.98
Q12.5BB	Previous degree location within Canada	ANS	3.16		
Q9.4	Resources adequately available for ta duties	NR		25.34	
Q10.1	Financial Support- All years All Sources	NR		14.64	
Q8.9A	Conference presentation funded fully	NR		11.99	
Q10.3	Equitable financial support distribution	NR		11.77	
Q2.3A	Satisfaction with authorship issues	NR		11.34	
Q11.2A	Seeking employement	ANS		5.04	
Q10.4	International student registration	NR		3.61	4.97
Q12.4	Stated as visible minority	NR			99.99
Q12.5BB	Previous degree location within Canada	NR			6.88

ANS = Answered; NR = Item nonresponse (no answer, skipped, etc)

Table C-7 NGS 95 - GSES 2000-02 Combined Global Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	1719	5971	989	800			
		Percent	18.1%	63.0%	10.4%	8.4%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q3.1L	Computing facilities	V SAT	84.37						
Q3.1D	Quality of teaching	V SAT	16.30	4.36					
Q6.9AA	I would choose Waterloo again	YES	13.42	22.96					
Q6.9B	I would choose the same program	YES	8.34	14.88					
Q3.1L	Computing facilities	SAT		34.39					
Q3.1D	Quality of teaching	SAT		27.03					
Q3.1L	Computing facilities	NR		14.64					
Q3.1L	Computing facilities	DIS		12.86	8.76				
Q3.1D	Quality of teaching	DIS			71.28	2.62			
Q6.9AA	I would choose Waterloo again	NO			22.08				
Q6.9B	I would choose the same program	NO			10.09				
Q6.9AA	I would choose Waterloo again	NR				52.45			
Q6.9B	I would choose the same program	NR				50.78			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-8 NGS 95 - GSES 2000-02 Combined Item Nonresponse Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES		
			1	2	3
		Count	1163	7930	386
		Percent	12.3%	83.7%	4.1%
SECT.	ITEM LABEL	SCALE	1	2	3
Q12.4	Stated as visible minority	NR	81.24		
Q6.1	Doctoral Program	NR	9.85		56.63
Q10.4	International student registration	NR	4.24		8.65
Q6.6	Deviation from normal program duration	NR	3.78		42.44
Q12.4	Stated as visible minority	ANS		75.13	
Q6.1	Doctoral Program	ANS		38.45	
Q6.6	Deviation from normal program duration	ANS		23.50	
Q12.1	Alumni gender	ANS		5.16	

ANS = Answered; NR = Item nonresponse (no answer, skipped, etc)

Table C-9 GSES Factorial Global Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	79	6	20	4			
		Percent	72.5%	5.5%	18.4%	3.7%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q3.1D	Quality of teaching	SAT	4.64						
Q3.1I	General "climate" for graduate study	SAT	4.51						
Q3.1U	Concern for intellectual property issues	SAT	4.34						
Q3.1S	Evaluation of student performance	SAT	4.18						
Q3.1H	Relationship with supervisory committee	SAT	4.12						
Q2.2A	Spent time to advise me on academic matters	AGR	4.01						
Q1.1A	Thesis submission	SAT	3.78						
Q1.1D	Admissions procedures	SAT	3.68						
Q3.1B	Program advising/guidance	SAT	3.66						
Q3.1F	Appropriateness of degree requirements	SAT	3.48						
Q2.2H	Overall relationship was good	AGR	3.30						
Q1.1E	Admissions staff responsiveness	SAT	3.28						
Q3.1P	Contacts with chair	SAT	3.25						
Q3.1Q	Contacts with office staff	SAT	3.20						
Q2.2C	Encouraged submission of work	AGR	2.97						
Q3.1R	Oral defence process	SAT	2.96						
Q3.1T	Concern for equity issues	SAT	2.96						
Q2.2E	Provided constructive criticism of work	AGR	2.85						
Q1.1F	Registration procedures	SAT	2.78						
Q3.1O	Contacts with graduate officer	SAT	2.74						
Q1.1B	Advice accessibility and helpfulness	SAT	2.72						
Q3.1E	Quality of research supervision	SAT	2.64						
Q4.1D	Safety on campus	SAT	2.63						

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-9 GSES Factorial Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q1.1G	Concern for equity issues	SAT	2.45						
Q3.1K	Field work facilities	SAT	2.40						
Q2.2D	Knowledgeable about the field	NR		5.19		3.96			
Q2.2G	Knowledgeable about GSO regulations	NR		5.19		3.96			
Q2.2E	Provided constructive criticism of work	NR		5.19		3.96			
Q2.2A	Spent time to advise me on academic matters	NR		5.19		3.96			
Q2.2F	Returned work in a timely manner	NR		5.19		3.96			
Q2.2H	Overall relationship was good	NR		5.19		3.96			
Q2.2B	Encouraged regular contacts	NR		5.19		3.96			
Q2.2C	Encouraged submission of work	NR		5.04		3.85			
Q1.1A	Thesis submission	NOP		4.67					
Q3.1E	Quality of research supervision	NOP		3.95					
Q3.1B	Program advising/guidance	NOP		2.91					
Q1.1E	Admissions staff responsiveness	NOP		2.74					
Q3.1H	Relationship with supervisory committee	NOP		2.58					
Q1.1G	Concern for equity issues	NOP		2.56					
Q3.1S	Evaluation of student performance	V SAT			6.47				
Q3.1I	General "climate" for graduate study	V SAT			6.39				
Q1.1D	Admissions procedures	V SAT			5.90				
Q3.1H	Relationship with supervisory committee	V SAT			5.83				
Q3.1P	Contacts with chair	V SAT			5.81				
Q3.1U	Concern for intellectual property issues	V SAT			5.70				
Q1.1F	Registration procedures	V SAT			5.42				
Q3.1F	Appropriateness of degree requirements	V SAT			5.14				
Q3.1T	Concern for equity issues	V SAT			5.04				
Q3.1D	Quality of teaching	V SAT			5.03				
Q1.1E	Admissions staff responsiveness	V SAT			4.80				

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-9 GSES Factorial Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1R	Oral defence process	V SAT			4.73				
Q3.1B	Program advising/guidance	V SAT			4.71				
Q1.1G	Concern for equity issues	V SAT			4.39				
Q3.1A	Appropriateness of course offerings	V SAT			4.32				
Q1.1A	Thesis submission	V SAT			4.27				
Q3.1N	Assistance in finding employment	SAT			4.20				
Q1.1C	Scholarships advice	V SAT			4.16				
Q3.1O	Contacts with graduate officer	V SAT			4.16				
Q1.1B	Advice accessibility and helpfulness	V SAT			4.10				
Q3.1Q	Contacts with office staff	V SAT			3.73				
Q3.1C	Advice during orientation	V SAT			3.71				
Q9.5	Satisfaction with instructors supervision	V SAT			3.69				
Q2.2C	Encouraged submission of work	S AGR			3.12				
Q3.1L	Computing facilities	V SAT			2.85				
Q2.2E	Provided constructive criticism of work	S AGR			2.83				
Q2.2H	Overall relationship was good	S AGR			2.83				
Q2.2A	Spent time to advise me on academic matters	S AGR			2.82				
Q3.1E	Quality of research supervision	V SAT			2.72				
Q6.9AA	I would choose Waterloo again	YES			2.66				
Q2.2F	Returned work in a timely manner	S AGR			2.62				
Q3.1O	Contacts with graduate officer	NR				5.09			
Q3.1L	Computing facilities	NR				5.09			
Q3.1P	Contacts with chair	NR				5.09			
Q3.1A	Appropriateness of course offerings	NR				5.09			
Q3.1I	General "climate" for graduate study	NR				5.09			
Q3.1Q	Contacts with office staff	NR				5.09			
Q3.1N	Assistance in finding employment	NR				5.09			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-9 GSES Factorial Global Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1M	Office space	NR				5.09			
Q3.1S	Evaluation of student performance	NR				5.09			
Q3.1J	Laboratory/studies facilities & equipment	NR				4.77			
Q3.1B	Program advising/guidance	NR				4.77			
Q3.1T	Concern for equity issues	NR				4.77			
Q3.1E	Quality of research supervision	NR				4.77			
Q3.1U	Concern for intellectual property issues	NR				4.77			
Q3.1C	Advice during orientation	NR				4.77			
Q3.1D	Quality of teaching	NR				4.77			
Q3.1H	Relationship with supervisory committee	NR				4.77			
Q3.1F	Appropriateness of degree requirements	NR				4.77			
Q3.1R	Oral defence process	NR				4.37			
Q3.1K	Field work facilities	NR				4.21			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-10 GSES Factorial Supervisor Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	56	3	40	10			
		Percent	51.4%	2.8%	36.7%	9.2%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q2.2E	Provided constructive criticism of work	S AGR	9.48						
Q2.2D	Knowledgeable about the field	S AGR	8.70						
Q2.2H	Overall relationship was good	S AGR	8.48						
Q2.2F	Returned work in a timely manner	S AGR	8.10						
Q2.2A	Spent time to advise me on academic matters	S AGR	7.99						
Q2.2B	Encouraged regular contacts	S AGR	7.17						
Q2.2C	Encouraged submission of work	S AGR	6.71						
Q2.2G	Knowledgeable about GSO regulations	S AGR	4.99						
Q2.2A	Spent time to advise me on academic matters	S DSG		4.43					
Q2.2F	Returned work in a timely manner	S DSG		3.46					
Q2.2B	Encouraged regular contacts	S DSG		2.96					
Q2.2H	Overall relationship was good	DSG		2.75					
Q2.2E	Provided constructive criticism of work	DSG		2.43					
Q2.2E	Provided constructive criticism of work	AGR			7.32				
Q2.2H	Overall relationship was good	AGR			7.24				
Q2.2A	Spent time to advise me on academic matters	AGR			6.52				
Q2.2D	Knowledgeable about the field	AGR			5.80				
Q2.2B	Encouraged regular contacts	AGR			4.38				
Q2.2F	Returned work in a timely manner	AGR			4.23				
Q2.2F	Returned work in a timely manner	DSG			4.14				
Q2.2C	Encouraged submission of work	DSG			3.68				
Q2.2B	Encouraged regular contacts	DSG			2.84				
Q2.2C	Encouraged submission of work	AGR			2.81				

S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-10 GSES Factorial Supervisor Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q2.2H	Overall relationship was good	NR				7.54			
Q2.2D	Knowledgeable about the field	NR				7.54			
Q2.2B	Encouraged regular contacts	NR				7.54			
Q2.2G	Knowledgeable about GSO regulations	NR				7.54			
Q2.2A	Spent time to advise me on academic matters	NR				7.54			
Q2.2F	Returned work in a timely manner	NR				7.54			
Q2.2E	Provided constructive criticism of work	NR				7.54			
Q2.2C	Encouraged submission of work	NR				7.22			

S AGR = Strongly Agree; AGR = Agree; DSG = Disagree; S DSG = Strongly Disagree; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-11 GSES Factorial Department Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	72	10	23	4			
		Percent	66.1%	9.2%	21.1%	3.7%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q3.1B	Program advising/guidance	SAT	5.23						
Q3.1I	General "climate" for graduate study	SAT	5.05						
Q3.1Q	Contacts with office staff	SAT	4.84						
Q3.1S	Evaluation of student performance	SAT	4.77						
Q3.1T	Concern for equity issues	SAT	4.34						
Q3.1O	Contacts with graduate officer	SAT	4.31						
Q3.1N	Assistance in finding employment	NOP	4.26						
Q3.1D	Quality of teaching	SAT	4.15						
Q3.1U	Concern for intellectual property issues	SAT	4.11						
Q3.1L	Computing facilities	SAT	3.85						
Q3.1P	Contacts with chair	SAT	3.33						
Q3.1F	Appropriateness of degree requirements	SAT	2.80						
Q3.1A	Appropriateness of course offerings	SAT	2.79						
Q3.1C	Advice during orientation	SAT	2.44						
Q3.1U	Concern for intellectual property issues	DIS		3.96					
Q3.1A	Appropriateness of course offerings	V DIS		3.94					
Q3.1N	Assistance in finding employment	V DIS		3.85					
Q3.1Q	Contacts with office staff	DIS		3.70					
Q3.1B	Program advising/guidance	V DIS		3.57					
Q3.1I	General "climate" for graduate study	V DIS		3.57					
Q3.1L	Computing facilities	V DIS		3.57					
Q3.1O	Contacts with graduate officer	DIS		3.57					
Q3.1J	Laboratory/studies facilities & equipment	V DIS		3.28					

