Original Research

Sociodemographic Differences in Respondent Preferences for Survey Formats: Sampling Bias and Potential Threats to External Validity

Szu-Wei Chen, PhD, OTR/L, Marian Keglovits, OTD, MSCI, OTR/L, Megen Devine, MA, Susan Stark, PhD, OTR/L, FAOTA

School of Medicine, Program in Occupational Therapy, Washington University in St Louis, St Louis, Missouri, United States

Abstract

Objective: To explore sampling bias as a result of survey format selection by examining associations between characteristics of people aging with long-term physical disability (PAwLTPD) and their preferences for phone or web-based survey format.

Design: A cross-sectional study using a secondary data analysis approach.

Setting: Data were from an ongoing longitudinal cohort study conducted in the community.

Participants: Convenience sampling was used. PAwLTPD who participated in year 2 of the longitudinal cohort study were included. Inclusion criteria were age 45-65 years, English speaking, and self-reported onset of a physical disability at least 5 years prior to study recruitment. Two participants completed the survey using both phone and web formats and were thus excluded; 387 participants (N=387) were included in the analysis.

Interventions: Not applicable.

Main Outcome Measures: Choice of survey format and demographics (age, sex, race and ethnicity, marital status, living arrangement, socioeconomic status) were collected in addition to self-rated physical health.

Results: Participants were on average 58.2±5.6 years old. A total of 33% were male, and 62% were White. Approximately 40% of participants completed phone surveys. The phone survey...
Sampling bias can impair the external validity of a study and limit the generalizability of its findings.\(^1\) When conducting surveys, researchers often choose 1 data collection format (eg, phone or web-based survey) based on cost or other logistical considerations.\(^2\) Although the decision to use a single survey format can couple with statistical adjustments or other special methods (eg, using random digit dialing to decrease sampling bias),\(^3\) providing a survey in only 1 format can also introduce sampling bias, resulting from accessibility issues such as internet access, telephone and/or mobile phone ownership, service disruption, and participants’ physical abilities (eg, visual impairment, hearing difficulties).

Phone and web-based surveys both have the advantages of low cost and wide reach.\(^4\) Use of these survey formats has become increasingly prevalent with technology development and may have been influenced by the rise of patient-reported outcomes. Phone and web-based surveys gained further prominence as social distancing research tools during the COVID-19 pandemic, particularly among vulnerable populations such as older adults and people with disabilities. Prior studies comparing the use of different survey formats among the general public found that phone surveys had higher response rates and better representation of the study target population than web-based surveys but also demonstrated a social desirability bias,\(^4,5\) which refers to the tendency of survey respondents to provide socially desirable responses instead of responses that truly reflect their situations. Each survey format has its own strengths, and the choice of survey format is often made in consideration of other practical factors (eg, sensitivity of study topic, technology readiness of a particular target population), but researchers may consider providing multiple format options to eliminate the potential for sampling bias when it is feasible to do so.\(^6,7\)

Previous studies exploring the effect of survey format on research were conducted in the general public. To our knowledge, the potential for sampling bias relating to use of survey formats among aging populations and/or populations with disability and the survey format preferences of these populations are unknown.

This study aimed to explore how survey format selection may introduce sampling bias in studies of people aging with long-term physical disability (PAwLTPD). Specifically, we examined the associations between characteristics of PAwLTPD and their preferences for a phone or web-based survey. Findings could inform rehabilitation researchers on potential biases in sampling that survey format selection may introduce and assist them in making decisions about survey strategies.

**Methods**

**Study design and setting**

This study used existing data collected from year 2 of an ongoing 3-year longitudinal cohort study (2018-2021) investigating the trajectory of function, community participation, and use of long-term supportive services in PAwLTPD. The 3-year longitudinal study was approved by Washington University in St Louis Institutional Review Board (IRB) (IRB no.: 201710186). PAwLTPD are a group of people who have different ages of onset of their primary disabling conditions and who live with these conditions throughout the rest of their lives. The disabling conditions can begin in early stages of life (eg, cerebral palsy, muscle degeneration), midlife (eg, multiple sclerosis, spinal cord injury), or later stages of life (eg, stroke, chronic obstructive pulmonary disease). PAwLTPD often experience accelerated aging and functional declines earlier in life than the healthy aging population because of their existing physical conditions.

