The Canadian Journal of Program Evaluation Vol. 24 No. 2 Pages 131–156
ISSN 0834-1516

THE DOWNWARD TREND OF SURVEY RESPONSE RATES: IMPLICATIONS AND CONSIDERATIONS FOR EVALUATORS

Teresa L. Bladon
Department of Sociology, University of Alberta
Edmonton, Alberta

Abstract: Rapidly declining response rates and the associated threat of nonresponse bias call into question the validity of data obtained through telephone surveys, a tool often used in evaluation. This article explores changes in nonresponse bias over time by examining three data points (1991, 1996, and 2002) from an annual household telephone survey conducted by the University of Alberta's Population Research Lab. Results demonstrate a substantial decline in response rates accompanied by an increasing level of bias in variables related to respondent education. Implications of these results are investigated through regression analyses and suggest that declining representation of individuals with less education could significantly impact a variety of survey variables, thus creating opportunity for opinions of the more educated to become more heavily weighted in evaluation results. In turn, such results could be used to inform government policies and programs in ways that advantage the educated middle class.

Résumé : Des taux de réponse diminuant rapidement et la menace afférente de biais dû à la non réponse mettent en doute la validité des données obtenues par l'intermédiaire d'enquêtes téléphoniques, un outil souvent utilisé en matière d'évaluation. Cet article analyse les changements observés dans le biais dû à la non réponse au fil du temps en l'examinant les données à trois points dans le temps (1991, 1996, et 2002) d'une enquête téléphonique annuelle auprès des ménages menée par le laboratoire de recherche sur la population (*Population Research Lab*) de l'Université de l'Alberta. Les résultats démontrent une baisse importante des taux de réponse accompagné d'un niveau croissant de biais dans les variables associées à la scolarité des répondants. Les répercussions de ces résultats sont étudiées par l'intermédiaire d'analyses de régression et suggèrent qu'une représentation décroissante de répondants ayant une scolarité inférieure pourrait avoir un impact significatif sur différentes variables d'enquête, permettant ainsi aux opinions des répondants plus instruits d'occuper

Corresponding author: Teresa L. Bladon, Department of Sociology, 5-21 HM Tory Building, University of Alberta, Edmonton, AB, T6G 2H4; <t_roeske@hotmail.com>

une place prépondérante dans les résultats d'évaluation. Par ricochet, de tels résultats pourraient alimenter des politiques et programmes gouvernementaux de manière à avantager la classe moyenne instruite.

INTRODUCTION

For decades surveys have been used extensively by evaluators, and in years past greater use of this method has been encouraged (see, for example, Henry, 1996). Indeed, the survey as a research and evaluation method has become firmly entrenched in contemporary society. Thus, the quality of data obtained through surveys is an important consideration, and concern, for today's evaluators. In order for surveys to provide high quality data, and thus enable the users of evaluations to make good decisions, they must be representative of the population of interest. At a time when survey response rates are on the decline (see, for example, Curtin, Presser, & Singer, 2005; de Leeuw & de Heer, 2002; Dillman et al., 2009; Groves & Couper, 1998; Groves, Dillman, Eltinge, & Little, 2002; O'Rourke et al., 1998; Steeh, Kirgis, Cannon, & DeWitt, 2001), the degree to which nonresponse rates lead to nonresponse bias and, as a result, negatively impact the representativeness of data obtained is an issue of key importance. Evaluators must be aware of the challenges associated with low response rates and remain sensitive to the potential impact on integrity of results.

The impact of nonresponse rates on nonresponse bias is an area that has received much attention in the survey research literature (see Curtin, Presser, & Singer, 2000; Groves, 2006; Groves et al., 2006; Groves & Peytcheva, 2008; Keeter, Miller, Kohut, Groves, & Presser, 2000; Merkle & Edelman, 2002; Novo, Hammarström, & Janlert, 1999; Voogt & Van Kempen, 2002). However, what has received less attention is the manner in which the nonresponse bias affecting particular survey items changes over time as survey response rates decline, particularly in a Canadian context (see Keeter, Kennedy, Dimock, Best, & Craighill, 2006, for an examination of the issue in the American context). Moreover, the implications of such trends for evaluators have not yet been fully explored.

This article brings to light the potential impact of rising nonresponse on survey variables, and implications for evaluators, through a critical examination of change over time in (a) response rates and (b) the degree to which nonresponse bias affects five sociodemographic vari-

ables commonly considered important in program evaluations (age, marital status, total years of schooling, highest level of education, and individual income). The study employs data from three points in time (1991, 1996, and 2002) collected through the Alberta Survey, an annual random-digit-dialed (RDD) household telephone survey administered by the Population Research Lab at the University of Alberta, Canada. In doing so, this article addresses several gaps in the literature as well as areas of important concern for those using telephone surveys as an evaluation method.

The first contribution made by this article relates to the dearth of studies on household survey nonresponse employing non-governmental Canadian survey data as a primary data source. Currently, the majority of studies on the topic are based on American data or, to a lesser extent, on data from European countries (see Novo et al., 1999; Voogt & Van Kempen, 2002, for examples of the latter). However, previous research suggests that there may be national differences in response rates, thereby supporting the need for Canadian-based research on survey nonresponse (see, e.g., de Heer, 1999; de Leeuw & de Heer, 2002; Goyder, 1985). The research that does employ at least some Canadian data is based largely on government surveys (see de Heer, 1999; Groves & Couper, 1998; Smith, 1995), which tend to achieve a higher response rate than do surveys conducted by universities and private survey research firms (de Heer, 1999; Groves & Couper, 1998)—organizational types that often undertake evaluative activities—thereby limiting the generalizability of such work (Smith, 1995). The current study begins to bridge this gap by analyzing the issue of nonresponse through the use of data obtained in an Alberta-based telephone survey conducted by a university research lab.

Second, this research measures changes in both the magnitude and the direction of nonresponse bias over time across a variety of dimensions. Rather than focusing on the bias present in a survey at a single point in time, this study extends such an analysis to obtain a more longitudinal picture of how nonresponse bias has changed from 1991 to 2002 in the Alberta Survey.