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-11 GSES Factorial Department Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1D	Quality of teaching	DIS		3.16					
Q3.1C	Advice during orientation	DIS		3.02					
Q3.1F	Appropriateness of degree requirements	DIS		2.88					
Q3.1T	Concern for equity issues	V DIS		2.85					
Q3.1S	Evaluation of student performance	V DIS		2.85					
Q3.1T	Concern for equity issues	DIS		2.85					
Q3.1C	Advice during orientation	V DIS		2.85					
Q3.1S	Evaluation of student performance	V SAT			6.53				
Q3.1B	Program advising/guidance	V SAT			6.11				
Q3.1I	General "climate" for graduate study	V SAT			5.72				
Q3.1P	Contacts with chair	V SAT			5.41				
Q3.1C	Advice during orientation	V SAT			5.38				
Q3.1U	Concern for intellectual property issues	V SAT			5.25				
Q3.1D	Quality of teaching	V SAT			5.00				
Q3.1Q	Contacts with office staff	V SAT			5.00				
Q3.1T	Concern for equity issues	V SAT			4.69				
Q3.1F	Appropriateness of degree requirements	V SAT			4.53				
Q3.1A	Appropriateness of course offerings	V SAT			4.53				
Q3.1O	Contacts with graduate officer	V SAT			4.27				
Q3.1N	Assistance in finding employment	SAT			3.66				
Q3.1N	Assistance in finding employment	V SAT			3.43				
Q3.1L	Computing facilities	V SAT			3.15				
Q3.1J	Laboratory/studies facilities & equipment	NOP			2.70				
Q3.1L	Computing facilities	NOP			2.62				
Q3.1M	Office space	V SAT			2.53				
Q3.1O	Contacts with graduate officer	NR				5.09			
Q3.1I	General "climate" for graduate study	NR				5.09			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-11 GSES Factorial Department Experience Cluster Loadings (cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q3.1N	Assistance in finding employment	NR				5.09			
Q3.1M	Office space	NR				5.09			
Q3.1P	Contacts with chair	NR				5.09			
Q3.1Q	Contacts with office staff	NR				5.09			
Q3.1A	Appropriateness of course offerings	NR				5.09			
Q3.1S	Evaluation of student performance	NR				5.09			
Q3.1L	Computing facilities	NR				5.09			
Q3.1B	Program advising/guidance	NR				4.77			
Q3.1U	Concern for intellectual property issues	NR				4.77			
Q3.1D	Quality of teaching	NR				4.77			
Q3.1J	Laboratory/studies facilities & equipment	NR				4.77			
Q3.1T	Concern for equity issues	NR				4.77			
Q3.1F	Appropriateness of degree requirements	NR				4.77			
Q3.1C	Advice during orientation	NR				4.77			

V SAT = Very Satisfied; SAT = Satisfied; DIS = Dissatisfied; V DIS = Very Dissatisfied; NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-12 GSES Factorial Involvement Experience Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES		
			1	2	3
		Count	66	19	24
		Percent	60.6%	17.4%	22.0%
SECT.	ITEM LABEL	SCALE	1	2	3
Q9.1	Assigned teaching assistantship duties	YES	8.50		
Q8.1	Research thesis degree	YES	7.04	2.55	
Q9.2CA	Participation in teaching certificate	NO	6.06		
Q9.3	Teaching certificate completion	NO	5.84		
Q8.7A	Conference attendance	YES	5.26		
Q8.6	Other research presentation at Waterloo	YES	5.09		
Q9.2AA	Participation in department seminar	YES	5.09		
Q8.8A	Conference paper presentation	YES	4.97		
Q9.2BA	Participation in trace programs	NO	4.22		
Q9.2BA	Participation in trace programs	YES	3.27		
Q8.10A	Publications in refereed journal	NO	2.91		
Q8.10A	Publications in refereed journal	YES	2.75		
Q9.2AA	Participation in department seminar	NO	2.40		
Q9.2AA	Participation in department seminar	NR		6.92	2.45
Q9.2BA	Participation in trace programs	NR		6.80	
Q9.2CA	Participation in teaching certificate	NR		6.49	2.45
Q9.3	Teaching certificate completion	NR		4.68	2.78
Q9.1	Assigned teaching assistantship duties	NO		4.68	2.78
Q8.6	Other research presentation at Waterloo	NO		3.37	
Q8.7A	Conference attendance	NO		3.05	
Q9.1	Assigned teaching assistantship duties	NR		2.94	
Q8.7A	Conference attendance	NR			10.18
Q8.6	Other research presentation at Waterloo	NR			10.18

NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-12 GSES Factorial Involvement Experience Cluster Loadings (Cont'd)

			SATISFACTION CLUSTERS						
SECT.	ITEM LABEL	SCALE	1	2	3	4	5	6	7
Q8.10A	Publications in refereed journal	NR			9.86				
Q8.1	Research thesis degree	NO			7.97				
Q8.8A	Conference paper presentation	NR			5.54				
Q8.1	Research thesis degree	NR			3.82				
Q6.5	Term Enrolment	PT			3.05				

NOP = No Opinion; NR = Item nonresponse (no answer, skipped, etc)

Table C-13 GSES Factorial Reasons to Choose Waterloo Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES						
			1	2	3	4			
		Count	80	17	6	6			
		Percent	73.4%	15.6%	5.5%	5.5%			
SECT.	ITEM LABEL	SCALE	1	2	3	4			
Q6.8I	Chose Waterloo - other	NO	7.53						
Q6.8F	Chose Waterloo - research expertise	NO	5.52						
Q6.8H	Chose Waterloo - co- op	NO	5.22						
Q6.8A	Chose Waterloo - reputation	YES	4.26						
Q6.8I	Chose Waterloo - other	YES		7.35					
Q6.8H	Chose Waterloo - co- op	YES		3.88					
Q6.8A	Chose Waterloo - reputation	NO		2.66					
Q6.8F	Chose Waterloo - research expertise	YES			6.11				
Q6.8B	Chose Waterloo - supervisor	YES			2.90				
Q6.8C	Chose Waterloo - program/courses	NR				6.11			
Q6.8E	Chose Waterloo - previous degree	NR				6.11			
Q6.8H	Chose Waterloo - co- op	NR				6.11			
Q6.8G	Chose Waterloo - financial support	NR				6.11			
Q6.8B	Chose Waterloo - supervisor	NR				6.11			
Q6.8D	Chose Waterloo - location	NR				6.11			
Q6.8F	Chose Waterloo - research expertise	NR				6.11			
Q6.8A	Chose Waterloo - reputation	NR				6.11			
Q6.8I	Chose Waterloo - other	NR				6.11			

NR = Item nonresponse (no answer, skipped, etc)

Table C-14 GSES Factorial Item Nonresponse Cluster Loadings

			SATISFACTION CLUSTERS TEST-VALUES		
			1	2	3
		Count	15	65	29
		Percent	13.8%	59.6%	26.6%
SECT.	ITEM LABEL	SCALE	1	2	3
Q11.2A	Seeking employement	NR	8.76		
Q10.3	Equitable financial support distribution	ANS		5.67	
Q10.1	Financial Support- All years All Sources	ANS		5.24	
Q11.2A	Seeking employement	ANS		5.05	2.48
Q9.4	Resources adequately available for ta duties	ANS		4.32	
Q8.9A	Conference presentation funded fully	ANS		3.60	
Q10.1	Financial Support- All years All Sources	NR			6.72
Q9.4	Resources adequately available for ta duties	NR			5.39
Q10.3	Equitable financial support distribution	NR			5.10
Q8.9A	Conference presentation funded fully	NR			4.52
Q2.3A	Satisfaction with authorship issues	NR			2.59

ANS = Answered; NR = Item nonresponse (no answer, skipped, etc)

Appendix D

Multiple Discrepancies Model Descriptive Statistics

Table D-1 GSES 2000-02 Combined Student-Level Model Descriptive Statistics

Student-Level Model Variables	Alumni Level				
	Mean	sd	N	Min. Score	Max. Score
Satisfaction Outcomes					
Cluster 1: Global Experience - Satisfied (PhD)	0.143	0.350	812	0	1
Cluster 2: Global Experience - Very Satisfied	0.190	0.392	812	0	1
Cluster 3: Global Experience - Satisfied (Master's)	0.229	0.420	812	0	1
Cluster 4: Global Experience - Very Dissatisfied	0.067	0.249	812	0	1
Cluster 5: Global Experience - Satisfied (Master's)	0.188	0.391	812	0	1
Cluster 7: Global Experience - Satisfied (Macc)	0.161	0.368	812	0	1
Cluster 1: Supervisor Experience - Satisfied	0.421	0.494	812	0	1
Cluster 2: Supervisor Experience - Dissatisfied	0.111	0.314	812	0	1
Cluster 3: Supervisor Experience - Very Satisfied	0.283	0.451	812	0	1
Cluster 1: Department Experience - Satisfied	0.402	0.491	812	0	1
Cluster 2: Department Experience - No Answer	0.090	0.286	812	0	1
Cluster 3: Department Experience - Very Dissatisfied	0.064	0.245	812	0	1
Cluster 4: Department Experience - No Opinion	0.228	0.420	812	0	1
Cluster 5: Department Experience - Very Satisfied	0.195	0.396	812	0	1
Subjective Discrepancies					
Program completion longer than expected	0.351	0.477	790	0	1
Not satisfied with superv. on authorship issues	0.091	0.288	550	0	1
Financial support distribution not fair/equitable	0.281	0.450	598	0	1
Resources not adequately available for TA duties	0.183	0.387	471	0	1
University expectations before enrolment	0.772	0.420	812	0	1
Coop program expectations before enrolment	0.059	0.236	812	0	1
Research expert. expectations before enrolment	0.047	0.211	812	0	1

373

Table D-1 GSES 2000-02 Combined Student-Level Model Descriptive Statistics (cont'd)

Student-Level Model Variables	Alumni-Level				
	Mean	sd	N	Min. Score	Max. Score
Objective Discrepancies					
Conference expenses partially/not funded	0.332	0.472	319	0	1
Reported loans at graduation	0.384	0.487	812	0	1
Reported unemployment at convocation	0.157	0.364	688	0	1
Deviation from avrg program duration (mths)	1.019	14.052	753	-62	108
Student with visa status	0.128	0.334	775	0	1
Not an UW alumni before enrolment	0.571	0.495	759	0	1
Conditionners					
Women alumni	0.448	0.498	801	0	1
Visible minority status	0.291	0.455	780	0	1
Married during degree program	0.363	0.481	787	0	1
With dependents during degree program	0.183	0.387	787	0	1
Master's of accounting alumni	0.128	0.334	804	0	1
Doctoral alumni	0.193	0.395	804	0	1
Science degree	0.589	0.492	774	0	1
Financial support (K $)	32.833	35.456	535	0.7	333
Frequent visits to the Graduate House	0.429	0.495	795	0	1
Full research or teaching involvement	0.103	0.305	812	0	1
No research or teaching involvement	0.137	0.344	812	0	1
Item Nonresponse					
Missingness on program characteristics	0.403	0.491	812	0	1
Missingness on socio- demographics	0.039	0.195	812	0	1

Table D-2 GSES 2000-02 Combined Student-Level NGS-GSES Model Descriptive Statistics

Student-Level Model Variables	GSES Alumni-Level				
	Mean	sd	N	Min. Score	Max. Score
Satisfaction Outcomes					
Cluster 1: Global Experience - Very Satisfied	0.193	0.395	812	0	1
Cluster 2: Global Experience - Satisfied	0.596	0.491	812	0	1
Cluster 3: Global Experience - Very Dissatisfied	0.094	0.291	812	0	1
Cluster 4: Global Experience - No Answer	0.117	0.322	812	0	1
Subjective Discrepancies					
Objective Discrepancies					
Reported loans at graduation	0.384	0.487	812	0	1
Reported unemployment at convocation	0.157	0.364	688	0	1
Deviation from avrg program duration (mths)	1.019	14.052	753	-62	108
Student with visa status	0.128	0.334	775	0	1
Conditionners					
Women alumni	0.448	0.498	801	0	1
Visible minority status	0.291	0.455	780	0	1
Master's of accounting alumni	0.128	0.334	804	0	1
Doctoral alumni	0.193	0.395	804	0	1
Science degree	0.589	0.492	774	0	1
Part- time studies (proxy for no involvement)	0.068	0.251	753	0	1
Item Nonresponse					
Missingness on program characteristics	0.010	0.099	812	0	1
Missingness on socio- demographics	0.038	0.192	812	0	1

Table D-3 NGS Student-Level NGS-GSES Model Descriptive Statistics

Student-Level Model Variables	NGS Alumni-Level				
	Mean	sd	N	Min. Score	Max. Score
Satisfaction Outcomes					
Cluster 1: Global Experience - Very Satisfied	0.205	0.404	8667	0	1
Cluster 2: Global Experience - Satisfied	0.611	0.487	8667	0	1
Cluster 3: Global Experience - Very Dissatisfied	0.095	0.293	8667	0	1
Cluster 4: Global Experience - No Answer	0.089	0.285	8667	0	1
Subjective Discrepancies					
Objective Discrepancies					
Reported loans at graduation	0.264	0.441	8667	0	1
Reported unemployment at convocation	0.568	0.495	6730	0	1
Deviation from avrg program duration (mths)	7.310	17.102	7684	-49	170
Student with visa status	0.055	0.227	6327	0	1
Conditionners					
Women alumni	0.484	0.500	8667	0	1
Visible minority status	0.178	0.383	7538	0	1
Master's of accounting alumni	0.050	0.217	8561	0	1
Doctoral alumni	0.225	0.418	8561	0	1
Science degree	0.391	0.488	8567	0	1
Part- time studies (proxy for no involvement)	0.191	0.393	8667	0	1
Item Nonresponse					
Missingness on program characteristics	0.012	0.110	8667	0	1
Missingness on socio- demographics	0.128	0.334	8667	0	1