Participants of the cohort study were provided with information related to the longitudinal cohort study using the IRB-approved script. The informed consent was obtained before any data collection. Cohort participants completed an annual survey for 3 years via phone or internet based on their preference. The cohort was recruited through referrals from Area Agencies on Aging and Centers for Independent Living in Missouri as well as from social media. Inclusion criteria were age 45-65 years, English speaking, and self-reported onset of a physical disability at least 5 years before study recruitment. Individuals were excluded if they had a cognitive impairment that could interfere with their ability to participate.
to answer survey questions reliably. The phone survey was conducted by trained raters, and the web-based survey was sent to participants via Research Electronic Data Capture 7. The survey takes approximately 1 hour to complete by phone and contains a series of questionnaires regarding personal background, services and/or resources used, general health status, disability and comorbid conditions, activity participation and satisfaction, environmental barriers, mental health status, resilience, and social support. The reason for using year 2 survey data for analysis is because year 2 data may more accurately reflect participants’ survey format preferences with prior knowledge from year 1 regarding survey content and length. We asked participants about their preference for survey format in year 1 during recruitment, but some participants did choose to switch from phone to web-based survey for the year 2 survey calls.

Participants

Participants in the second year of the cohort study were included in the current study; 2 participants were excluded because of survey completion using both phone and web-based formats (N=387). Among this year 2 cohort of individuals with physical disabilities, the number of years living with one’s primary disabling condition ranged from 5-65 years. Approximately 44% of the cohort self-reported having neurologic-related conditions, including cerebral palsy, multiple sclerosis, spinal cord injury, polio, and stroke; 20% reported having musculoskeletal-related conditions, such as total knee replacement, arthritis, or back pain; another 20% reported having other conditions, such as respiratory-, cardiovascular-, immune system—, or genitourinary-related conditions; and the rest of participants self-identified having multiple conditions. The cohort had a diverse constitution of health conditions, so their functional levels also varied. In general, about 29% of participants reported not being able to walk 25 feet on a level surface with or without support. Over half of the participants had difficulty with activities of daily living such as showering, getting in and out of bed, bending down/picking up items from the floor, getting things from up high, or pushing open a heavy door. Approximately 50% of the participants learned about the cohort study in noninternet sources (eg, Facebook, Twitter, website, newsletter), and about 40% of participants learned about the study from noninternet sources (eg, word of mouth, call lists, program officer, flyers).

Measures

Demographic characteristics including age, sex, race and ethnicity, socioeconomic status (SES) (ie, annual personal income, education, employment status), marital status, living arrangement, and physical health were examined. Race was originally a “choose all that apply” variable and was recoded into “White” and “non-White.” Participants who indicated White as their only race were recoded as “White.” Participants who reported more than 1 race or who reported only 1 non-White race were recoded as “non-White.” Marital status was recoded as “married/long-term partnered” and “not in a long-term relationship (ie, single/divorced/separated/widowed).” Education and employment status were recoded into 3 and 4 levels, respectively (Table 1). Personal annual income was collected as a dichotomized variable using the Missouri poverty level cutoff ($10,008). Living arrangement was measured as living at one’s primary residence alone or with others. Physical health was collected using a single question asking participants to rate their overall physical health at the time of data collection on a 5-point scale ranging from “excellent” to “poor.” It was recoded into 4 levels by combining “excellent” and “very good” into 1 category because of insufficient cell counts. Higher scores indicate worse physical health.

Data analysis

Data were analyzed using SPSS. An independent t test was conducted to compare age differences between phone and web-based survey groups. Because of the ordinal scale of variables, Mann-Whitney U tests were conducted to compare differences in education and self-rated physical health between survey groups. A chi-square test of independence was conducted to explore associations between survey format preference and categorical variables (ie, nominal and ordinal variables). If the omnibus chi-square test was significant and the df was >1, post hoc tests were performed by calculating each cell’s standardized residual and its squared value (ie, squared standardized residual is equal to $X^2$) to determine cells that contributed to the significant associations found in the omnibus chi-square test. All tests that were conducted were 2-tailed with a 0.05 significance level. The effect sizes of the test results were calculated using Hedges’ g for t test with unbalanced sample sizes, r for Mann-Whitney U test, and odds ratio (OR) and Cramer’s V for chi-square test. The post hoc power analysis on the chi-

| Variable                          | Phone Survey | Web Survey | t Test or Mann-Whitney U Test | z Score | P Value |
|-----------------------------------|--------------|------------|-----------------------------|---------|---------|
| Age (y), mean ± SD                | 59.8±5.0     | 57.2±5.7   | -4.76*                      | -       | <.001†  |
| Level of education                | 149.74       | 222.63     | 11133                       | -6.65   | <.001†  |
| Self-rated physical health        | 210.06       | 183.61     | 15420                       | -2.38   | .017‡   |

* t Test result.
† P<.001.
‡ P<.05.
square test with the highest number of cells ($df=3$) showed that, given the total sample size ($N=387$), the power to detect a medium effect size ($w=0.3$) with $\alpha=0.05$ is far greater than 0.8 ($\beta=1.00$).