Finally, this study attempts to demonstrate the significant research, evaluation, and policy implications of nonresponse bias through examining the relationships between the sociodemographic variables under analysis and a wide range of topic areas. In doing so, the study highlights the practical implications of nonresponse for survey areas that may be of interest to evaluators, as well as those that are

pertinent in matters of policy and public discourse. Moreover, the discussion explores possible ways in which evaluators can address, or lessen the impact of, nonresponse bias and examines current societal trends impacting the role of the RDD household telephone survey.

METHOD

The Alberta Survey

The research presented in this article was conducted using data obtained from the Alberta Survey, an annual omnibus household telephone survey administered since 1987 by the Population Research Laboratory housed within the University of Alberta. The Alberta Survey is a survey of Alberta adults (over 18 years of age) where households are selected through random digit dialing. The survey includes research questions that interrogate a wide range of issues related to public policy, although the specific subject matter of the survey varies from year to year.

Each year approximately 1,200 Alberta adults are included in the sample—400 from each of three geographic areas: Edmonton, Calgary, and other Alberta. Attempts are made to interview one respondent from within each household, and sampling quotas are implemented on the basis of geographic region (as outlined above) and gender (50% males, 50% females).

The minimum number of callback attempts to each household varied from 10 to 15 over the time period examined. Interviews were typically administered from 9:00 a.m to 9:00 p.m. weekdays, and during shorter time blocks on Saturdays and Sundays. Each year data collection occurred over a roughly two-month period, typically between February and May, although from 1997 to 2002 the time of administration shifted to between October and February.

Calculating Call Disposition Rates

Response rates to the Alberta Survey were calculated by dividing the number of completed interviews by the number of eligible households interviewers attempted to contact (telephone numbers that are nonresidential or not in service are considered ineligible). This method corresponds to that defined in the American Association for Public Opinion Research (AAPOR) *Standard Definitions* as "Response Rate

1" (AAPOR, 2004). Refusal, noncontact, and other nonresponse rates were calculated by dividing the call disposition under analysis (i.e., refusal, noncontact, or other nonresponse) by the number of eligible households where contact was attempted.

Measurement of Nonresponse Bias

To investigate the extent to which particular survey items were affected by nonresponse bias, a three-part analysis process was employed.

1. Comparison of Alberta Survey Data with Canadian Census Data

First, the distributions of five sociodemographic variables from the Alberta Survey were compared with Canadian census data, including age, marital status, total years of schooling, highest level of education, and individual income. Comparisons were drawn for the years 1991, 1996, and 2002 to correspond with Canadian census years (at the time of initial research, 2006 census data were not available).[1] Comparisons were made using data for individuals over 18 years of age residing in the province of Alberta to ensure correspondence of the sampling frames for the two data sets.

2. Calculation of the Index of Dissimilarity

Following comparison of Alberta Survey and census data, the index of dissimilarity (ID) was calculated for each of the five sociodemographic variables under analysis. The ID has been used as a measure of the difference between two age distributions (Hobbs, 2004, p. 157), a use that is replicated in this study and extended to the other four background variables analyzed.

The index of dissimilarity can be interpreted as the proportion of cases in one group that would have to be moved to other categories to make two distributions identical (Agresti, 2002). For example, if the age distributions of two groups, X and Y, are being compared and the ID is equal to 7%, then 7% of the cases in group X would have to be relocated to other age categories in order to make the distributions of X and Y equivalent. The greater the proportion of cases that must be relocated, the less alike the two distributions are. Therefore, in this study, smaller values of the ID are more desirable as they indicate more representative data (i.e., a higher degree of similarity between population and survey data). (For more information on the ID see Kuha & Firth, 2005; McKibben & Faust, 2004.)

The index of dissimilarity is calculated for age, marital status, highest level of education, total years of schooling, and individual income for the years 1991, 1996, and 2002. In addition, the ID is calculated for the variable age for the year 2005, comparing Alberta Survey data with postcensal estimates (obtained from Statistics Canada, 2006a).

Note that while this approach provides an estimate of the nonresponse bias present along a number of demographic variables, it does not allow one to control for the influence of other types of survey error, such as measurement error, which could potentially impact these variables. Measurement error occurs when there is a discrepancy between the true value of a variable and the value obtained through data collection. One type of measurement error can occur when there is a desire by respondents to provide answers that they feel would be viewed positively by others as opposed to answering survey questions truthfully, thereby resulting in social desirability bias.

Although both sources of demographic variables used here (Alberta Survey and census data) are self-reported, it is possible that the former could be impacted more by social desirability bias as a result of the mode of data collection used. Specifically, while Alberta Survey data were gathered through telephone interviews, census data are collected through self-administered questionnaires, and research has demonstrated that self-administered modes of data collection produce less social desirability effects than do interviewer-administered modes (see, for example, Kreuter, Presser, & Tourangeau, 2008; Presser & Stinson, 1998; Richman, Kiesler, Weisband, & Drasgow, 1999; Tourangeau & Yan, 2007).

However, evidence also suggests that the impact of social desirability becomes more pronounced as the perceived sensitivity of the topic increases (Catania, Gibson, Chitwood, & Coates, 1990). Thus, social desirability effects on the demographic variables investigated in this study are likely less than may be the case for topics perceived as more sensitive such as drinking, alcohol abuse, and views on abortion or gay rights, for example. Nonetheless, it is important to keep this potentially confounding bias in mind when interpreting results of this study; as such, there is further discussion of the issue in the Results section.

3. Ordinary Least Squares Regression

Another limitation of the ID is that direct comparison between survey and population data is only possible for variables that are measured

in both data sets. Thus, it is only possible to calculate the index of dissimilarity for a limited selection of sociodemographic variables. To extend the index of dissimilarity analyses to items representing a broader range of topics, ordinary least squares (OLS) regression was employed to regress a variety of attitudinal and behavioural items on sociodemographic variables. These regressions were conducted to explore the relationships between sociodemographic variables and respondent attitudes and behaviours while controlling for other potentially related variables. As Bethlehem (2002, p. 287) states, "if there are sufficiently strong relationships between auxiliary variables and target variables, … conclusions with respect to a possible bias carry over to the target variables."

A total of 18 regression analyses were conducted with a wide range of attitude/behaviour items as the dependent variables, chosen to represent a wide range of topics that could serve as variables of interest in evaluations. Topics investigated are summarized in the Results section in Table 2.