Table D-4 GSES Factorial Student-Level Model Descriptive Statistics

Student-Level Model Variables	Alumni Level				
	Mean	sd	N	Min. Score	Max. Score
Satisfaction Outcomes					
Cluster 1: Global Experience - Satisfied	0.725	0.449	109	0	1
Cluster 3: Global Experience - Very Satisfied	0.184	0.389	109	0	1
Cluster 1: Supervisor Experience - Very Satisfied	0.514	0.502	109	0	1
Cluster 2: Supervisor Experience - Dissatisfied	0.028	0.164	109	0	1
Cluster 3: Supervisor Experience - Satisfied	0.367	0.484	109	0	1
Cluster 1: Department Experience - Satisfied	0.661	0.476	109	0	1
Cluster 2: Department Experience - Very Dissatisfied	0.092	0.290	109	0	1
Cluster 3: Department Experience - Very Satisfied	0.211	0.410	109	0	1
Subjective Discrepancies					
Program completion longer than expected	0.391	0.490	105	0	1
Not satisfied with superv. on authorship issues	0.071	0.258	85	0	1
Financial support distribution not fair/equitable	0.324	0.471	74	0	1
Resources not adequately available for TA duties	0.172	0.381	58	0	1
University expectations before enrolment	0.734	0.444	109	0	1
Coop program expectations before enrolment	0.156	0.365	109	0	1
Research expert. expectations before enrolment	0.055	0.229	109	0	1

Table D-4 GSES Factorial Student-Level Model Descriptive Statistics (cont'd)

Student-Level Model Variables	Alumni Level				
	Mean	sd	N	Min. Score	Max. Score
Objective Discrepancies					
Conference expenses partially/not funded	0.350	0.483	40	0	1
Reported loans at graduation	0.312	0.465	109	0	1
Reported unemployment at convocation	0.202	0.404	94	0	1
Deviation from avrg program duration (mths)	0.980	11.322	102	-33	40
Student with visa status	0.165	0.373	103	0	1
Not an UW alumni before enrolment	0.647	0.480	102	0	1
Conditionners					
Women alumni	0.454	0.500	108	0	1
Visible minority status	0.274	0.448	106	0	1
Married during degree program	0.402	0.493	107	0	1
With dependents during degree program	0.215	0.413	107	0	1
Master's of accounting alumni	0.000	0.000	107	0	0
Doctoral alumni	0.178	0.384	107	0	1
Science degree	0.673	0.471	101	0	1
Financial support (K $)	42.217	44.655	78	1.6	333
Frequent visits to the Graduate House	0.426	0.497	108	0	1
Some research or teaching involvement	0.606	0.491	109	0	1
No research or teaching involvement	0.174	0.381	109	0	1
Item Nonresponse					
Missingness on program characteristics	0.138	0.346	109	0	1
Missingness on socio- demographics	0.266	0.444	109	0	1

Table D-5 GSES Department-Level Model Descriptive Statistics

Department-Level Model Variables	Departmental-Level				
	Mean	sd	N	Min. Score	Max. Score
Conditionners					
Departmental response rate	0.377	0.118	36	0.0909	0.714
Alumni pool (proxy for dept. size)	60.200	63.830	36	8	363
Alumni to faculty ratio (graduate and undergraduate combined)	17.012	13.482	36	1.5	66.12
Graduate to undergraduate ratio	0.346	0.436	36	0.0372	2.263
Average departmental budget (averaged 2000- 02, $K)	2838.882	1495.285	36	917.85	7249
Proportion of doctoral alumni	0.177	0.133	36	0	0.524
Proportion of master's of accounting alumni	0.020	0.142	36	0	1
Average grade point average	85.409	2.251	36	80.659	90.11
Average deviation from normal program duration (mths)	0.133	5.273	36	-8.143	18
Proportion of women alumni	0.457	0.213	36	0.1263	0.875
Proportion of alumni with visa status	0.098	0.121	36	0	0.581
Average alumni age	29.272	1.633	36	25.375	32.91

Appendix E

Multiple Discrepancies Model Correlation Statistics

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics

Predictors	1	2	3	4	5	6	7	8
1. Cluster 1: Global Experience - Satisfied (PhD)	1.00							
2. Cluster 2: Global Experience - Very Satisfied	-0.198***	1.00						
3. Cluster 3: Global Experience - Satisfied (Master's)	-0.223***	-0.264***	1.00					
4. Cluster 4: Global Experience - Very Dissatisfied	-0.109**	-0.129***	-0.145***	1.00				
5. Cluster 5: Global Experience - Satisfied (Master's)	-0.197***	-0.233***	-0.263***	-0.129***	1.00			
6. Cluster 7: Global Experience - Satisfied (Macc)	-0.179***	-0.212***	-0.239***	-0.117***	-0.211***	1.00		
7. Cluster 1: Supervisor Experience - Satisfied	0.094**	-0.343***	0.188***	-0.078*	0.501***	-0.374***	1.00	
8. Cluster 2: Supervisor Experience - Dissatisfied	-0.088*	-0.151***	0.172***	0.457***	-0.090**	-0.155***	-0.301***	1.00
9. Cluster 3: Supervisor Experience - Very Satisfied	0.110**	0.679***	-0.102**	-0.135***	-0.289***	-0.276***	-0.536***	-0.222***
10. Cluster 1: Dept. Experience - Satisfied	0.111**	-0.025	0.175***	-0.098**	0.042	-0.209***	0.161***	0.007
11. Cluster 2: Dept. Experience - No Answer	-0.104**	-0.152***	-0.171***	0.020	0.267***	0.178***	0.046	-0.001
12. Cluster 3: Dept. Experience - Very Dissatisfied	-0.078*	-0.114***	-0.071*	0.617***	-0.087*	-0.033	0.001	0.212***
13. Cluster 4: Dept. Experience - No Opinion	-0.088*	-0.053	0.179***	-0.121***	-0.262***	0.329***	-0.208***	-0.033
14. Cluster 5: Dept. Experience - Very Satisfied	0.102**	0.294***	-0.209***	-0.131***	0.113***	-0.173***	0.028	-0.094**
15. Program completion longer than expected	0.055	0.038	0.067	0.130***	0.050	-0.289***	0.132***	0.135***
16. Not satisfied with superv. on authorship issues	-0.034	-0.155***	-0.009	0.389***	-0.051	-0.003	-0.078*	0.360***
17. Financial support distribution not fair/equitable	-0.056	-0.067	0.019	0.186***	-0.073*	0.069*	-0.069	0.117***
18. Resources not adequately available for TA duties	-0.076*	-0.009	-0.005	0.102**	-0.029	0.058	-0.019	0.050
19. University expectations before enrolment	0.004	0.120***	0.024	0.027	-0.076*	-0.081*	-0.096**	0.070*
20. Coop program expectations before enrolment	-0.102**	-0.095**	0.012	-0.046	-0.014	0.231***	-0.023	-0.055

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	1	2	3	4	5	6	7	8
21. Research expert. expectations before enrolment	0.093**	-0.018	-0.024	-0.012	0.057	-0.097**	0.059	0.015
22. Conference expenses partially/not funded	-0.027	0.019	-0.032	0.118***	-0.017	0.004	-0.044	0.107**
23. Reported loans at graduation	-0.040	-0.046	-0.075*	0.094**	0.040	0.101**	-0.002	0.052
24. Reported unemployment at convocation	-0.058	0.011	0.075*	0.098**	0.038	-0.155***	0.042	0.069*
25. Deviation from avrg program duration (mths)	-0.053	0.049	0.096**	-0.019	0.049	-0.127***	0.091**	0.052
26. Student with visa status	0.094**	0.031	-0.053	0.019	0.040	-0.128***	0.016	-0.013
27. Not an UW alumni before enrolment	0.013	0.039	0.051	0.030	0.074*	-0.201***	0.092**	0.047
28. Women alumni	-0.083*	-0.057	-0.009	0.027	0.050	0.100**	-0.035	0.075*
29. Visible minority status	-0.028	-0.075*	-0.071*	0.027	0.037	0.126***	-0.018	-0.072*
30. Married during degree program	0.295***	0.007	-0.059	0.015	0.001	-0.220***	0.160***	-0.063
31. With dependents during degree program	0.182***	0.011	-0.050	-0.030	0.024	-0.125***	0.117***	-0.092**
32. Master's of accounting alumni	-0.157***	-0.185***	-0.207***	-0.103**	-0.183***	0.839***	-0.326***	-0.134***
33. Doctoral alumni	0.840***	-0.172***	-0.226***	0.007	-0.161***	-0.205***	0.049	0.008
34. Science degree	0.160***	0.064	0.109**	0.090**	0.079*	-0.487***	0.191***	0.035
35. Financial support (K $)	0.358***	-0.112***	-0.122***	0.002	0.041	-0.134***	0.122***	-0.033
36. Frequent visits to the Graduate House	0.022	0.069*	0.030	0.059	0.019	-0.186***	0.060	0.090**
37. Full research or teaching involvement	0.104**	0.042	-0.022	0.007	0.033	-0.149***	0.062	-0.030
38. No research or teaching involvement	-0.070*	-0.028	0.099**	-0.005	0.092**	-0.145***	0.103**	0.042
39. Missingness on program characteristics	-0.220***	-0.083*	0.054	-0.068	-0.190***	0.459***	-0.258***	-0.034
40. Missingness on socio- demographics	-0.010	-0.033	0.010	0.048	0.000	-0.037	0.058	0.070*

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	14	16
1. Cluster 1: Global Experience - Satisfied (PhD)								
2. Cluster 2: Global Experience - Very Satisfied								
3. Cluster 3: Global Experience - Satisfied (Master's)								
4. Cluster 4: Global Experience - Very Dissatisfied								
5. Cluster 5: Global Experience - Satisfied (Master's)								
6. Cluster 7: Global Experience - Satisfied (Macc)								
7. Cluster 1: Supervisor Experience - Satisfied								
8. Cluster 2: Supervisor Experience - Dissatisfied								
9. Cluster 3: Supervisor Experience - Very Satisfied	1.00							
10. Cluster 1: Dept. Experience - Satisfied	0.020	1.00						
11. Cluster 2: Dept. Experience - No Answer	-0.188***	-0.257***	1.00					
12. Cluster 3: Dept. Experience - Very Dissatisfied	-0.109**	-0.214***	-0.082*	1.00				
13. Cluster 4: Dept. Experience - No Opinion	0.010	-0.445***	-0.171***	-0.142***	1.00			
14. Cluster 5: Dept. Experience - Very Satisfied	0.202***	-0.403***	-0.154***	-0.129***	-0.267***	1.00	1.00	
15. Program completion longer than expected	0.013	0.052	-0.002	0.044	-0.084*	0.020	0.020	
16. Not satisfied with superv. on authorship issues	-0.162***	-0.010	-0.010	0.193***	-0.020	-0.079*	-0.079*	1.00
17. Financial support distribution not fair/equitable	-0.054	0.012	-0.066	0.159***	0.067	-0.123***	-0.123***	0.143***
18. Resources not adequately available for TA duties	-0.059	-0.009	0.066	0.076*	0.049	-0.128***	-0.128***	0.088*
19. University expectations before enrolment	0.126***	0.014	-0.086*	0.010	-0.006	0.059	0.059	-0.018
20. Coop program expectations before enrolment	-0.123***	-0.003	0.122***	0.041	-0.024	-0.084*	-0.084*	0.016

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	14	16
21. Research expert. expectations before enrolment	0.003	-0.003	-0.029	-0.010	-0.023	0.053	0.053	-0.022
22. Conference expenses partially/not funded	-0.017	0.018	0.049	0.040	0.024	-0.087*	-0.087*	0.123***
23. Reported loans at graduation	-0.092**	0.061	0.062	0.062	-0.079*	-0.043	-0.043	0.050
24. Reported unemployment at convocation	0.030	0.044	-0.034	0.082*	-0.108**	0.030	0.030	0.043
25. Deviation from avrg program duration (mths)	-0.007	-0.033	0.037	-0.048	0.000	0.066	0.066	-0.018
26. Student with visa status	0.090**	0.120***	-0.042	-0.052	-0.087*	-0.005	-0.005	0.036
27. Not an UW alumni before enrolment	0.049	0.038	-0.061	0.006	-0.086*	0.094**	0.094**	0.010
28. Women alumni	-0.084*	0.001	0.094**	0.027	-0.006	-0.059	-0.059	0.063
29. Visible minority status	-0.037	0.046	-0.045	0.042	0.006	-0.062	-0.062	-0.017
30. Married during degree program	0.063	0.066	-0.016	0.005	-0.046	-0.012	-0.012	0.000
31. With dependents during degree program	0.048	0.019	0.011	-0.070*	-0.053	0.077*	0.077*	-0.021
32. Master's of accounting alumni	-0.240***	-0.207***	0.164***	-0.055	0.305***	-0.159***	-0.159***	0.001
33. Doctoral alumni	0.091**	0.101**	-0.078*	-0.039	-0.129***	0.079*	0.079*	0.029
34. Science degree	0.164***	0.118***	-0.165***	0.062	-0.123***	0.067	0.067	-0.067
35. Financial support (K $)	-0.012	0.006	0.012	-0.017	-0.040	0.031	0.031	-0.016
36. Frequent visits to the Graduate House	0.028	0.051	-0.057	0.069*	-0.161***	0.110**	0.110**	0.042
37. Full research or teaching involvement	0.074*	0.093**	-0.064	0.010	-0.107**	0.048	0.048	0.040
38. No research or teaching involvement	-0.043	-0.048	0.075*	-0.016	-0.011	-0.005	-0.005	0.069*
39. Missingness on program characteristics	-0.093**	-0.191***	0.102**	-0.020	0.308***	-0.175***	-0.175***	0.028
40. Missingness on socio- demographics	-0.085*	0.015	-0.019	0.025	-0.019	-0.036	-0.036	0.027