**Results**

The study included 387 participants. Their mean age was 58.2±5.6 years. A total of 33% were male and 62% were White. Approximately 40% of participants completed phone surveys. The phone survey group was older (59.8±5.0 years) than the web-based survey group (57.2±5.7 years) ($t=-4.76$, $P<.001$), with a small effect size ($r=0.11$). Self-rated physical health in the web-based survey group (mean rank=183.61) was statistically significantly higher than in the phone survey group (mean rank=149.74) ($U=11133$, $z=-6.65$, $P<.001$), with a small effect size ($r=0.11$). Education in the web-health for phone and web-based survey groups were not similar, as assessed by visual inspection. Education in the web-health for phone surveys, as shown by the higher observed count than non-White counterparts, the majority of whom chose phone surveys. This finding corresponds with the phenomenon of a “digital divide” among different races.

In terms of SES, we found that PAwLTPD with a bachelor’s degree or higher and those with an income above the state’s poverty threshold were more likely to choose the web-based survey. Conversely, participants with a high-school–level education or below and those with an income at or below the state’s poverty threshold were more likely to choose the phone survey. These findings are not surprising because studies of older adults have shown that social disparities closely align with tech disparities; e-literacy and one’s ability to afford digital devices could account for these differences.

Additionally, employed PAwLTPD tended to choose the web-based survey. This could be because of the length of the survey (~1 hour) in the original study; people with paid employment may have limited personal time to talk on the phone for an hour and appreciate the flexibility provided by a web-based format. Individuals on disability leave were more likely to choose the phone survey. This might be because of worsening health status or loneliness. In contrast with people who are retired, those on disability leave may have newly acquired or worsening conditions that caused them to leave the workforce. Phone surveys provide more human interaction than web-based surveys, which may be appealing to individuals adjusting to these changes (see supplemental appendix S1 for the ancillary test).

Furthermore, we found that PAwLTPD who lived alone were more likely to choose the phone survey than those who lived with others. Higher levels of loneliness among people who live alone may explain this finding; speaking with a phone surveyor could relieve loneliness (see supplemental appendix S1). Regarding relationship status, PAwLTPD who were married or in a long-term relationship tended to choose the web-based survey; this is in line with a study by Dupлага investigating internet use among people with disabilities, which found that married participants were more likely to use the internet than those who were widowed or never married. Our ancillary test on loneliness and marital status failed to explain the association we found (see supplemental appendix S1). An alternative explanation should be further explored, including investigating whether health status or patterns of time use vary among people with different marital statuses.

Previous studies have shown that older adults who report better health have greater technology use, including email,
Table 2  Characteristics of participants and relationships with survey format preference (N=387)