Eight control variables were used across the regression analyses: sex, age, total years of schooling, marital status, whether the respondent lived in an urban or rural location,[2] number of people in the household, whether the respondent owned or rented a home, and individual income.

RESULTS

Trends in Alberta Survey Call Disposition Rates

Trends in response rates as well as in refusal, noncontact, and other nonresponse rates were examined for the Alberta Survey from 1991 to 2007.[3] The other nonresponse rate includes nonresponse due to circumstances such as language barriers and physical illness or disability. The trends for these four call disposition rates are summarized in Figure 1.

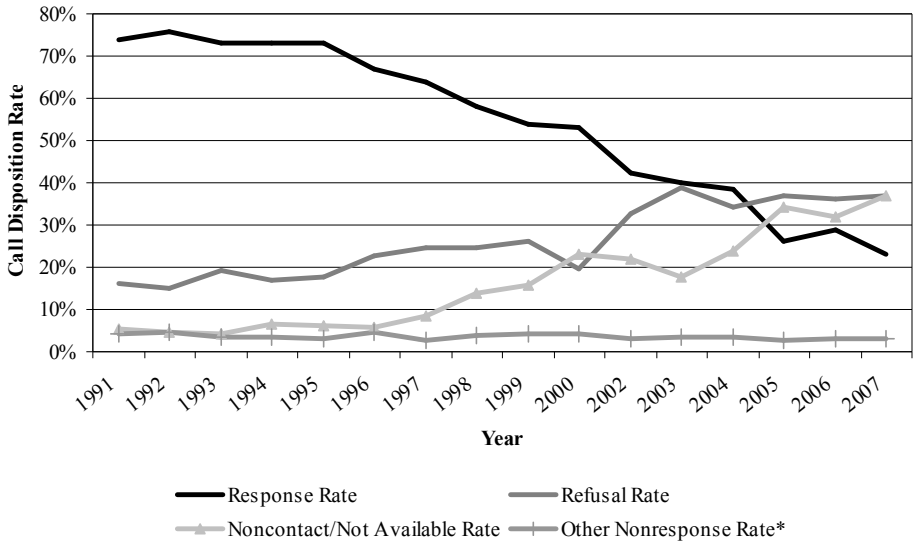
Figure 1 clearly demonstrates a decline in response rates, particularly since the mid-1990s, mirrored by a rise in both refusals and noncontacts. The Alberta Survey response rate dropped from 74% in 1991 to only 23% in 2007—a substantial decline of 51%. Refusal rates more than doubled over the same time period, beginning at 16% in 1991 and rising to 37% by 2007. Noncontact rates experienced an even greater increase, starting at only 5% in 1991 and jumping to 37% in

2007. On the other hand, the other nonresponse rate remained fairly consistent from 1991 to 2007.

Figure 1 indicates that refusal rates have been increasing fairly steadily over time despite several isolated years where a slight drop occurred (e.g., 1994–1995 and 2000). The sharpest increase in refusals occurred from 2000 to 2003. While the trend for noncontact rates also shows a substantial increase over time, the trend is of a slightly different form than is the case for refusals. Specifically, while refusal rates show an overall steady increase from 1991 to 2005 followed by a levelling off, noncontact rates remain relatively stable in the period from 1991 to 1996. It is from 1996 onwards that noncontact rates begin the steep ascent that continues to the present, despite a slight drop in 2003. The other nonresponse rate does not display any overall trend for the time period examined and accounts for a very small proportion of total nonresponse.

Alberta Survey Sampling Reports from 1991 to 2007 (Population Research Laboratory, 1991–2007) were reviewed to determine if changes in response rates could be accounted for, at least in part, by changes in data collection methods. Variation in the number of callback attempts, time of year of survey administration, length of

Figure 1
Alberta Survey Call Disposition Rates: 1991–2007



*Includes nonresponse due to circumstances such as language barriers and physical illness or disability.

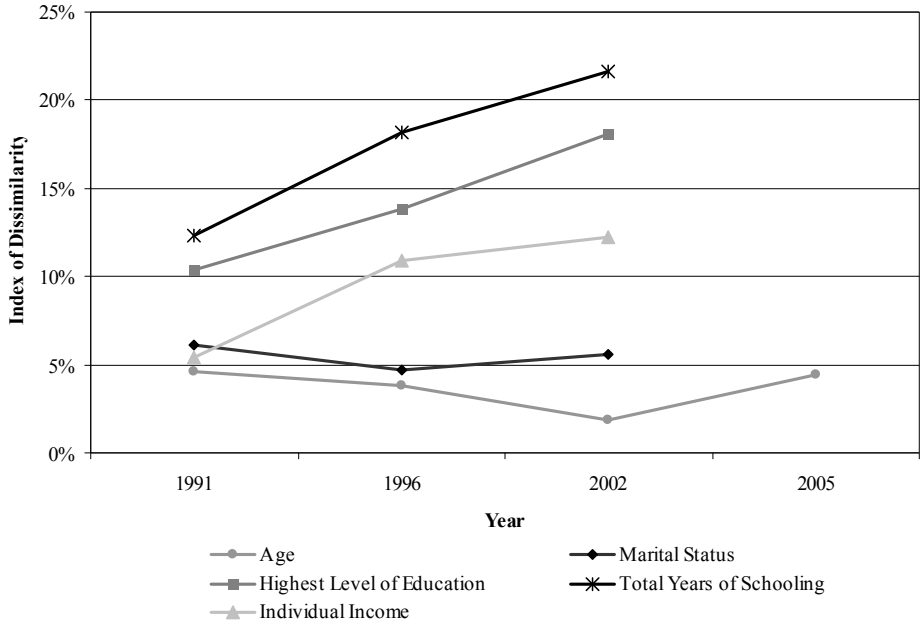
data collection period, and mean interview time were examined. The review of the Sampling Reports revealed that Alberta Survey methods did undergo some minor changes; however, none were able to account for the consistent reduction in response rates or the rise in refusals and noncontacts observed (see Roeske, 2007, for more detail). Thus, there are other causes independent of the survey itself resulting in the trends observed in call disposition rates.