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	17	18	19	20	21	22	23	24
1. Cluster 1: Global Experience - Satisfied (PhD)								
2. Cluster 2: Global Experience - Very Satisfied								
3. Cluster 3: Global Experience - Satisfied (Master's)								
4. Cluster 4: Global Experience - Very Dissatisfied								
5. Cluster 5: Global Experience - Satisfied (Master's)								
6. Cluster 7: Global Experience - Satisfied (Macc)								
7. Cluster 1: Supervisor Experience - Satisfied								
8. Cluster 2: Supervisor Experience - Dissatisfied								
9. Cluster 3: Supervisor Experience - Very Satisfied								
10. Cluster 1: Dept. Experience - Satisfied								
11. Cluster 2: Dept. Experience - No Answer								
12. Cluster 3: Dept. Experience - Very Dissatisfied								
13. Cluster 4: Dept. Experience - No Opinion								
14. Cluster 5: Dept. Experience - Very Satisfied								
15. Program completion longer than expected								
16. Not satisfied with superv. on authorship issues								
17. Financial support distribution not fair/equitable	1.00							
18. Resources not adequately available for TA duties	0.199***	1.00						
19. University expectations before enrolment	-0.011	-0.048	1.00					
20. Coop program expectations before enrolment	0.062	0.109**	-0.461***	1.00				

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	17	18	19	20	21	22	23	24
21. Research expert. expectations before enrolment	-0.043	-0.053	-0.408***	-0.056	1.00			
22. Conference expenses partially/not funded	0.086*	0.083*	-0.020	0.012	-0.039	1.00		
23. Reported loans at graduation	0.078*	0.061	0.007	0.103**	0.005	0.110**	1.00	
24. Reported unemployment at convocation	0.019	-0.028	0.067	-0.050	-0.007	0.017	-0.017	1.00
25. Deviation from avrg program duration (mths)	-0.054	0.011	0.062	-0.050	-0.053	-0.002	-0.048	0.083*
26. Student with visa status	-0.060	-0.065	0.069*	-0.062	-0.051	-0.111***	-0.202***	0.023
27. Not an UW alumni before enrolment	-0.093**	-0.046	0.000	-0.057	0.061	0.002	-0.090**	0.046
28. Women alumni	0.062	0.090**	-0.044	0.090*	0.011	0.099**	0.028	-0.053
29. Visible minority status	0.040	0.018	-0.009	0.101**	-0.079*	-0.023	-0.024	0.041
30. Married during degree program	-0.039	-0.012	0.022	-0.100**	-0.018	-0.032	-0.057	0.017
31. With dependents during degree program	-0.056	-0.028	0.031	-0.077*	-0.012	-0.042	-0.022	0.043
32. Master's of accounting alumni	0.079*	0.005	-0.097**	0.188***	-0.083*	0.000	0.123***	-0.153***
33. Doctoral alumni	-0.041	-0.089*	0.020	-0.121***	0.087*	0.001	-0.026	-0.068
34. Science degree	-0.089*	-0.106**	0.075*	-0.186***	0.027	-0.227***	-0.163***	0.062
35. Financial support (K $)	0.021	-0.082*	0.037	-0.116***	0.094**	-0.034	-0.076*	0.015
36. Frequent visits to the Graduate House	-0.054	0.017	0.041	-0.018	0.061	0.067	0.066	0.020
37. Full research or teaching involvement	-0.008	-0.100**	-0.018	-0.051	0.097**	0.010	0.073*	0.081*
38. No research or teaching involvement	0.016	0.021	0.011	-0.039	0.014	-0.044	-0.108**	-0.069*
39. Missingness on program characteristics	0.064	0.011	-0.021	0.103**	-0.099**	0.006	-0.060	-0.101**
40. Missingness on socio- demographics	0.023	-0.020	-0.056	0.003	-0.045	-0.064	-0.043	-0.003

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	25	26	27	28	29	30	31	32
21. Research expert. expectations before enrolment								
22. Conference expenses partially/not funded								
23. Reported loans at graduation								
24. Reported unemployment at convocation								
25. Deviation from avrg program duration (mths)	1.00							
26. Student with visa status	-0.067	1.00						
27. Not an UW alumni before enrolment	0.031	0.205***	1.00					
28. Women alumni	-0.004	-0.083*	0.007	1.00				
29. Visible minority status	-0.044	0.115***	-0.030	-0.086*	1.00			
30. Married during degree program	0.193***	0.077*	0.138***	-0.079*	-0.064	1.00		
31. With dependents during degree program	0.174***	0.094**	0.107**	-0.110**	-0.012	0.551***	1.00	
32. Master's of accounting alumni	-0.130***	-0.129***	-0.283***	0.075*	0.186***	-0.272***	-0.172***	1.00
33. Doctoral alumni	-0.079*	0.087*	-0.019	-0.060	-0.059	0.312***	0.187***	-0.187***
34. Science degree	0.110**	0.162***	0.065	-0.215***	0.061	0.137***	0.084*	-0.468***
35. Financial support (K $)	0.078*	0.041	-0.010	-0.065	-0.039	0.153***	0.095**	-0.119***
36. Frequent visits to the Graduate House	0.000	-0.002	0.047	-0.024	-0.178***	-0.065	-0.120***	-0.211***
37. Full research or teaching involvement	-0.029	0.069	0.079*	0.023	0.056	0.061	0.064	-0.131***
38. No research or teaching involvement	0.093**	-0.030	0.050	-0.057	-0.036	0.097**	0.134***	-0.151***
39. Missingness on program characteristics	0.047	-0.104**	-0.124***	0.005	0.095**	-0.078*	0.025	0.450***
40. Missingness on socio- demographics	-0.034	0.101**	0.038	-0.049	0.000	0.045	0.036	-0.057

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-1 GSES 2000-02 Combined Student-Level Model Correlation Statistics (Cont'd)

Predictors	33	34	35	36	37	38	39	40
21. Research expert. expectations before enrolment								
22. Conference expenses partially/not funded								
23. Reported loans at graduation								
24. Reported unemployment at convocation								
25. Deviation from avrg program duration (mths)								
26. Student with visa status								
27. Not an UW alumni before enrolment								
28. Women alumni								
29. Visible minority status								
30. Married during degree program								
31. With dependents during degree program								
32. Master's of accounting alumni								
33. Doctoral alumni	1.00							
34. Science degree	0.114***	1.00						
35. Financial support (K $)	0.379***	0.168***	1.00					
36. Frequent visits to the Graduate House	0.045	0.002	-0.007	1.00				
37. Full research or teaching involvement	0.142***	0.115***	0.041	0.000	1.00			
38. No research or teaching involvement	-0.055	0.001	0.023	-0.087*	-0.135***	1.00		
39. Missingness on program characteristics	-0.219***	-0.270***	-0.090**	-0.258***	-0.221***	0.331***	1.00	
40. Missingness on socio- demographics	0.000	0.073*	-0.015	-0.019	-0.027	-0.025	-0.166***	1.00

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-2 GSES 2000-02 Student-Level NGS-GSES Model Correlation Statistics

Predictors	1	2	3	4	5	6	7	8
1. Cluster 1: Global Experience - Very Satisfied	1.00							
2. Cluster 2: Global Experience - Satisfied	-0.595***	1.00						
3. Cluster 3: Global Experience - Very Dissatisfied	-0.157***	-0.390***	1.00					
4. Cluster 4: Global Experience - No Answer	-0.178***	-0.442***	-0.117***	1.00				
5. Reported loans at graduation	-0.066	0.041	0.033	-0.012	1.00			
6. Reported unemployment at convocation	0.006	-0.035	0.044	0.006	-0.017	1.00		
7. Deviation from avrg program duration (mths)	0.028	-0.003	-0.036	0.003	-0.048	0.083*	1.00	
8. Student with visa status	0.073*	-0.094**	0.004	0.052	-0.202***	0.023	-0.067	1.00
9. Women alumni	-0.041	0.039	-0.043	0.029	0.028	-0.053	-0.004	-0.083*
10. Visible minority status	-0.007	0.006	0.065	-0.060	-0.024	0.041	-0.044	0.115***
11. Master's of accounting alumni	-0.083*	0.146***	-0.046	-0.080*	0.123***	-0.153***	-0.130***	-0.129***
12. Doctoral alumni	0.041	-0.069	0.017	0.040	-0.026	-0.068	-0.079*	0.087*
13. Science degree	0.084*	-0.058	0.009	-0.023	-0.163***	0.062	0.110**	0.162***
14. Part- time studies (proxy for no involvement)	0.016	0.022	-0.063	0.004	-0.134***	-0.029	0.179***	-0.085*
15. Missingness on program characteristics	0.014	-0.045	0.011	0.041	-0.053	-0.004	0.000	0.009
16. Missingness on socio- demographics	0.000	-0.019	-0.064	0.087*	-0.038	-0.003	-0.034	0.102**

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-2 GSES 2000-02 Student-Level NGS-GSES Model Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	15	16
1. Cluster 1: Global Experience - Very Satisfied								
2. Cluster 2: Global Experience - Satisfied								
3. Cluster 3: Global Experience - Very Dissatisfied								
4. Cluster 4: Global Experience - No Answer								
5. Reported loans at graduation								
6. Reported unemployment at convocation								
7. Deviation from avrg program duration (mths)								
8. Student with visa status								
9. Women alumni	1.00							
10. Visible minority status	-0.086*	1.00						
11. Master's of accounting alumni	0.075*	0.186***	1.00					
12. Doctoral alumni	-0.060	-0.059	-0.187***	1.00				
13. Science degree	-0.215***	0.061	-0.468***	0.114***	1.00			
14. Part- time studies (proxy for no involvement)	-0.009	-0.027	-0.069*	-0.097**	0.032	1.00		
15. Missingness on program characteristics	0.022	-0.001	0.000	0.000	0.000	0.000	1.00	
16. Missingness on socio- demographics	-0.050	0.000	-0.057	0.000	0.074*	0.003	-0.020	1.00

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=812; Missing cases mean imputation

Table E-3 NGS Student-Level NGS-GSES Model Correlation Statistics

Predictors	1	2	3	4	5	6	7	8
1. Cluster 1: Global Experience - Very Satisfied	1.00							
2. Cluster 2: Global Experience - Satisfied	-0.616***	1.00						
3. Cluster 3: Global Experience - Very Dissatisfied	-0.161***	-0.451***	1.00					
4. Cluster 4: Global Experience - No Answer	-0.139***	-0.391***	-0.102***	1.00				
5. Reported loans at graduation	0.001	-0.021	0.021	0.012	1.00			
6. Reported unemployment at convocation	0.044***	-0.067***	0.023*	0.030**	0.427***	1.00		
7. Deviation from avrg program duration (mths)	-0.070***	0.079***	-0.010	-0.028**	-0.111***	-0.233***	1.00	
8. Student with visa status	0.038***	-0.071***	0.047***	0.019	0.024*	0.040***	-0.039***	1.00
9. Women alumni	-0.078***	0.031**	0.024*	0.027*	-0.009	-0.046***	0.091***	-0.059***
10. Visible minority status	0.054***	-0.057***	0.004	0.019	-0.021*	0.104***	-0.089***	0.204***
11. Master's of accounting alumni	-0.047***	0.060***	-0.007	-0.030**	-0.032**	-0.089***	0.279***	-0.043***
12. Doctoral alumni	0.056***	-0.064***	-0.006	0.041***	-0.089***	0.042***	-0.132***	0.021
13. Science degree	0.115***	-0.068***	-0.037***	0.000	-0.026*	0.122***	-0.053***	0.105***
14. Part- time studies (proxy for no involvement)	-0.086***	0.106***	-0.024*	-0.039***	-0.241***	-0.334***	0.360***	-0.087***
15. Missingness on program characteristics	-0.054***	0.011	0.082***	-0.036***	-0.049***	0.025*	-0.008	0.023*
16. Missingness on socio- demographics	0.005	-0.021*	0.011	0.018	0.038***	0.037***	-0.049***	0.000

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=8667; Missing cases mean imputation

Table E-3 NGS Student-Level NGS-GSES Model Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	15	16
1. Cluster 1: Global Experience - Very Satisfied								
2. Cluster 2: Global Experience - Satisfied								
3. Cluster 3: Global Experience - Very Dissatisfied								
4. Cluster 4: Global Experience - No Answer								
5. Reported loans at graduation								
6. Reported unemployment at convocation								
7. Deviation from avrg program duration (mths)								
8. Student with visa status								
9. Women alumni	1.00							
10. Visible minority status	-0.116***	1.00						
11. Master's of accounting alumni	0.099***	-0.073***	1.00					
12. Doctoral alumni	-0.129***	0.109***	-0.120***	1.00				
13. Science degree	-0.161***	0.204***	-0.086***	0.186***	1.00			
14. Part- time studies (proxy for no involvement)	0.083***	-0.142***	0.224***	-0.134***	-0.205***	1.00		
15. Missingness on program characteristics	-0.045***	-0.006	-0.035***	0.076***	-0.051***	0.057***	1.00	
16. Missingness on socio- demographics	0.000	0.025*	-0.046***	0.015	0.025*	-0.070***	-0.083***	1.00

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=8667; Missing cases mean imputation