| Characteristic                        | Phone survey | Web survey | Omnibus or Post Hoc $\chi^2$ | df | $P$ Value | Odds Ratio or Cramer’s $V$ |
|---------------------------------------|--------------|------------|-------------------------------|----|----------|---------------------------|
|                                       | (n=152, 39.3%) | (n=235, 60.7%) |                             |    |          |                           |
|                                       | n (%)        | n (%)      |                               |    |          |                           |
| Sex                                   |              |            |                               |    |          |                           |
| Male                                  | 57 (37.5)    | 72 (30.6)  | 1.96                         | 1  | .162     | 1.36                      |
| Female                                | 95 (62.5)    | 163 (69.4) |                               |    |          |                           |
| Race/ethnicity                        |              |            |                               |    |          |                           |
| White                                 | 58 (38.2)    | 182 (77.4) | 60.69                        | 1  | <.001$^b$| 0.18                      |
| Non-White                             | 92 (60.5)    | 51 (21.7)  |                               |    |          |                           |
| Marital status                        |              |            |                               |    |          |                           |
| Currently married/long-term          | 25 (16.4)    | 115 (48.9) | 42.20                        | 1  | <.001$^b$| 0.21                      |
| partnered                             |              |            |                               |    |          |                           |
| Not in a long-term relationship      | 127 (83.6)   | 120 (51.1) |                               |    |          |                           |
| Living arrangement                   |              |            |                               |    |          |                           |
| Living alone                          | 93 (61.2)    | 71 (30.2)  | 36.26                        | 1  | <.001$^b$| 3.64                      |
| Living with others                   | 59 (38.8)    | 164 (69.8) |                               |    |          |                           |
| Personal annual income                |              |            |                               |    |          |                           |
| ≤$10,008                              | 83 (54.6)    | 44 (18.7)  |                               |    |          |                           |
| ≥$10,009                              | 69 (45.4)    | 191 (81.3) |                               |    |          |                           |
| Level of education                   |              |            |                               |    |          |                           |
| ≤High school graduation              | 74 (48.7)    | 40 (17.0)  | 48.32                        | 2  | <.001$^b$| 0.35$^a$                 |
| Some college/tech degree/associate    | 47 (30.9)    | 90 (38.3)  | 2.25$^c$                     | 1  | .13      |                           |
| degree of school                     |              |            |                               |    |          |                           |
| Bachelor’s degree/graduate school     | 53.8         | 83.2       |                               |    |          |                           |
| degree                               |              |            |                               |    |          |                           |
| Employment status                    |              |            |                               |    |          |                           |
| Paid work full-/part-time            | 13 (8.6)     | 59 (25.2)  | 17.95                        | 3  | <.001$^b$| 0.22$^a$                 |
| Expected counts                      | 28.2         | 43.8       |                               |    |          |                           |
| Seeking paid work                    | 4 (2.6)      | 9 (3.8)    | 0.36$^c$                     | 1  | .549     |                           |
| Expected counts                      | 5.1          | 7.9        |                               |    |          |                           |
| Retired/other                        | 28 (18.5)    | 39 (16.7)  | 0.25$^c$                     | 1  | .617     |                           |
| Expected counts                      | 26.3         | 40.7       |                               |    |          |                           |
| Disability leave                     | 106 (70.2)   | 127 (54.3) | 9.61$^c$                     | 1  | .002$^d$|                           |
| Expected counts                      | 91.4         | 141.6      |                               |    |          |                           |
| Self-rated physical health            |              |            |                               |    |          |                           |
| Excellent/very good                  | 21 (13.8)    | 33 (14.0)  | 0.00$^c$                     | 1  | .952     |                           |
| Expected counts                      | 21.2         | 32.8       |                               |    |          |                           |
| Good                                  | 35 (23.0)    | 84 (35.7)  | 7.02$^c$                     | 1  | .008$^d$|                           |
| Expected counts                      | 46.7         | 72.3       |                               |    |          |                           |
| Fair                                  | 61 (40.1)    | 84 (35.7)  | 0.76$^c$                     | 1  | .384     |                           |
| Expected counts                      | 57.0         | 88.1       |                               |    |          |                           |
| Poor                                  | 35 (23.0)    | 34 (14.5)  | 4.62$^c$                     | 1  | .032$^d$|                           |
| Expected counts                      | 27.1         | 41.9       |                               |    |          |                           |

NOTE. Table does not show standardized residuals because of limited space and instead shows post hoc $\chi^2$ values (ie, squared standardized residuals) to indicate significant cells. Expected counts are presented to show the direction of significant relationships.

$^1$ Post hoc $\chi^2$ value of each cell. Both cells in the same row had the same $\chi^2$ value (because df=1). $^2$ Cramer’s V effect size for $\chi^2$ cell numbers larger than 2 £ 2.

$^a$ Race had 4 missing values; employment had 1 missing value.

$^b$ $P$<.001.

$^c$ $P$<.01.

$^d$ $P$<.05.
internet, and text messages. This may explain our finding that PAwLTPD with good physical health were more likely to choose the web-based survey and those with poor physical health were more likely to choose the phone survey.

PAwLTPD are a unique population representing the intersection of aging and physical disability. This study is among the first to provide preliminary findings on how sociodemographics and physical health associate with survey format preferences in PAwLTPD. One strength of this study is that our participants self-selected their survey format. This approach provides a less biased sample of participants because no one was excluded because of difficulty with one format or the other because we have demonstrated that each survey format is likely to appeal to participants who have certain characteristics. Although sampling bias can be addressed using, for example, poststratification weighting during statistical analysis, efforts decreasing sampling bias by using better study design should not be ignored. The study findings could provide insight for planning research recruitment and retention among PAwLTPD with varying characteristics. Providing options for a survey format that participants are more comfortable using could decrease dropout rates in longitudinal studies. ¹⁷

Study limitations

A limitation of this study is that it was a cross-sectional study; therefore, causation should not be assumed. This means that inverse relationships between participant characteristics and survey preferences are also possible. This study also used a convenience sample recruited from community organizations and social media; therefore, the potential for sampling bias cannot be excluded. In addition, as an exploratory study that did not involve any decision making, we did not control the type I error inflation. Future replication studies are warranted and should consider using a multivariate approach such as logistic regression, which has a greater ability to control for type I error and covariates, as well as to clarify multicollinearity issues among predictors.