Index of Dissimilarity Analyses: Investigating Changes in the Magnitude and Direction of Nonresponse Bias Present in Alberta Survey Data

Index of dissimilarity scores were calculated for reported age, marital status, highest level of education, total years of schooling, and individual income for the years 1991, 1996, and 2002. The results of these calculations pertain to the magnitude of nonresponse bias present in the variables analyzed and how it changes over time, and are summarized in Figure 2.

Figure 2 reveals that the index of dissimilarity scores for the two education variables—highest level of education and total years of

Figure 2
Index of Dissimilarity Scores for Age, Marital Status, Highest Level of Education, Total Years of Schooling, and Individual Income



schooling—are consistently higher than the indices for the other variables examined and increase steadily from 1991 to 2002. The index of dissimilarity for highest level of education rises from 10% in 1991 to 18% by 2002, an increase of 8%. For total years of schooling the ID reaches 22% by 2002, the highest value attained among all the variables examined and an increase of 9% from 1991 levels. The overall steady rise in index of dissimilarity scores for these two variables suggests that Alberta Survey data are becoming, over time, less representative of the actual level of education held by Albertans.

While ID scores for individual income are lower than those for either of the education variables, the scores do reveal an overall increase from 1991 to 2002. The ID for individual income increases by 7% over the 11-year period, beginning at 5% in 1991 and reaching 12% by 2002. This trend indicates that, as is the case for education, Alberta Survey data are becoming increasingly unrepresentative of the Alberta population along the dimension of individual income.

The ID scores for age and marital status are the lowest among the five variables analyzed. Moreover, unlike the previous variables examined, Figure 2 demonstrates that trends in ID scores for these two variables do not indicate an overall upward or downward pattern in the time period examined. Thus, these analyses suggest that there is less bias present in either marital status or age than is the case for highest level of education, total years of schooling, and individual income.

Discussion of index of dissimilarity scores has thus far focused on changes in the magnitude of nonresponse bias present in the five variables analyzed. However, in order to determine the manner in which bias is impacting Alberta Survey results it is necessary to determine the direction of bias. To do so, the proportion of cases in each category for the three variables demonstrating the greatest potential for bias based on the ID were compared with Canadian census data for all Alberta (Difference = Alberta Survey proportion – Canadian census proportion). A negative value indicates that the category is underrepresented in the Alberta Survey and a positive value that it is overrepresented. The results are presented in Table 1.

Table 1 reveals that, overall, individuals with the lowest reported levels of education are becoming increasingly underrepresented in the Alberta Survey while those with higher reported levels of education are increasingly overrepresented. These results suggest that, over time, Alberta Survey data have become progressively less representa-

tive of the actual level of education of the Alberta population and progressively more biased toward individuals with more schooling.

Table 1
Differences[a] Between the Distribution of Variables in the Alberta Survey and Canadian Census for All Alberta

	1991	1996	2002
Highest Level of Education			
Less than high school	-9.9%	-13.8%	-14.7%
High school graduate	7.0%	7.9%	8.6%
Some non-university	-0.6%	0.6%	-3.4%
Some university	3.4%	5.4%	9.5%
Total Years of Schooling			
Less than 9 years	-2.8%	-3.8%	-4.6%
9–12 years	-9.5%	-14.4%	-17.0%
13–17 years	11.3%	14.0%	16.3%
18 or more years	1.0%	4.3%	5.3%
Individual Income			
< $16,000	-5.3%	-10.9%	-9.3%
$16,000–25,999	0.1%	0.0%	-3.0%
$26,000–35,999	1.7%	3.2%	0.6%
$36,000–45,999	0.4%	2.4%	1.9%
$46,000–54,999	0.9%	0.6%	3.5%
$55,000–64,999	0.7%	0.3%	1.6%
$65,000–74,999	0.5%	0.9%	1.5%
$75,000–84,999	0.4%	0.8%	0.3%
$85,000–99,999	0.0%	0.9%	0.6%
$100,000+	0.7%	1.6%	2.1%

[a]Differences were calculated as (Alberta Survey proportion) – (Canadian census proportion).

While the differences between individual categories are much smaller than is the case for the two education variables, a similar trend emerges when individual income is examined. Specifically, Table 1 demonstrates that those with the lowest reported income are underrepresented in Alberta Survey data while those with average and high levels of reported income are overrepresented. Given the findings for the two variables related to education, the trend apparent for individual income is not surprising as, in general, those with more education also tend to have a higher income.

The index of dissimilarity results lend themselves to two possible explanations: either those with higher levels of education and income are overrepresented in the data set (nonresponse bias), or respondents are increasingly reporting higher levels of education and income than they actually possess (social desirability bias). While both respondent education and income could be impacted by social desirability bias, income is typically viewed as more sensitive than level of education and thus one would expect greater misreporting by respondents along this variable (see, for example, Tourangeau & Yan, 2007). However, results demonstrate a greater discrepancy between Alberta Survey and census data for respondent education, suggesting at least some level of nonresponse bias present in the data. Overall, it seems that Alberta Survey data are becoming more biased towards those with a higher income as well as toward those with a higher level of education.

Regression Results: Relationships of Sociodemographic Variables and Measures of Behaviour and Attitude

Ordinary least squares regression was used to determine the relationship between the 5 sociodemographic variables analyzed and 18 attitude and behaviour items falling into four broad categories: health issues, social issues, environmental issues, and crime- and surveillance-related issues. The control variables used across all 18 regressions were sex, age, total years of schooling, marital status, whether the respondent lived in an urban or rural location, number of people in the household, whether the respondent owned or rented a home, and individual income. Total years of schooling demonstrated the greatest number of significant associations (10) followed by age (9) and sex (8). The urban/rural variable demonstrated 6 significant associations, household size and income 3, and marital status and home ownership each displayed only 1 significant association across all 18 regression analyses. The beta values, standard errors, and levels of significance for the 4 control variables where 5 or more significant associations were observed across the regression analyses (i.e., were significantly related to at least 25% of the dependent variables investigated) are presented in Table 2.