Table E-4 GSES 2000-02 Department-Level Contextual Model Correlation Statistics

Predictors	1	2	3	4	5	6
1. Alumni pool (proxy for dept. size)	1.00					
2. Departmental response rate	-0.075	1.00				
3. Alumni to faculty ratio (graduate and undergraduate combined)	0.595***	-0.069	1.00			
4. Graduate to undergraduate ratio	0.099	0.000	-0.306	1.00		
5. Average departmental budget (averaged 2000- 02, $K)	0.417*	0.019	-0.149	0.027	1.00	
6. Average grade point average	-0.238	0.139	-0.168	-0.008	-0.045	1.00
7. Average deviation from normal program duration (mths)	-0.115	0.452**	0.105	-0.030	-0.144	0.161
8. Average alumni age	-0.301	0.013	-0.154	0.005	-0.139	0.044
9. Proportion of master's of accounting alumni	0.814***	-0.051	0.625***	-0.013	-0.010	-0.260
10. Proportion of doctoral alumni	-0.018	0.125	-0.224	0.043	0.477**	0.394*
11. Proportion of women alumni	-0.190	0.261	-0.016	-0.014	-0.447**	0.201
12. Proportion of alumni with visa status	-0.069	-0.155	-0.241	0.594***	0.026	-0.126

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=36

Table E-4 GSES 2000-02 Department-Level Contextual Model Correlation Statistics (Cont'd)

Predictors	7	8	9	10	11	12
1. Alumni pool (proxy for dept. size)						
2. Departmental response rate						
3. Alumni to faculty ratio (graduate and undergraduate combined)						
4. Graduate to undergraduate ratio						
5. Average departmental budget (averaged 2000-02, $K)						
6. Average grade point average						
7. Average deviation from normal program duration (mths)	1.00					
8. Average alumni age	0.225	1.00				
9. Proportion of master's of accounting alumni	-0.151	-0.409*	1.00			
10. Proportion of doctoral alumni	-0.063	0.059	-0.213	1.00		
11. Proportion of women alumni	0.096	0.120	0.029	-0.116	1.00	
12. Proportion of alumni with visa status	-0.247	-0.238	-0.132	-0.165	-0.063	1.00

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded listwise, N=36

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics

Predictors	1	2	3	4	5	6	7	8
1. Cluster 1: Supervisor Experience - Very Satisfied	1.00							
2. Program completion longer than expected	-0.051	1.00						
3. Not satisfied with superv. on authorship issues	-0.229*	-0.145	1.00					
4. Financial support distribution not fair/equitable	-0.231*	0.043	0.039	1.00				
5. Resources not adequately available for TA duties	-0.321*	-0.035	0.238	0.240	1.00			
6. University expectations before enrolment	-0.004	0.041	-0.047	-0.158	-0.147	1.00		
7. Coop program expectations before enrolment	0.013	-0.068	0.031	0.087	0.309*	-0.714***	1.00	
8. Research expert. expectations before enrolment	0.154	0.055	-0.069	0.006	-0.140	-0.401***	-0.104	1.00
9. Conference expenses partially/not funded	0.137	-0.157	-0.161	-0.073	-0.048	-0.018	-0.015	0.105
10. Reported loans at graduation	-0.058	0.149	0.100	0.177	0.070	0.002	0.093	-0.076
11. Reported unemployment at convocation	0.006	0.284**	-0.040	0.216	0.092	-0.131	0.142	-0.106
12. Deviation from avrg program duration (mths)	-0.045	0.121	-0.077	0.142	-0.126	0.037	-0.042	-0.181
13. Student with visa status	0.170	0.072	-0.015	-0.045	0.004	-0.151	0.187	0.001
14. Not an UW alumni before enrolment	-0.012	-0.112	0.023	-0.180	-0.018	0.143	-0.076	-0.077
15. Women alumni	-0.090	-0.202*	0.171	-0.058	0.290*	-0.119	0.168	0.023
16. Visible minority status	-0.044	0.108	-0.181	0.057	-0.151	0.032	-0.022	-0.150
17. Married during degree program	0.072	0.101	0.030	-0.179	-0.105	0.141	-0.096	-0.117

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics (Cont'd)

Predictors	1	2	3	4	5	6	7	8
18. With dependents during degree program	0.099	0.022	-0.026	0.017	-0.026	0.104	-0.041	-0.029
19. Doctoral alumni	0.207*	0.047	0.099	-0.074	-0.177	0.054	-0.126	0.099
20. Science degree	0.028	0.168	-0.154	-0.064	-0.392**	0.152	-0.184	-0.004
21. Financial support (K $)	-0.014	-0.007	-0.039	0.160	-0.082	0.128	-0.111	-0.043
22. Frequent visits to the Graduate House	0.022	-0.062	-0.060	0.027	0.084	-0.070	0.039	0.118
23. Some research or teaching involvement	0.379***	0.135	-0.198	-0.198	-0.215	0.066	-0.067	0.113
24. No research or teaching involvement	-0.279**	-0.072	0.373***	0.095	0.290*	0.003	0.002	-0.005
25. Missingness on unemployment	0.229*	-0.103	-0.117	-0.055	-0.196	-0.061	-0.025	0.137
26. Missingness on program characteristics	-0.245**	-0.101	-0.023	0.317**	0.306*	0.081	-0.087	-0.145

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	15	16
1. Cluster 1: Supervisor Experience - Very Satisfied								
2. Program completion longer than expected								
3. Not satisfied with superv. on authorship issues								
4. Financial support distribution not fair/equitable								
5. Resources not adequately available for TA duties								
6. University expectations before enrolment								
7. Coop program expectations before enrolment								
8. Research expert. expectations before enrolment								
9. Conference expenses partially/not funded	1.00							
10. Reported loans at graduation	-0.023	1.00						
11. Reported unemployment at convocation	-0.057	0.020	1.00					
12. Deviation from avrg program duration (mths)	0.058	-0.201*	0.079	1.00				
13. Student with visa status	-0.127	-0.201*	-0.003	-0.032	1.00			
14. Not an UW alumni before enrolment	-0.122	-0.136	-0.034	-0.076	0.186	1.00		
15. Women alumni	0.137	0.023	-0.026	-0.139	-0.040	-0.017	1.00	
16. Visible minority status	0.149	-0.184	0.249*	0.083	0.177	0.107	-0.145	1.00
17. Married during degree program	-0.081	-0.027	-0.129	0.046	0.040	0.167	-0.088	0.035

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics (Cont'd)

Predictors	9	10	11	12	13	14	15	16
18. With dependents during degree program	0.149	0.180	-0.058	0.065	-0.077	0.108	-0.060	0.045
19. Doctoral alumni	-0.170	-0.098	-0.235*	-0.145	0.152	0.067	0.001	-0.106
20. Science degree	-0.467**	-0.314***	-0.047	0.188	0.142	-0.016	-0.293**	0.198
21. Financial support (K $)	-0.106	-0.151	-0.124	0.311**	0.044	-0.012	-0.116	0.157
22. Frequent visits to the Graduate House	0.126	0.242*	-0.160	-0.098	0.079	-0.130	0.052	-0.277**
23. Some research or teaching involvement	0.245	0.219*	0.026	-0.184	0.131	-0.062	-0.095	-0.096
24. No research or teaching involvement	-0.245	-0.048	-0.121	0.070	-0.127	-0.121	-0.177	-0.011
25. Missingness on unemployment	0.132	0.076	#VALUE!	-0.067	-0.024	-0.063	0.036	-0.165
26. Missingness on program characteristics	-0.118	-0.316***	0.008	0.319***	-0.078	0.178	0.035	0.240*

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics (Cont'd)

Predictors	17	18	19	20	21	22	23	24
18. With dependents during degree program	0.499***	1.00						
19. Doctoral alumni	0.212*	0.142	1.00					
20. Science degree	0.129	-0.098	-0.007	1.00				
21. Financial support (K $)	0.146	0.003	0.404***	0.204	1.00			
22. Frequent visits to the Graduate House	-0.204*	-0.128	0.041	-0.197*	-0.023	1.00		
23. Some research or teaching involvement	0.112	0.048	0.281**	0.126	0.026	0.127	1.00	
24. No research or teaching involvement	-0.032	0.055	-0.152	-0.097	-0.121	-0.103	-0.569***	1.00
25. Missingness on unemployment	0.078	0.067	-0.047	0.096	-0.129	0.033	0.268**	-0.184
26. Missingness on program characteristics	0.076	0.103	-0.228*	-0.009	0.246*	-0.311***	-0.618***	0.325***

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Table E-5 GSES 2002 Factorial Student-Level Correlation Statistics (Cont'd)

Predictors	25	26						
18. With dependents during degree program								
19. Doctoral alumni								
20. Science degree								
21. Financial support (K $)								
22. Frequent visits to the Graduate House								
23. Some research or teaching involvement								
24. No research or teaching involvement								
25. Missingness on unemployment	1.00							
26. Missingness on program characteristics	-0.241*	1.00						

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$
Cases excluded pairwise, N Max = 109

Appendix F

Power Calculations for Intercept-Only Models

Power calculations are based on the approximation

$ES/_{s.e.} \approx \left(Z_{1-\alpha} + Z_{1-\beta}\right)$ see (Hox, 2002: 177-179)

where

ES is the effect size, the anticipated between-group variance (Tau)

S.E. is the square root of the anticipated sampling variance of the between-group variance

$Z_{1-\alpha}$ is the probability of accepting the null hypothesis without Type I error (1 - critical region)

$Z_{1-\beta}$ is the probability of rejecting the null hypothesis without Type II error (power)

The sampling variance (s.e.) for the intercept-only model is equal to

$$\text{var}(\sigma_u^2) = \frac{2\sigma_e^4}{kn_{clus}}\left(\frac{1}{n_{clus}-1} + 2\omega + n_{clus}\omega^2\right)$$ see (Hox, 2002: 184-196)

Where

$\omega = \sigma_u^2/\sigma_e^2$ is the ratio of between-group variance (Tau) divided by within-group variance (Sigma)

k is the number of clusters

n_{clus} is the average cluster size

The effect size (Tau) is based on one's anticipated size of the between-group variance, or may be drawn from the anticipated intraclass correlation. The later can be computed as $\rho = \sigma_u^2/\left(\sigma_u^2 + \sigma_e^2\right)$; for a logistic model, the within-group variance is fixed to $\sigma_e^2 = \pi^2/_3 \approx 3.29$ see (Hox, 2002: 117)

Bibliography

AAPOR. (2006). *Standard Definitions: Final Dispositions of Case and Outcome Rates for Surveys*. Retrieved May 2006, from http://www.aapor.org/pdfs/satndarddefs_4.pdf.

Allen, H. (1998). Anticipating Market Demand: Tracking Enrollee Satisfaction and Health Over Time. *International Journal for Quality in Health Care, 10*(6), 521-530.

Allison, P. (1999). Comparing Logit and Probit Coefficients Across Groups. *Sociological Methods and Research, 28*(2), 186-208.

Ancis, J., Sedlacek, W., & Mohr, J. (2000). Student Perceptions of Campus Cultural Climate by Race. *Journal of Counseling and Development, 78*(2), 180-185.

Andres, L., & Carpenter, S. (1997). *Today's Higher Education Students: Issues of Admission, Retention, Transfer and Attrition in Relation to Changing Demographics*: British Columbia Council on Admissions and Transfer.

Astin, A. (1999). Student Involvement: A Developmental Theory for Higher Education. *Journal of College Student Development, 40*(5), 518-529.

Atrostic, B. K., Bates, N., Burt, G., & Silberstein, A. (2001). Nonresponse in U.S. Government Household Surveys: Consistent Measures, Recent Trends, and New Insights. *Journal of Official Statistics, 17*(2), 209-226.

Babbie, E. (1998). *The Practice of Social Research* (Eight ed.). Belmont: Wadsworth Publishing Company.

Babin, B., & Griffin, M. (1998). The Nature of Satisfaction: An Updated Examination and Analysis. *Journal of Business Research, 41*, 127-136.

Baker, S. (2000). Analyzing a Randomized Cancer Prevention Trial with a Missing Binary Outcome, an Auxiliary Variable, and All-or-None Compliance. *Journal of the American Statistical Association, 95*, 43-50.

Barber, B. (1983). *The Logic and Limits of Trust*. New Brunswick: Rutgers University Press.

Barsky, J., & Huxley, S. (1992). A Customer-Survey Tool: Using the "Quality Sample". *Cornell Hotel and Restaurant Administration Quarterly, 33*(6), 18-25.

Baruch, Y. (1999). Response Rate in Academic Studies -- A Comparative Analysis. *Human Relations, 52*(4), 421-438.

Befu, H. (1980). Structural and Motivational Approaches to Social Exchange. In K. Gergen, M. Greenberg & R. Willis (Eds.), *Social Exchange: Advances in Theory and Research* (pp. 197-214). New York: Plenum Press.

Benjamin, M. (1994). The Quality of Student Life: Toward a Coherent Conceptualization. *Social Indicators Research, 31*, 205-264.

Benson, L. (1946). Mail Surveys Can Be Valuable. *Public Opinion Quarterly, 10*(2), 234-241.

Berinsky, A. J. (2004). *Silent Voices: Public Opinion and Political Participation in America*. Princeton: Princeton University Press.

Billings-Gagliardi, S., Barrett, S., & Mazor, K. (2004). Interpreting Course Evaluation Results: Insights From Thinkaloud Interviews with Medical Students. *Medical Education, 38*, 1061-1070.

Blaikie, N. W. (1991). A Critique of the USe of Triangulation in Social Research. *Quality and Quantity, 25*, 115-136.

Blalock, H. M. (1984). Contextual Effects Models: Theoretical and Methodological Issues. *Annual Review of Sociology, 10*, 353-372.

Blasius, J., & Thiessen, V. (2001). Methodological Artefacts in Measures of Political Efficacy and Trust: A Multiple Correspondence Analysis [Electronic Version]. *Political Analysis, 9*. Retrieved May 2006 from http://polmeth.wustl.edu/polanalysis/vol/9/pas91-1-20.pdf.