Conclusions

In summary, individuals’ selection of survey format is associated with sociodemographic characteristics. These findings provide evidence to the existing understanding of user characteristics for technologies. Findings demonstrate that sampling bias can be easily introduced in a convenience sample when using 1 survey format: offering only 1 method for survey response may result in lower participation rates of certain sociodemographic groups. When it is feasible (eg, logistics allow, measurement equivalence across different methods of instrument administration have been established), researchers should consider collecting survey data using more than 1 format to improve external validity.
Suppliers

a. REDCap Version 7; Research Electronic Data Capture hosted at Washington University in St Louis. b. SPSS 2017 version; IBM.

Corresponding author

Szu-Wei Chen, PhD, OTR/L, Washington University in St Louis, School of Medicine, Program in Occupational Therapy, 4444 Forest Park Ave, St Louis, MO 63108. E-mail address: szu-wei.chen@wustl.edu.

Acknowledgments

We thank all team members for their feedback during the revisions of the manuscript. Special thanks to Katelyn Storey for her editorial assistance in preparation for submission.

References

1. Khorsan R, Crawford C. External validity and model validity: a conceptual approach for systematic review methodology. Evid Based Complement Alternat Med 2014;2014:694804.
2. Pruchno R, Brill JE, Shands Y, et al. Convenience samples and caregiving research: how generalizable are the findings? Gerontologist 2008;28:820-7.
3. Greenacre Z. The importance of selection bias in internet surveys. Open J Stat 2016;6:397.
4. Szolnoki G, Hoffman D. Online, face-to-face and telephone surveys—comparing different sampling methods in wine consumer research. Wine Econ Policy 2013;2:57-66.
5. Cobanoglu C, Moreo PJ, Warde B. A comparison of mail, fax and web-based survey methods. Int J Mark Stud 2001;43:1-15.
6. Kays K, Gathercoal K, Buhrow W. Does survey format influence self-disclosure on sensitive question items? Comput Human Behav 2012;28:251-6.
7. Tijdens K, Steinmetz S. Is the web a promising tool for data collection in developing countries? An analysis of the sample bias of 10 web and face-to-face surveys from Africa, Asia, and South America. Int J Soc Res Methodol 2016;19:461-79.
8. Sharpe D. Chi-square test is statistically significant: Now what? Pract Assess Res Eval 2015;20:8.
9. Hartkopf AD, Graf J, Simoes E, et al. Electronic-based patient-reported outcomes: willingness, needs, and barriers in adjuvant and metastatic breast cancer patients. JMIR Cancer 2017;3:e6996.
10. Horevoorts NJ, Vissers PA, Mols F, Thong MS, van de Poll-Franse LV. Response rates for patient-reported outcomes using web-based versus paper questionnaires: comparison of two invitacional methods in older colorectal cancer patients. J Med Internet Res 2015;17:e3741.
11. Keurentjes J, Fiocco M, So-Osman C, et al. Hip and knee replacement patients prefer pen-and-paper questionnaires: implications for future patient-reported outcome measure studies. Bone Joint Res 2013;2:238-44.
12. Hunsaker A, Hargittai E. A review of internet use among older adults. New Media Soc 2018;10:3937-54.
13. Fairlie R. Race and the digital divide. Santa Cruz: Universiy of California Santa Cruz (UCSC): Department of Economics; 2014.
14. Larsson K, Wallroth V, Schröder A. You never get used to loneliness”—older adults’ experiences of loneliness when applying for going on a senior summer camp. J Gerontol Soc Work 2019;62:892-911.
15. Duplaga M. Digital divide among people with disabilities: analysis of data from a nationwide study for determinants of internet use and activities performed online. PLoS One 2017;12:e0179825.
16. Gell N, Rosenberg DE, Demiris G, LaCroix AZ, Patel KV. Patterns of technology use among older adults with and without disabilities. Gerontologist 2015;55:412-21.
17. Booker C, Harding S, Benzeval M. A systematic review of the effect of retention methods in population-based cohort studies. BMC Public Health 2011;11:1-12.