The results from the regression analyses indicate that total years of schooling, age, and sex are significantly associated with the greatest number of dependent variables across a wide range of topics (see Table 2). Thus, any bias present in these three demographic variables also has the greatest potential to impact data related to the widest range of attitude and behaviour items. When considered in

conjunction with results from the index of dissimilarity analyses, particularly those related to total years of schooling, these findings are of significant consequence. The index of dissimilarity calculations revealed not only that Alberta Survey data were biased toward those with more schooling, but that data were becoming less representative of the actual distribution of Albertans' level of education over time. Results of the regression analyses suggest that the bias present in total years of schooling, which is already high and experiencing a steady upward climb, has the potential to influence more attitude and behaviour items than is the case for any of the other sociodemographic variables investigated. Thus, it is likely that the 10 variables for which total years of schooling is significantly related are biased, at least to some degree.

Age was significantly related to 50% of the dependent variables analyzed while sex was related to 44%, and each had at least one significant effect in all four attitude/behaviour categories included in the regressions. These results suggest that any bias present in the variables sex and age is likely to impact a wide range of other variables as well, thereby negatively affecting the representativeness of results. While it was impossible to examine the degree of bias in the variable sex based on Alberta Survey data due to the use of a quota sample, some research does suggest that sex, or gender, may be a correlate of nonresponse, thereby opening up space for bias to enter into survey results (see Goyder, 1987; Groves & Couper, 1998; Smith, 1983).

Index of dissimilarity analyses indicated that there was no clear trend for age in terms of representativeness of the Alberta population. However, results did suggest that age may become more biased over time if the trend observed from 2002 to 2005 continues. Given that age is significantly related to a wide range of variables (see Table 2), there is potential for bias in this variable to have a substantial impact. Therefore, the representativeness of survey data along the dimension age should be closely monitored in future years.

DISCUSSION

Implications for Evaluators

The analyses presented here have significant implications for evaluators. Results of this study show that response rates to the Alberta Survey are indeed rapidly declining, with a more pronounced decline in later years. A steady rise in refusals and a more recent increase

Table 2
Beta Values, Standard Errors, and Significance Levels for Total Years of Schooling, Sex, Age, and Urban or Rural Residence

Dependent Variable	Sex (male = 1; female = 0)			Age			Years of Schooling			Urban/Rural (urban = 1; rural = 0)			
	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β	R^2 [a]
Health issues													
Long-term care regulation	.25**	.08	.11**	-.00	.00	-.02	-.01	.01	-.02	-.21*	.08	-.09*	.01
Long-term care subsidies	.27**	.10	.10**	.00	.00	.03	-.00	.02	-.00	-.09	.10	-.03	.01
Physical activity	.00	.08	.00	-.01**	.00	-.15**	.04**	.01	.10**	.16	.09	.06	.06
Social issues													
Funding for public education	-.22	.13	-.06	-.01	.00	-.06	.08**	.02	.13**	.13	.13	.04	.03
Health of democracy	.14	.08	.06	.00	.00	.05	-.02	.01	-.04	.02	.08	.01	.00
Influence of big business	-.04	.06	-.02	.00	.00	.03	.03**	.01	.12**	-.03	.07	-.02	.01
Government hides information	-.13	.09	-.06	.01*	.00	.09*	.01	.01	.02	-.03	.09	-.01	.01
Support for US war on terrorism	.35**	.10	.12**	.01**	.00	.12**	-.06**	.02	-.13**	-.24*	.11	-.07*	.07
Openness to immigrants after Sept. 11, 2001	.11	.09	.04	-.01	.00	-.06	.06**	.01	.14**	.48**	.10	.16**	.05
Welfare recipients	.29*	.12	.08*	-.02**	.00	-.13**	-.12**	.02	-.20**	-.16	.13	-.04	.07
Same-sex marriages	-.53**	.11	-.16**	-.03**	.00	-.27**	.09**	.02	.17**	.47**	.11	.13**	.15

Dependent Variable	Sex (male = 1; female = 0)			Age			Years of Schooling			Urban/Rural (urban = 1; rural = 0)			
	B	SE B	β	B	SE B	β	B	SE B	β	B	SE B	β	$R^{2\,a}$
Environmental issues													
Health of environment in Alberta	.18**	.07	.10**	.01*	.00	.09*	-.01	.01	-.03	.04	.07	.02	.01
Influence of environment-alists	.14	.08	.07	.01**	.00	.14**	-.06**	.01	-.17**	-.25**	.08	-.11**	.08
Protected areas	-.10*	.05	-.07*	-.00	.00	-.06	.01	.01	.03	.07	.05	.05	.01
Risks from drinking tap water	-.09	.07	-.05	-.00	.00	-.01	-.02	.01	-.06	.03	.07	.02	.00
Crime-related issues													
Support for video surveillance	-.40**	.10	-.14**	.00	.00	.05	-.03*	.02	-.07*	-.16	.10	-.05	.04
Concern with new surveillance technologies	-.15	.09	-.06	.01**	.00	.11**	.00	.01	.00	-.08	.10	-.03	.01
Neighbourhood crime	-.11	.10	-.04	.01**	.00	.11**	-.05**	.02	-.11**	.43**	.10	.14**	.03
Total # of significant effects	8		8	9		9	10		10	6		6	

[a]Adjusted R^2
*$p < .05$ **$p < .01$

in noncontacts appear to be the primary culprits for falling response rates. However, rising refusal rates are also a concern of private survey research firms (see Allen, 2001). As evaluators, it is important to recognize that while response rates to surveys administered by university-affiliated organizations are much lower than those administered by government, response rates to surveys administered by private-sector organizations are, in some cases, experiencing even greater difficulties. For example, drawing on information from the Council for Marketing and Opinion Research (CMOR), Steeh et al. (2001, p. 228) state that response rates to RDD commercial telephone surveys average around 12%. Note that this article was published a number of years ago, and it is likely that average response rates have declined even further since that time. Such low response rates could significantly impact the integrity of data obtained.

However, it is not the response rate itself that is inherently problematic, but rather the implications of declining response rates. Specifically, as response rates to the Alberta Survey have been falling, index of dissimilarity scores, representing (at least in part) nonresponse bias, have been increasing. Of the variables examined in this study, education and individual income appear to be impacted most by nonresponse bias and thus provide the greatest cause for concern. Moreover, the bias present in these two dimensions is growing over time, rendering Alberta Survey data increasingly biased toward individuals with higher levels of both education and income.