Blau, P. M. (1967). *Exchange and Power in Social Life*. New York: John Wiley & Sons.

Bone, A. (2002). Pursuing a Master's Degree: Opportunity Cost and Benefit. *Education Quarterly Review, 8*(4), 16-42.

Brehm, J. (1993). The Phantom Respondents. Opinion Surveys and Political Representation. Ann Arbor: University of Michigan Press.

Brenders, D. A., Hope, P., & Ninnan, A. (1999). A Systemic Student-Centered Study of University Service. *Research in Higher Education, 40*(6), 665-685.

Campanelli, P. C., Martin, E., & Rothgeb, J. M. (1991). The Use of Respondent and Interviewer Debriefing Studies as a Way to Study Response Error in Survey Data. *The Statistician, 40*(3), 253-264.

Carr-Hill, R. A. (1992). The Measurement of Patient Satisfaction. *Journal of Public Health Medicine, 14*(3), 236-249.

Chang, M. (1996). *Racial Diversity in Higher Education: Does a Racially Mixed Population Affect Educational Outcomes?* , University of California, Los Angeles.

Cheng, Y., & Tam, W. (1997). Multi-models of Quality in Education. *Quality Assurance in Education, 5*(1), 22-31.

Clark, S., & Boser, J. (1995). *Reexamining Traditional Issues in Survey Research: Just How Evil is the Anathema of Low Response Rate?* Paper presented at the American Education Research Association.

Cochran, W. (1953). *Sampling Techniques*. New York: Wiley.

Cohen, G., & Duffy, J. C. (2002). Are Nonrespondents to health Surveys Less Healthy than Respondents. *Journal of Official Statistics, 18*(1), 13-23.

Cohen, J., & Cohen, P. (1983). *Applied Rergression/Correlation Analysis for the Behavioral Sciences* (Second ed.). Hillsdale: Lawrence Erlbaum Associates.

Collins, D. (2003). Pretesting Survey Instruments: An Overview of Cognitive Methods. *Quality of Life Research, 12*, 229-238.

Connelly, N., Brown, T., & Decker, D. (2003). Factors Affecting Response Rates to Natural Resource - Focused Mail Surveys: Empirical Evidence of Declining Rates Over Time. *Society and Natural Resources, 16*, 541-549.

Conrad, F., & Blair, J. (2001). *Interpreting Verbal Reports in Cognitive Interviews.* Paper presented at the Proceedings of the American Statistical Association.

Conrad, F., Blair, J., & Tracy, E. (1999). *Verbal Reports are Data! A theoretical Approach to Cognitive Interviews.* Paper presented at the Federal Committee on Statistical Methodology.

Converse, P. E. (1970). Attitudes and Non-Attitudes: Continuation of a Dialogue. In E. TUFTE (Ed.), *The Quantitative Analysis of Social Problems* (pp. 168-189). California: Addison-Wesley Publishing.

Cotten, S., & Wilson, B. (2006). Student-Faculty Interactions: Dynamics and Determinants. *Higher Education, 51*, 487-519.

Coyle-Shapiro, J., Kessler, I., & Purcell, J. (2004). Exploring Organizationally Directed Citizenship Behaviour: Reciprocity or 'It's My Job'? *Journal of Management Studies, 41*(1), 85-106.

Cronin, J., & Taylor, S. (1992). Measuring Service Quality: A Reexamination and Extension. *Journal of Marketing, 56*(3), 55-68.

Cronin, J., & Taylor, S. (1994). SERVPERF versus SERVQUAL: Reconciling Performance-Based and Perceptions-Minus-Expectations Measurement of Service Quality. *Journal of Marketing, 58*(1), 125-131.

Cropanzano, R., & Mitchell, M. S. (2005). Social Exchange Theory: An Interdisciplinary Review. *Journal of Management, 31*(6), 874-900.

Cummings, S. M., Savitz, L., & Konrad, T. (2001). Reported Response Rates to Mailed Physician Questionnaires. *Health Services Research, 35*(6), 1347-1355.

Curtin, R., Presser, S., & Singer, E. (2000). The Effects of Response Rate Changes on the Index of Consumer Sentiment. *Public Opinion Quarterly, 64*(4), 413-428.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in Telephone Survey Nonresponse Over the Past Quarter Century. *Public Opinion Quarterly, 69*(1), 87-98.

De Heer, W. (1999). International Response Trends: Results of an International Survey. *Journal of Official Statistics, 15*(2), 129-142.

De Leeuw, E. D. (1999). How Do Successful and Less Successful Interviewers Differ in Tactics for Combating Survey Nonresponse. *Bulletin de Methodologie Sociologique, 62*, 29-43.

De Leeuw, E. D., Mellenbergh, G. J., & Hox, J. J. (1996). The Influence of Data Collection Method on Structural Models. A Comparison of a Mail, a Telephone, and a Face-to-Face Survey. *Sociological Methods & Research, 24*(4), 443-472.

De Vaus, D. A. (2002). *Social surveys* London: Thousand Oaks, Calif. : SAGE.

Delucchi, M. (2003). *Student Satisfaction with Higher Education During the 1970s – A decade of Social Change*. Lewiston: Mellen Studies in Education.

Deming, E. (1944). On Errors in Surveys. *American Sociological Review, 9*(4), 359-369.

Dey, E. L. (1997). Working with Low Response Rates: The Efficacy of Weighting Adjustments. *Research in Higher Education, 38*(2), 215-227.

Dillman, D. A. (2000). *Mail and Internet Surveys*. New York: John Wiley & Sons.

Dinsdale, G., & Marson, B. (1999). *Citizen/Client Surveys: Dispelling Myths and Redrawing Maps*. Retrieved. from.

Donald, M. (1960). Implications of Nonresponse for the Interpretation of Mail Questionnaire Data. *Public Opinion Quarterly, 24*(1), 99-114.

Dumaresq, C., & Lambert-Maherly, A. (2002). *Graduate Student Exit Survey 2000-01*. Ubc: Planning and Institutional Research.

Duncan, O. (1968). Inheritance of Race or Inheritance of Poverty? In D. Moynihan (Ed.), *On Understanding Poverty* (pp. 85-110). New York: Basic Books.

Durand, C., Gagnon, M.-E., Doucet, C., & Lacourse, E. (2006). An Inquiry into the Efficacy of a Contemporary Training Session for Telephone Survey Interviewers. *Bulletin de Methodologie Sociologique, 92*, 5-27.

Dyke, T., Kappelman, L., & Prybutok, V. (1997). Measuring Information Systems Service Quality: Concerns on the Use of the SERVQUAL Questionnaire. *MIS Quarterly, 21*(2), 195-208.

Einarson, M. K., & Matier, M. W. (2005). Exploring Race Differences in Correlates of Seniors' Satisfaction with Undergraduate Education. *Research in Higher Education, 46*(6), 641-676.

Elliott, K. M., & Healy, M. (2001). Key Factors Influencing Student Satisfaction Related to Recruitment and Retention. *Journal of Marketing for Higher Education, 10*(4), 1-12.

Ericson, A. (2003). Valid and Non-Reactive Verbalization of Thoughts During Performance of Tasks. Towards a Solution to the Central Problems of Introspection as a Source of Scientific Data. *Journal of Consciousness Studies, 10*(9-10), 1-18.

Ericson, A., & Simon, H. (1980). Verbal Reports as Data. *Psychological Review, 8*, 215-251.

Ericson, A., & Simon, H. (1993). *Protocol analysis: verbal reports as data* Cambridge: MIT Press.

Etter, J.-F., Perneger, T. V., & Rougemont, A. (1996). Does Sponsorship Matter in Patient Satisfaction Surveys? A Randomized Trial. *Medical Care, 34*(4), 327-335.

Etzioni, A. (1958). Administration and the Consumer. *Administrative Science Quarterly, 3*(2).

Fathi, D., Schooler, J., & Loftus, E. (1984). *Moving Survey Problems into the Cognitive Psychology Laboratory.* Paper presented at the Proceedings of the American Statistical Association.

Foddy, W. (1993). *Constructing Questions for Interviews and Questionnaires*. Cambridge: Cambridge University Press.

Folz, D. (1996). *Survey Research for Public Administration*. London: Sage Publications.

Forsyth, B., Levin, K., & Fisher, S. (1999). *Test of an Appraisal Method for Establishment Survey Questionnaires.* Paper presented at the Presented at the Joint Statistical Meetings of the American Statistical Association, Baltimore, MD.

Frankel, M. R., & Frankel, L. R. (1987). Fifty Years of Survey Sampling in the United States. *Public Opinion Quarterly, 51*(Part 2: Supplement 50th Anniversary Issue), S127-S138.

Gallagher, P. M., Fowler, F. J., & Stringfellow, V. L. (2005). The Nature of Nonresponse in Medicaid Survey: Causes and Consequences. *Journal of Official Statistics, 21*(1), 73-87.

Gallup, G. (1951). The Gallup Poll and the 1950 Election. *Public Opinion Quarterly, 15*(1), 16-22.

Gallup , G. (1953). The Future Direction of Election Polling. *Public Opinion Quarterly, 17*(2), 202-207.

Gallup , G. (1957). The Changing Climate for Public Opinion Research. *Public Opinion Quarterly, 21*(1), 23-57.

Gallup , G. (1965). Polls and the Political Process-Past, Present, and Future. *Public Opinion Quarterly, 29*(4), 544-549.

Gasquet, I., Falissard, B., & Ravaud, P. (2001). Impact of Reminders and Method of Questionnaire Distribution on Patient Response to Mail-Back Satisfaction Survey. *Journal of Clinical Epidemiology, 54*, 1174-1180.

Goudy, W. J. (1976). Nonresponse Effects on Relationships Between Variables. *Public Opinion Quarterly, 40*(3), 360-369.

Goudy, W. J. (1978). Interim Response to a Mail Questionnaire: Impacts on a Variable Relationships. *Sociological Quarterly, 19*, 253-265.

Goyder, J. (1981). Income Differences between the Sexes: Findings from a National Canadian Survey. *Canadian Review of Sociology and Anthropology, 18*(3).

Goyder, J. (1987). *The Silent Minority. Nonrespondents on Sample Surveys*. Boulder: Westview Press.

Goyder, J., Boyer, L., & Martinelli, G. (2006). Integrating Exchange and Heuristic Theories of Survey Nonresponse. *Bulletin de Methodologie Sociologique, 92*, 28-44.

Grayson, J., & Myles, R. (2005). How Research Ethics Boards are Undermining Research on Canadian University Students. *Journal of Academic Ethics, 2*(4), 293-314.

Grayson, P. (1999). *The Student Experience at York University: The Effects of Income, Race, and Gender Over Four Years*. Toronto: Institute for Social Research, York University.

Green, I. (1991). *Nonresponse Bias in a Patient Satisfaction Survey.* Claremont Graduate School, Claremont.

Green, J., Caracelli, V., & Graham, W. (1989). Toward a Conceptual Framework for Mixed-Method Evaluation Designs. *Educational Evaluation and Policy Analysis, 11*(3), 255-274.

Grosset, J. (1994). *The Biasing Effects of Nonresponse on Information Gathered by Mail Surveys. Institutional Research Report*. Philadelphia: Philadelphia Community College. Office of Institutional Research.

Groves, R. (1987). Research on Survey Data Quality. *Public Opinion Quarterly, 51*(Part 2: Supplement" 50th Anniversary), S156-S172.

Groves, R. (1989). *Survey Errors and Survey Costs*. New York: John Wiley and Sons.

Groves, R. (1991). *Towards a Useful Theory of Survey Participation*. from www.amstat.org/sections/SRMS/Proceedings/papers/1991_013.pdf

Groves, R. (2006a). Experiments in Producing Nonresponse Bias. *Public Opinion Quarterly, 70*(5 (Special Issue)), 720-736.

Groves, R. (2006b). Nonresponse Rates and Nonresponse Bias in Household Surveys. *Public Opinion Quarterly, 70*(5), 646-675.

Groves, R., Cialdini, R., & Couper, M. (1992). Understanding the Decision to Participate in a Survey. *Public Opinion Quarterly, 56*, 475-495.

Groves, R., & Couper, M. (1998). *Nonresponse in Households Interview Surveys*. New York: John Wiley and Sons.

Groves, R., & Couper, M. (2002). Designing Surveys Acknowledging Nonresponse. In M. Ver Ploeg, R. A. Moffitt & C. F. Citro (Eds.), *Studies of Welfare Populations: Data Collection and Research Issues*: National Research Council.

Groves, R., Fowler, F., Couper, M., Lepkowski, J., Singer, E., & Tourangeau, R. (2004). *Survey Methodology*. Hoboken: John Wiley & Sons.

Groves, R., Presser, S., & Dipko, S. (2004). The Role of Topic Interest in Survey Participation Decisions. *Public Opinion Quarterly, 68*(1), 2-31.

Groves, R., Singer, E., & Corning, A. (2000). Leverage-Saliency Theory of Survey Participation: Description and Illustration. *Public Opinion Quarterly, 64*(299-308).

Gruber, J. (1987). *Controlling Bureaucracies. Dilemmas in Democratic Governance*. Berkeley: University of California Press.

Gwinner, K. P., & Stephens, N. (1998). Why Don't Some People Complain? A Cognitive-Emotive Process Model of Consumer Complaint Behavior. *Journal of the Academy of Marketing Science, 26*(3), 172-189.

Healey, J. F. (1996). *Statistics: A Tool for Social Research*. Belmont: Wadsworth Publishing Company.