Finding such an education bias is not unique to this research. A recent study examining nonresponse errors occurring across a number of survey modes (mail, telephone, interactive voice response, and the Internet) found that respondents were in all cases significantly more likely to have higher levels of education than were nonrespondents (Dillman et al., 2009). Researchers in this study were able to obtain demographic characteristics for all members of the population and as such were not relying on reported levels of education, meaning that this difference cannot be attributed to social desirability bias. In another study, Keeter et al. (2006, p. 764) found that respondents to a survey conducted by the Pew Research Center in 2003 were also more likely to have completed high school and more likely to have a college degree when compared to the adult population of the United States, further supporting the finding of an education bias. Thus, evaluators should pay particular attention to the distribution of respondents' level of education and maintain sensitivity to possible bias in this and related variables.

The education bias is particularly disconcerting as level of education is significantly associated with a wide range of attitudes related to health, social, environmental, and crime- and surveillance-related issues. Overrepresentation of individuals with higher levels of education could lead to biased results for a wide range of topics, thereby undermining the very purpose of a random sample survey, that is, to accurately represent the views of the population sampled. As many evaluations are targeted at determining impact of government policy and program initiatives, what may be measured most is impact on the educated middle class. In turn, it is possible that household telephone survey results could help inform government policies and programs in ways that advantage, or have the most positive impact on, the educated middle class.

Addressing the Problem

While nonresponse bias poses a potentially significant problem for evaluators, steps can be taken to reduce its impact at two stages of an evaluation: data collection and data analysis. When considering data collection, results of this study can be used to inform a sampling design to account for the overrepresentation of individuals with higher levels of reported education. Specifically, adopting a stratified random sampling approach would allow evaluators to ensure that the sample is representative of the population in terms of educational attainment. Stratification is an effective method to ensure that the distribution of a sample is representative of the population along a variable of interest, such as level of education (Fowler, 2002). Stratification is particularly useful when one suspects that a simple random sample may not accurately represent the population of interest along a variable that is important to the study, and thus provides one relatively simple method to minimize some of the effects of nonresponse bias. The relative size of each stratum (i.e., level of educational attainment) could be determined by examining census data for the population of interest.

As it is almost certain that nonresponse will be a component of all surveys regardless of sampling design, post-survey adjustment is a popular method of minimizing the effects of nonresponse bias at the data analysis stage. Post-survey adjustment generally includes methods such as weighting the data (including poststratification), imputation, or modelling nonresponse (Groves, Cialdini, & Couper, 1992; Weisberg, 2005). While the difficulty of both assessing and subsequently correcting for the "negative effects of nonresponse" is

emphasized by Bethlehem (2002, p. 287), weighting adjustments, such as poststratification, are often viewed as providing greater accuracy to survey estimates. Gelman and Carlin (2002, p. 291) state that "considerable gains can be made by poststratifying on variables that are predictive of survey outcomes."

However, there are limitations to post-survey adjustment techniques. For example, in their study of nonresponse bias in the Dutch national election, Voogt and Van Kempen (2002) found that bias in political and voting behaviours of respondents was larger than bias in background variables, meaning that there was not a clear relationship between demographic variables and the key variables of interest (political and voting behaviours). Hence, weighting the data based on demographic information would not correct for the bias in the results. In order for post-survey statistical adjustments such as weighting to be effective in reducing bias, the weighting variables must be predictive of, or related to, key variables of interest (see Gelman & Carlin, 2002; Weisberg, 2005).

Thus, while in many cases post-survey adjustment enables a reduction of the discrepancies between particular survey estimates and population values, it is not without flaws or limitations. Nonetheless, such adjustment is perhaps the best method currently available to evaluators attempting to minimize the effects of nonresponse bias in their data sets and can be effective, particularly if the degree to which background variables are related to key variables of interest is considered prior to adjustment.

Considering the Future of the RDD Telephone Survey

The overrepresentation of those with higher levels of education and/or income in survey data is fundamentally at odds with the purpose of evaluation and the RDD telephone survey, both of which aim to collect data that are representative of the characteristics and opinions of the population of interest. Data are to transcend the boundaries of class through equal representation of all individuals within the population, at least to the greatest extent possible when possession of a landline telephone is a requirement for survey participation. Whatever the reason for the overrepresentation of those with higher levels of education in household telephone surveys, it is a trend that should provide much cause for concern among evaluators who use this method in their work. The goal of obtaining representative data is still at the heart of survey research, but the degree to which this goal is being

achieved may be declining over time and creating a situation where the weight of the opinions of those with less education is diminished.

There are other current trends that challenge the efficacy of the RDD telephone survey as a research tool, a primary one being the growth of cellphone-only households. This trend serves to exclude an expanding proportion of the population from random-digit-dialed sampling frames, thereby opening up space for coverage bias to impact survey results (Brick, Dipko, Presser, Tucker, & Yuan, 2006; Singer, 2006). Specifically, there is evidence to suggest that there are differences between cellphone-only households and other households. Canadians who rely solely on cellphones tend to be lower-income and live in large urban centres (Statistics Canada, 2006b). American studies indicate that individuals who are lower-income, younger, renting, and living alone or single are more likely to opt for a cellphone instead of a landline (see Blumberg, Luke, & Cynamon, 2006; de Leeuw, Lepkowski, & Kim, 2002; Keeter, 2006; Pew Research Center, 2006). These demographic differences have the potential to produce biased results for survey items with which they are correlated, particularly if the proportion of cellphone-only households continues to grow.

Trends such as this, coupled with the results of this study, suggest that telephone survey research does not have an easy road ahead. One may even ask the question of whether we are witnessing the end of an era, so to speak, where the relevance and validity of the RDD telephone survey as an evaluation tool is diminished. While perhaps overly dramatic in nature, this question is certainly pertinent today, particularly in the context of the current study. Despite the somewhat sobering results of this study, it is difficult to provide either a clearly positive or cleary negative response to the question posed. The challenges telephone survey researchers are facing today may be different in form and/or degree than those faced in the past, but that is not to say there is no hope of contending with them.

In fact, new obstacles and challenges are often a driving force behind innovation, and thus should not be viewed in a purely negative light. Rather, they provide opportunity for change and growth. As Cliff Zukin, recent president of AAPOR, states, "we need to innovate and adapt so as to continue to be relevant" (Zukin, 2006, p. 432). To contend with the challenges posed by declining response rates, evaluators relying on this research method must display flexibility and creativity while maintaining a commitment to rigorous methods and designs. This is, to say the least, a tall order.

In light of the widespread and rapid decline of response rates to telephone surveys, what may be required is a reconsideration of research design to both increase response rates and lessen the degree of potential non-coverage of the population. For example, declining response rates have already resulted in increased use of mixed mode designs (Dillman & Christian, 2005). These designs may employ a variety of mediums to achieve survey response, such as the telephone, mail, and increasingly, the Internet. However, while mixed mode research designs provide an alternative to the telephone-only survey method, they do have their own set of challenges of which evaluators should be aware. For example, mode effects and the issue of cognitive equivalence of survey items can pose challenges in mixed mode surveys (see, for example, Cobben, Schouten, & Bethlehem, 2006; Dillman & Christian, 2005; Dillman et al., 2009). In addition, recent studies have found that telephone survey methods actually result in higher response rates than do other methods such as interactive voice recognition (IVR) and Internet surveys (Dillman et al., 2009; Kreuter et al., 2008). One study found that while switching to a second mode was effective in improving response rate, it did not necessarily reduce nonresponse error based on demographics, such as respondent education (Dillman et al., 2009). Thus, while the prevalence of mixed mode and alternative modes of surveying (such as online methods) will likely increase, evaluators should recognize that it is unlikely that such changes will address the problem of nonresponse bias discussed here.

While there is no clear solution to the challenges faced by evaluators relying on telephone surveys, what is clear is that the status quo is, in many cases, declining in effectiveness. As current trends continue and response rates remain at a very low level, the space for nonresponse bias to enter into survey results becomes ever wider and the threat to validity higher. Change and innovation of some form are required. At this point in time, the future of RDD household telephone surveys remains tenuous. This uncertain future is somewhat unsettling given the extent to which telephone survey data are currently collected and used for purposes ranging from planning and decision-making in both the private and public sectors (see Gandy, 2003; Herbst, 1993; Tourangeau, 2004), to the shaping of government communications strategies (Page, 2006), to scholarly research. Evaluators have a daunting but essential task ahead of them in attempting to address the response rate challenge threatening this widely popular research tool. Potential strategies, such as the use of stratified sampling designs and post-survey adjustment techniques, can help to minimize the impact of nonresponse bias, but a continued decline in response rates

coupled with other trends such as growing noncoverage will require both innovation and creativity on the part of evaluators relying on the results of RDD telephone surveys for their evaluations.

ACKNOWLEDGEMENTS

The author thanks Dr. Harvey Krahn for the ongoing support received throughout this research, as well as the staff at the Population Research Laboratory at the University of Alberta. Comments received from Dr. Krahn and Dr. Dale Howard on earlier drafts of this article are greatly appreciated.

NOTES

1. The Alberta Survey was not administered in 2001 and, therefore, Canadian census data from 2001 were compared with Alberta Survey data from 2002.

2. For the purposes of these analyses, "urban" was defined as city only, while towns and villages were included in the "rural" category.

3. In 2001 the Population Research Lab did not conduct the Alberta Survey. In 1995 and 1996 the Alberta Survey was conducted in three and two segments respectively, each consisting of approximately 1,200 respondents, due to the large number of questions submitted by researchers. Data from the first iteration of the Alberta Survey for both years were used for all analyses as time of year of survey administration corresponded most closely with that of previous Alberta Surveys.

REFERENCES

Agresti, A. (2002). *Categorical data analysis* (2nd ed.). New York: Wiley-Interscience.

Allen, M. (2001). *Telephone refusal rates continue to rise*. Marketing Research and Intelligence Association. Retrieved from http://www.mria-arim.ca/COMMITTEES/ResponseArticle02.asp

American Association for Public Opinion Research. (2004). *Standard definitions: Final dispositions of case codes and outcome rates for surveys* (3rd ed.). Lenexa, KS: Author.

Bethlehem, J. (2002). Weighting nonresponse adjustments based on auxiliary information. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little (Eds.), *Survey nonresponse* (pp. 275–288). New York, NY: Wiley.

Blumberg, S.J., Luke, J.V., & Cynamon, M.L. (2006). Telephone coverage and health survey estimates: Evaluating the need for concern about wireless substitution. *American Journal of Public Health, 96,* 926–931.

Brick, J.M., Dipko, S., Presser, S., Tucker, C., & Yuan, Y. (2006). Nonresponse bias in a dual frame sample of cell and landline numbers. *Public Opinion Quarterly, 70,* 780–793.

Catania, J.A., Gibson, D.R., Chitwood, D.D., & Coates, T.J. (1990). Methodological problems in AIDS behavioral research: Influences on measurement error and participation bias in studies of sexual behavior. *Psychological Bulletin, 108,* 339–362.

Cobben, F., Schouten, B., & Bethlehem, J. (2006, August). *Mixed mode and nonresponse: Opportunities and challenges.* Paper presented at the 17th International Workshop on Household Survey Nonresponse, Omaha, NE.

Curtin, R., Presser, S., & Singer, E. (2000). The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly, 64,* 413–428.

Curtin, R., Presser, S., & Singer, E. (2005). Changes in telephone survey nonresponse over the past quarter century. *Public Opinion Quarterly, 69,* 87–98.

de Heer, W. (1999). International response trends: Results of an international survey. *Journal of Official Statistics, 15,* 129–142.

de Leeuw, E., & de Heer, W. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little (Eds.), *Survey nonresponse* (pp. 41–54). New York: Wiley.

de Leeuw, E., Lepkowski, J., & Kim, S.W. (2002, August). *Have telephone surveys a future in the 21st century?* Paper presented at the International Conference on Improving Surveys, Copenhagen, Denmark. Retrieved from http://www.icis.dk/ICIS_papers/F3_2_2.pdf

Dillman, D.A., & Christian, L.M. (2005). Survey mode as a source of instability in responses across surveys. *Field Methods, 17,* 30–52.

Dillman, D.A., Phelps, G., Tortora, R., Swift, K., Kohrell, J., Berck, J., et al. (2009). Response rate and measurement differences in mixed-mode surveys using mail, telephone, interactive voice response (IVR) and the Internet. *Social Science Research, 38,* 1–18.

Fowler, F.J., Jr. (2002). *Survey research methods* (3rd ed.). Thousand Oaks, CA: Sage.