Heberlein, T., & Baumgartner, R. (1978). Factors Affecting Response Rates to Mailed Questionnaires: A Quantitative Analysis of the Published Literature. *American Sociological Review, 43*, 443-462.

Helgeson, J. G. (1994). Receiving and Responding to a Mail Survey: a Phenomenological Examination. *Journal of the Market Research Society, 36*(4), 339-346.

Herzberg, F. (1966). *Work and the Nature of Man*. Cleveland: The World Publishing Company.

Heskett, J., Sasser, E. J., & Schlesinger, L. A. (1997). *The Service Profit Chain*. New York: The Free Press.

Hirdes, J. P., Zimmerman, D. R., Hallman, K. G., & Soucie, P. S. (1998). Use of the MDS Quality Indicators to Assess Quality of Care in Institutional Settings. *Canadian Journal of Quality in Health Care, 14*(2), 5-11.

Hirschman, A. O. (1970). *Exit, Voice and Loyalty*. Cambridge: Harvard University Press.

Hodgson, G. (1986). Behind Methodological Individualism. *Cambridge Journal of Economics, 10*, 211-224.

Hoffmann, D., & Gavin, M. (1998). Centering Decisions in Hierarchical Linear Models: Implications for Research in Organizations. *Journal of Management, 24*(5), 623-641.

Hoffmann, D. A. (1997). An Overview of the Logic and Rationale of Hierarchical Linear Models. *Journal of Management, 23*(6), 723-744.

Howe, K. R. (2006). Against the Quantitative-Qualitative Incompatibility Thesis or Dogmas Die Hard. *Educational Researcher, 17*(8), 10-16.

Hox, J. J. (2002). *Multilevel Analysis. Techniques and Applications*. London: Lawrence Erlbaum Associates.

Hu, S., & Kuh, G. (2001). *Being Disengaged in Educationally Purposeful Activities: The Influences of Student and Institutional Characteristics*. Paper presented at the American Educational Research Association.

Hunt, K. H. (1977). CS/D Overview and Future Directions. In K. H (Ed.), *Conceptualization and Measurement of Consumer Satisfaction and Dissatisfaction*. Cambridge: Marketing Science Institute.

Iverson, G. R. (1991). *Contextual Analysis*. Newbury Park: Sage Publications.

Iyengar, S., Wells, R., & Schwarts, B. (2006). Doing Better but Feeling Worse. *Psychological Science, 17*(3), 143-150.

Jacobs, J. A. (1996). Gender Inequality and Higher Education. *Annual Review of Sociology, 22*, 153-185.

Jernigan, I., & Beggs, J. (2005). An Examination of Satisfaction with my Supervisor and Organizational Commitment. *Journal of Applied Psychology, 35*(10), 2171-2192.

Jobe, J. B. (2003). Cognitive Psychology and Self-Reports: Models and Methods. *Quality of Life Research, 12*, 219-227.

Johnson, T., Cho, Y. I., Campbell, R., & Holbrook, A. (2006). Using Community-Level Correlates to Evaluate Nonresponse Effects in a Telephone Survey. *Public Opinion Quarterly, 70*(5), 704-719.

Katz, E., & Danet, B. (1973). Introduction. In E. Katz & B. Danet (Eds.), *Bureaucracy and the Public* (pp. 3-27). New York: Basic Books Inc.

Keeter, S., Miller, C., Kohut, A., Groves, R., & Presser, S. (2000). Consequences of Reducing Nonresponse in a National Telephone Survey. *Public Opinion Quarterly, 64*, 125-148.

Kidwell, R., Mossholder, K., & Benett, N. (1997). Cohesiveness and Organizational Citizenship Behaviour: A Multilevel Analysis Using Work Groups and Individuals. *Journal Of Management, 23*(6), 775-793.

Klapp, E. (1986). *Overload and Boredom: Essays on the Quality of Life in the Information Society*. Westport: Greenwood Press.

Kojetin, B., Borgida, E., & Snyder, M. (1993). *Survey Topic Involvement and Nonresponse Bias.* Paper presented at the Proceedings of the Survey Research Methods Section, American Statistical Association.

Krahn, H., & Bowlby, J. W. (1997). Good Teaching and Satisfied University Graduates. *The Canadian Journal of Higher Education, 27*(2-3), 157-180.

Kramer, R. M. (1999). Trust and Distrust in Organizations: Emerging Persepctives, Enduring Questions. *Annual Review of Psychology, 50*, 569-598.

Kraut, A. (1996). An Overview of Organizational Surveys. In A. Kraut (Ed.), *Organizational Surveys. Tools for Assessment and Change* (pp. 1-14). San Francisco: Jossey-Bass Publishers.

Krosnick, J. A. (1991). Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology, 5*, 213-236.

Krushat, M., & Molnar, J. (1993). The Effect of Nonrespondents on a Self-Administered Mail Survey. *Evaluation Practices, 14*(1), 25-32.

Kunda, Z. (1999). *Social Cognition. Making Sense of People*. London: MIT press.

Larsson, B. W. (2000). Does the Method of Data Collection Affect Patients' Evaluations of Quality of Care. *International Journal of Nursing Practice, 5*, 284-291.

Lasek, R., Barkley, W., Harper, D., & Rosenthal, G. (1997). An Evaluation of the Impact of Nonresponse Bias on Patient Satisfaction Surveys. *Medical Care, 95*(6), 646-652.

Laurent, A. (1994). *L'Individualisme Méthodologique*. Paris: Presses universitaires de France.

Lemay, M., & Durand, C. (2002). The Effect of Interviewer Attitude on Survey Cooperation. *Bulletin de Methodologie Sociologique, 76*, 27-44.

Leonard, D. (1997). Gender Issues in Doctoral Studies. In N. Graves & V. Varma (Eds.), *Working for a Doctorate. A Guide for the Humanities and Social Sciences*. London: Routledge.

Lessler, J. T., & Kalsbeek, W. D. (1992). *Nonsampling Error in Surveys*. New York: Wiley-Interscience publication.

Levy, P. S. (1982). Surviving in a Predominantly White Male Institution. In S. Vartuli (Ed.), *The Ph.D. Experience. A Women's Point of View*. New York: Praeger Publishers.

Lewis, D., & Weigert, A. (1985). Trust as a Social Reality. *Social Forces, 63*(4), 967-985.

Lewis, I. (1984). *The Student Experience of Higher Education*. London: Croom Helm.

Ley, P., Bradshaw, P., Kincey, J., & Atherton, S. (1976). Increasing Patients' Satisfaction with Communications. *British Journal of social clinical Psychology, 25*, 403-413.

Lin, I.-F., & Schaeffer, N. C. (1995). Using Participants to Estimate the Impact of Nonparticipation. *Public Opinion Quarterly, 59*(2), 236-258.

Lipsey, M. (1998). Design Sensitivity: Statistical Power for Applied Experimental Research. In L. Bickman & D. Rog (Eds.), *Handbook of Applied Social Research Methods* (pp. 39-68). London: Sage Publications.

Low, G. (1999). What Respondents Do with Questionnaires: Accounting for Incongruity and Fluidity. *Applied Linguistics, 20*(4), 503-533.

Luhmann, N. (1979). *Trust and Power: Two Works by Niklas Luhmann*. Chichester: John Wiley.

Martin, C. (1994). The Impact of Topic Interest on Mail Survey Response Behaviour. *Journal of the Market Research Society, 36*(4).

McCarthy, J., & Beckler, D. (2000). *Survey Burden and its Impact on Attitudes Towards the Survey Sponsor*. Paper presented at the International Conference on Establishment Surveys II. from http://www.nass.usda.gov/Education_&_Outreach/Reseach_Reports/.

McDaniel, S., & Madden, C. (1987). Do Topic Differences Affect Survey Non-Reponse? *Journal of the Market Research Society, 29*(1), 55-66.

Menard, S. (2002). *Applied Regression Analysis* (Vol. 106). London: Sage Publications.

Merkle, D. M., & Edelman, M. (2002). Nonresponse in Exit Polls: A Comprehensive Analysis. In M. G. Robert (Ed.), *Survey Nonresponse*. New York: John Wiley & Sons.

Michalos, A. (1973). Rationality Between the Maximizers and the Satisficers'. *Policy Sciences, 4*, 229-244.

Michalos, A. (1985). Multiple Discrepancies Theory. *Social Indicators Research, 16*, 347-413.

Millem, J. (1998). Change in College Students: Examining the Effect of College Peer Groups and Faculty Normative Groups. *The Journal of Higher Education, 69*(2), 117-140.

Miller, J. (1983). *The American People and Social Policy*. New York: Pergamon Press.

Miller, T., Bender, B., & Schuh, J. (2005). *Promoting Reasonable Expectations. Aligning Student and Institutional Views of the College Experience*. San Francisco: Jose-Bass.

Mollering, G. (2001). The Nature of Trust: From Georg Simmel to a Theory of Expectation, Interpretation and Suspension. *Sociology, 35*(2), 403-420.

Moore, D. L., & Tarnai, J. (2002). Evaluating Nonresponse Error in Mail Surveys. In M. G. Robert (Ed.), *Survey Nonresponse*. New York: John Wiley & Sons.

Morgenbesser, S. (1964). *Psychologism and Methodological Individualism*: Voice of America Forum Lectures.

Mori. (2002). *Public Service Reform: Measuring & Understanding Customer Satisfaction*. Retrieved Feb 2005, from http://www.mori.com/publications/rd/opsr.shtml.

Morrison, Z., Bourke, M., & Kelley, C. (2005). 'Stop Making it Such a Big Issue': Perceptions and Experiences of Gender Inequality by Undergraduates at a British University. *Women's Studies International Forum, 28*, 150-162.

Moum, T. (1994). Is Subjective Well-Being A Predictor of Nonresponse in Broad Population Surveys. *Social Indicators Research, 32*, 1-20.

Nachmias-Frankfort, C., & Nachmias, D. (1996). *Research Methods in the Social Sciences* (5th ed.). New York: St. Martin's Press.

Nederhof, A. J. (1987). When Neutrality is Negative: Pleasantness of Most Recent Survey Experience and Nonresponse. *Quality and Quantity, 21*, 425-432.

Neuman, L. W. (2004). *Basics of Social Research. Qualitative and Quantitative Approaches*. Boston: Pearson Education Inc.

Nolin, M. J. (1996). *Use of Cognitive Laboratories and Recorded Interviews in the National Household Education Survey*: National Center for Education Statistics.

Nyer, P., & Gopinath, M. (2005). Effects of Complaining Versus Negative Word of Mouth on Subsequent Changes in Satisfaction: The Role of Public Commitment. *Psychology and Marketing, 22*(12), 937-953.

Oliver, R. (1981). Measurement and Evaluation of Satisfaction Processes in Retail Settings. *Journal of Retailing, 57*, 25-48.

Oliver, R. (1997). *Satisfaction: A Behavioral Perspective On the Consumer*. New York: McGraw-Hill Companies Inc.

Parashos, P., Morgan, M., & Messer, H. (2005). Response Rate and Nonresponse Bias in a Questionnaire Survey of Dentists. *Community Dentistry Oral Epidemiology, 33*, 9-16.

Parasuraman, A., Zeithami, V., & Berry, L. (1994). Reassessment of Expectations as a Comparison Standard in Measuring Service Quality: Implications for Further Research. *Journal of Marketing, 58*(1), 111-124.

Pearl, D., & Fairley, D. (1985). Testing for the Potential for Nonresponse Bias in Sample Surveys. *Public Opinion Quarterly, 49*(4), 553-560.

Penley, L., & Gould, S. (1988). Etzioni's Model of Organizational Involvement: A Perspective for Understanding Commitment to Organizations. *Journal of Organizational Behavior, 9*(1), 43-59.

Perneger, T., Charnot, E., & Bovier, P. (2005). Nonresponse Bias in a Survey of Patient Perceptions of Hospital Care. *Medical Care, 43*(4), 1-10.

Petty, Richard, & John T. Cacioppo. (1986). *Communication and Persuasion: Central and Peripheral Routes to Attitude Change*. New York: Springer-Verlag.

Petty, R. E., Cacioppo, J., & Schumann, D. (1983). Central and Peripheral Routes to Advertising Effectiveness: The Moderating Role of Involvement. *Journal of Consumer Research, 10*(2), 135-146.

Pike, G. (1991). The Effects of Background, Coursework, and Involvement on Students' Grades and Satisfaction. *Research in Higher Education, 32*(1), 15-31.

Planning, O. o. I. R. a. A. (1996). *Continuing Student Opinion Survey*. New Brunswick: Rutgers-New Brunswick University.

Podsakoff, P., Mackenzie, S., Paine, J. B., & Bachrach, D. (2000). Organizational Citizenship Behaviors: A Critical Review of the Literature and Suggestions for Future Research. *Journal of Management, 26*(3), 513-563.

Polkinghorne, D. (1983). *Methodology for the Human Sciences. Systems of Inquiry*. Albany: State University of New York Press.

Porter, S. (2004). Raising Response Rates: What Works. *New Directions for Institutional Research, 121*, 5-21.

Porter, S., & Umbach, P. D. (2006). Student Survey Response Rates Across Institutions: Why Do They Vary? *Research in Higher Education, 47*(2), 229-247.

Porter, S., & Whitcomb, M. (2005). Non-Response in Student Surveys: The Role of Demographics, Engagement and Personality. *Research in Higher Education, 46*(2), 127-152.

Presser, S., Couper, M., Lessler, J. T., Martin, E., Rothgeb, J. M., & Singer, E. (2004). Methods for Testing and Evaluating Survey Questions. *Public Opinion Quarterly, 68*(1), 109-130.