Gandy, O.H. (2003). Public opinion surveys and the formation of privacy policy. *Journal of Social Issues, 59,* 283–299.

Gelman, A., & Carlin, J.B. (2002). Poststratification and weighting adjustments. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little (Eds.), *Survey nonresponse* (pp. 289–302). New York, NY: Wiley.

Goyder, J. (1985). Nonresponse on surveys: A Canada-United States comparison. *Canadian Journal of Sociology, 10*(3), 231–251.

Goyder, J. (1987). *The silent minority: Nonrespondents on sample surveys*. Cambridge, UK: Polity Press.

Groves, R.M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly, 70,* 646–675.

Groves, R.M., Cialdini, R.B., & Couper, M.P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly, 56,* 475–495.

Groves, R.M., & Couper, M.P. (1998). *Nonresponse in household interview surveys*. New York, NY: Wiley.

Groves, R.M., Couper, M.P., Presser, S., Singer, E., Tourangeau, R., Piani Acosta, G., et al. (2006). Experiments in producing nonresponse bias. *Public Opinion Quarterly, 70,* 720–736.

Groves, R.M., Dillman, D.A., Eltinge, J.L., & Little, R.J. (2002). Preface. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little (Eds.), *Survey nonresponse* (pp. xiii–xv). New York, NY: Wiley.

Groves, R.M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly, 72,* 167–189.

Henry, G. (1996). Does the public have a role in evaluation? Surveys and democratic discourse. *New Directions for Evaluation, 70,* 3–15.

Herbst, S. (1993). *Numbered voices: How opinion polling has shaped American politics*. Chicago, IL: University of Chicago Press.

Hobbs, F.B. (2004). Age and sex composition. In J. Seigel & D. Swanson (Eds.), *The methods and materials of demography* (2nd ed., pp. 125–173). San Diego, CA: Elsevier Academic Press.

Keeter, S. (2006). The impact of cell phone noncoverage bias on polling in the 2004 presidential election. *Public Opinion Quarterly, 70,* 88–98.

Keeter, S., Kennedy, C., Dimock, M., Best, J., & Craighill, P. (2006). Gauging the impact of growing nonresponse on estimates from a national RDD telephone survey. *Public Opinion Quarterly, 70,* 759–779.

Keeter, S., Miller, C., Kohut, A., Groves, R.M., & Presser, S. (2000). Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly, 64,* 125–148.

Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly, 72,* 847–865.

Kuha, J., & Firth, D. (2005). *On the index of dissimilarity for lack of fit in loglinear and log-multiplicative models.* Retrieved from http://stats.lse.ac.uk/kuha/Publications/kuha_firth_Delta.pdf

McKibben, J.N., & Faust, K.A. (2004). Population distribution: Classification of residence. In J. Seigel & D. Swanson (Eds.), *The methods and materials of demography* (2nd ed., pp. 105–123). San Diego, CA: Elsevier Academic Press.

Merkle, D.M., & Edelman, M. (2002). Nonresponse in exit polls: A comprehensive analysis. In R.M. Groves, D.A. Dillman, J.L. Eltinge, & R.J. Little (Eds.), *Survey nonresponse* (pp. 243–257). New York, NY: Wiley.

Novo, M., Hammarström, A., & Janlert, U. (1999). Does low willingness to respond introduce a bias? Results from a socio-epidemiological study among young men and women. *International Journal of Social Welfare, 8,* 155–163.

O'Rourke, D., Chapa-Resendez, G., Hamilton, L., Lind, K., Owens, L., & Parker, V. (1998). An inquiry into declining RDD response rates. Part I: Telephone survey practices. *Survey Research: Newsletter from the Survey Research Laboratory, 29*(2). Retrieved from http://www.srl.uic.edu/Publist/Newsletter/1998/98v29n2.pdf

Page, C. (2006). *The roles of public opinion research in Canadian government*. Toronto, ON: University of Toronto Press.

Pew Research Center. (2006). *National polls not undermined by growing cell-only population: The cell phone challenge to survey research.* Retrieved from http://people-press.org/reports/pdf/276.pdf

Population Research Laboratory, Department of Sociology, University of Alberta. (1991–2007). *Alberta survey sampling report.* Edmonton, AB: Author.

Presser, S., & Stinson, L. (1998). Data collection mode and social desirability bias in self-reported religious attendance. *American Sociological Review, 63,* 137–145.

Richman, W.L., Kiesler, S., Weisband, S., & Drasgow, F. (1999). A meta-analytic study of social desirability distortion in computer-administered questionnaires, traditional questionnaires, and interviews. *Journal of Applied Psychology, 84,* 754–775.

Roeske, T.L. (2007). *The phenomenon of nonresponse in household telephone surveys: Trends and implications.* Unpublished master's thesis, University of Alberta, Edmonton, Alberta, Canada.

Singer, E. (2006). Introduction: Nonresponse bias in household surveys. *Public Opinion Quarterly, 70,* 637–645.

Smith, T.W. (1983). The hidden 25 percent: An analysis of nonresponse on the 1980 General Social Survey. *Public Opinion Quarterly, 47,* 386–404.

Smith, T.W. (1995). Trends in non-response rates. *International Journal of Public Opinion Research, 7,* 157–171.

Statistics Canada. (2006a). *Annual demographic statistics 2005*. Ottawa, ON: Statistics Canada.

Statistics Canada. (2006b, April 5). Residential telephone service. *The Daily*. Retrieved from http://www.statcan.ca/Daily/English/060405/d060405b.htm

Steeh, C., Kirgis, 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.

Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology, 55,* 775–801.

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*, 859–883.

Voogt, R.J.J., & Van Kempen, H. (2002). Nonresponse bias and stimulus effects in the Dutch National Election Study. *Quality and Quantity, 36,* 325–345.

Weisberg, H.F. (2005). *The total survey error approach: A guide to the new science of survey research*. Chicago, IL: University of Chicago Press.

Zukin, C. (2006). The future is here! Where are we now? And how do we get there? *Public Opinion Quarterly, 70,* 426–442.

Teresa L. Bladon holds an M.A. in sociology obtained from the University of Alberta in 2007. Involved in program evaluation for over seven years, she currently works as a consultant with an independent research and evaluation firm with a primary focus on human services programs.