Pryor, J. (2004). Conducting Surveys on Sensitive Topics. In S. Porter (Ed.), *Overcoming Survey Research Problems* (pp. 39-50). San Francisco: Jossey-Bass.

Puyo-Savary, F. (2004). *Réflexions sur les determinants socio-économiques de la protection Natura 20000 en France*. Retrieved May 2006, from http://crereg.eco.univ-rennes1.fr/.

Raudenbush, S. (1989). "Centering" Predictors in Multilevel Analysis: Choices and Consequences. *Multilevel Modelling Newsletter, 1*(2), 10-12.

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical Linear Modelling. Applications and Data Analysis Methods*. London: Sage Publications.

Raudenbush, S. W., & Liu, X. (2000). Statistical Power and Optimal Design for Multisite Randomized Trials. *Psychological Methods, 5*(2), 199-213.

Ray, J. J., & Still, L. V. (1987). Maximizing the Response Rate in Surveys May Be a Mistake. *Personality and Individual Differences, 8*(4), 571-573.

Redline, C., Smiley, R., Lee, M., & DeMaio, T. (1998). *Beyond Concurrent Interviews: An Evaluation of Cognitive Interviewing Techniques for Self-Administered Questionnaires.* Paper presented at the Proceedings of the American Statistical Association.

Research, B. M. (2003). *New Brunswick Community College Three Year graduate Follow-Up Survey of 1999-2000 Graduates*: Department of Training and Employment Development.

Reuss, C. (1943). Differences Between Persons Responding and Not Responding to a Mailed Questionnaire. *American Sociological Review, 8*, 433-438.

Robertson, M. T., & Sundstrom, E. (1990). Questionnaire Design, Return Rates and Response Favorableness in an Employee Attitude Questionnaire. *Journal of Applied Psychology, 75*(3), 354-357.

Rodeghier, M. (1996). *Surveys with Confidence*. Chicago: SPSS.

Rogelberg, S. G., Luong, A., Sederburg, M., & Cristol, D. (2000). Employee Attitude Surveys: Examining the Attitudes of Noncompliant Employees. *Journal of Applied Psychology, 85*(2), 284-293.

Rogelberg, S. G., Sederburg, M., Aziz, S., Conway, J., Spitzmuller, C., & Knight, W. (2003). Profiling Active and Passive Nonrespondents to an Organizational Survey. *Journal of Applied Psychology, 85*(6), 1104-1114.

Roose, H., Waege, H., & Agneessens. (2003). Respondent Related Correlates of Response Behaviour in Audience Research. *Quality and Quantity, 37*, 411-434.

Roseneau, J. (1974). *Citizenship Between Elections. An Inquiry Into the Mobilizable American*. New York: The Free Press.

Rosenthal, R., & Rosnow, R. L. (1975). *The Volunteer Subject*. New York: John Wiley & Sons.

Rossi, P. H., Wright, J. D., & Anderson, A. B. (1983). *Handbook of Survey Research*. New York: Academic Press.

Rubin, D. B. (1987). *Multiple Imputation for Nonresponse in Surveys*. New York: John Wiley & Sons.

Schiltz, M. (1988). Professional Standards for Survey Research. *Research in Higher Education, 28*(1), 67-75.

Schultz, W. (1995). Multiple-Discrepancies Theory Verus Resource Theory. *Social Indicators Research, 34*, 153-169.

Schuman, H., & Presser, S. (1981). *Questions and Answers in Attitude Surveys*. New york: Academic Press.

Schwarts, B., Ward, A., Monterosso, J., Lyubomirsky, S., White, K., & Lehman, D. (2002). Maximizing Versus Satisficing: Happiness Is a Matter of Choice. *Journal of Personality and Social Psychology, 83*(5), 1178-1197.

Senf, J. (1987). The Option to Refuse: A Tool in Understanding Nonresponse in Mailed Surveys. *Evaluation Review, 11*(6), 775-781.

Silva, A. d. (2000). *A Framework for Measuring Responsiveness. GPE Discussion Paper Series: No. 32. EIP discussion papers*. Retrieved Feb 2005, from http://www.who.int/responsiveness/papers/en/.

Simonoff, J. (2003). *Analyzing Categorical Data*. New York: Springer.

Singer, E., Mathiowetz, N. A., & Couper, M. (1993). The Impact of Privacy and Confidentiality Concerns on Survey Participation. The Case of the 1990 U.S. Census. *Public Opinion Quarterly, 57*, 465-482.

Singh, J. (1988). Consumer Complaint Intentions and Behaviour: Definitional and Taxonomical Issues. *Journal of Marketing, 52*, 93-107.

Singh, J. (1990). A Typology of Consumer Dissatisfaction Response Styles. *Journal of Retailing, 66*, 57-98.

Smith, D. (2000). *How Will I Know If There Is Quality? Report on Quality Indicators and Quality Enhancement in Universities: Issues and Experiences*. Toronto: Council of Ontario Universities

Smith, K., & Bers, T. (1987). Improving Alumni Survey Response Rates: An Experiment and Cost-Benefit Analysis. *Research in Higher Education, 27*(3), 218-225.

Snedecor, G., & Cochran, W. (1989). *Statistical Methods* (8th ed.). Ames: The Iowa State University Press.

Spitzmuller, C., Glenn, D. M., Barr, C., Rogelberg, S. G., & Daniel, P. (2006). "If you treat me right, I reciprocate": Examining the Role of Exchange in Organizational Survey Response. *Journal of Organizational Behavior, 27,* 19-35.

SPSS. (1999). *SPSS Regression Models 9.0*. Chicago: SPSS Inc.

SPSS. (2003). Ordinal Regression. In *SPSS Advanced Models* (pp. 81-89). Chicago, Illinois: SPSS Inc.

Squire, P. (1988). Why the 1936 Literary Digest Poll Failed. *Public Opinion Quarterly, 52*(1).

Stallard, P. (1996). The Role and Use of Consumer Satisfaction Surveys in Mental Health Services. *in Journal of Mental Health, 5*(4), 333-348.

Stassen, M., & Staats, S. R. (1988). Hope and Happiness: A Comparison of Some Discrepancies. *Social Indicators Research, 20*(20), 45-58.

Statistics Canada. (1995a). *Microdata Codebook for the Survey of 1995 Graduates*. Ottawa: Statistics Canada.

Statistics Canada. (1995b). *National Graduate Survey Users Guide*.

Steeh, C. (1981). Trends in Nonresponse Rates, 1952-1979. *Public Opinion Quarterly, 45*(1), 40-57.

Steeh, C., Firgis, N., Cannon, B., & DeWitt, J. (2001). Are They Really as Bad as They Seem? Nonresponse Rates at the End of the Twentieth Century. *Journal of Official Statistics, 17*(2), 227-247.

Stephens, N., & Gwinner, K. (1998). Why Don't Some People Complain? A Cognitive-Emotive Process Model of Consumer Complaint Behavior. *Journal of the Academy of Marketing Science, 26*(3), 172-189.

Stoop, I. A. L. (2005). *The Hunt for the Last Respondent: Nonresponse in Sample Surveys*. The Hague: Social and Cultural Planning Office of the Netherlands.

Stopher, P., & Sheskin, I. (1981). A Method for Determining and Reducing Nonresponse Bias. *American Statistical Association, 54*, 252-257.

Stouffer, S., Suchman, E., DeVinney, L., Star, S., & Williams, R. (1949). The American Soldier (Vol. 1). Princeton: Princeton University Press.

Sudman, S., & Bradburn, N. (1987). The Organizational Growth of Public Opinion Research in the United States. *Public Opinion Quarterly, 51*(Part 2: Supplement: 50th Anniversary Issue), S67-S78.

Sudman, S., Bradburn, N. M., & Schwarz, N. (1996). *Thinking About Answers. The Application of Cognitive Processes to Survey Methodology*. San Francisco: Jossey-Bass Publishers.

Takahashi, N. (2000). The Emergence of Generalized Exchange. *American Journal of Sociology, 105*(4), 1105-1134.

Tashakkori, A., & Teddlie, C. (1998). *Mixed Methodology. Combining Qualitative and Quantitative Approaches*. Thousand Oak: Sage publications.

Taylor, S. E., & Fiske, S. T. (1978). Salience, Attention and Attribution: Top of Head Phenomena. *Advances in Experimental Social Psychology, 11*, 250-288.

Teas, K. (1993). Expectations, Performance Evaluation, and Consumers' Perceptions of Quality. *Journal of Marketing, 57*(4), 18-34.

Teas, K. (1994). Expectations as Comparison Standard in Measuring Service Quality: An Assessment of a Reassessment. *Journal of Marketing, 58*(1), 132-139.

Teitler, J. O., Reichman, N., & Sprachman, S. (2003). Costs and Benefits of Improving Response Rates for a Hard-To-Reach Population. *Public Opinion Quarterly, 67*, 126-137.

Thomas, E. H., & Galambos, N. (2004). What Satisfies Students? Mining Student-Opinion Data with Regression and Decision Tree Analysis. *Research in Higher Education, 45*(3), 251-269.

Thorsten, H.-T., Langer, M. F., & Hansen, U. (2001). Modeling and Managing Student Loyalty. An Approach Based on the Concept of Relationship Quality. *Journal of Service Research, 3*(4), 331-344.

Tourangeau, R. (1999). Context Effects on Answers to Attitude Questions. In M. Sirken, D. Herrman, S. Schechter, N. Schwarz, J. Tanur & R. Tourangeau (Eds.), *Cognition and Survey Research*. New York: John Wiley and Sons.

Tourangeau, R. (2000). *The Psychology of Survey Response*. New York: Cambridge University Press.

Tourangeau, R. (2004). Survey Research and Societal Change. *Annual Review of Psychology, 55*, 775-801.

Trice, A., & Layman, W. H. (1984). Improving Guest Surveys. *The Cornell Hotel and Restaurant Administration Quarterly, 25*, 10-13.

Tsarenko, Y., & Mavondo, F. (2001, 3-6 December). *Resources and capabilities as determinants of student satisfaction: Do local and foreign student differ.* Paper presented at the Proceedings of the Australian and New Zealand Marketing Academy Conference, Auckland, NZ.

Tuckel, P., Leppo, E., & Kaplan, B. (1993). *The Motivational Basis of Attendance at Focus Groups and Its Effect on Participant Attitudes.* Paper presented at the Proceedings of the Survey Research Methods Section, American Statistical Association

Udehn, L. (2002). The Changing Face of Methodological Individualism. *Annual Review of Sociology, 28*, 479-507.

Umbach, P., & Porter, S. (2002). How do academic departments impact students satisfaction? Understanding the Contextual Effects of Departments. *Research in Higher Education, 43*(2), 209-234.

Van den Eeden, P., & Hüttner, H. J. M. (1982). MultiLevel Research. *Current Sociology, 30*, 1-178.

Van Kenhove, P. (2002). The Influence of Topic Involvement on Mail-Survey Response Behavior. *Psychology and Marketing, 19*(3), 293-301.

Volkov, M., Harker, D., & Harker, M. (2002). Complaint Behaviour: A Study of the Differences Between Complainants About Advertizing in Australia and the Population at Large. *Journal of Consumer Marketing, 19*(4), 319-332.

Warren, M. E. (1996). Deliberative Democracy and Authority. *American Political Science Review, 90*(1), 46-60.

Whipple, T., & Muffo, J. (1982). *Adjusting for Nonresponse Bias: The Case of an Alumni Survey*. Denver: Association for Institutional Research Annual Forum.

Williams, R. (2006a). Using Heterogeneous Choice Models to COmpare Logit and Probit Coefficents Across Groups. Part I Description of the Problem & Problem with the Solution [Electronic Version] from http://www.nd.edu/~rwilliam/stats3/.

Williams, R. (2006b). Using Heterogeneous Choice Models to COmpare Logit and Probit Coefficents Across Groups. Part II A Better Solution to the Problem [Electronic Version] from http://www.nd.edu/~rwilliam/stats3/.

Willis, G. (1999). *Cognitive Interviewing. A How-to Guide.* Paper presented at the Reducing Survey Error through Research on the Cognitive and Decision Processes in Surveys, American Statistical Association.

Willis, G., DeMaio, T., & Harris-Kojetin, B. (1999). Is the Bandwagon Headed to the Methodological Promised Land? Evaluating the Validity of Cognitive Interviewing Techniques. In M. Sirken, D. Herrmann, S. Schechter, N. Schwarz, J. M. Tanur & R. Tourangeau (Eds.), *Cognition and Survey Research*. New York: John Wiley & Sons.

Willis, G., Schechter, S., & Whitaker, K. (1999). *A Comparison of Cognitive Interviewing: Expert Review and Behavior Coding: What Do They Tell Us?* Paper presented at the Proceedings of the American Statistical Association.

Winship, C., & Mare, R. (1992). Models for Sample Selection Bias. *Annual Review of Sociology, 18*, 327-350.

Wirtz, J. (2003). Halo in Customer Satisfaction Measures: The Role of Purpose of Rating, Number of Attributes and Customer Involvement. *Internation Journal of Service Industry Management, 14*(1), 96-119.

Yin, R. (2003). *Case Study Research*. Thousands Oaks: Sage Publications.

Zaller, J., & Feldman, S. (1992). A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences. *American Journal of Political Science, 36*(3), 579-616.

Zimowski, M., Tourangeau, R., Ghadialy, R., & Pedlow, S. (1997). *Nonresponse in Household Travel Surveys*: Federal Highway Administration