



Practice of Epidemiology

The Effect of Question Order on Self-rated General Health Status in a Multilingual Survey Context

Sunghee Lee and David Grant

Initially submitted October 16, 2008; accepted for publication March 3, 2009.

Current practices recommend placing a self-rated health question before specific health items in survey questionnaires to minimize potential order effects. Because this recommendation is based on data collected in English, its applicability to other languages is unknown. This study examines whether there is an order effect associated with self-rated health for interviews conducted in English and Spanish languages. An experiment was conducted by using the 2007 California Health Interview Survey, where questions on self-rated health were inserted in 1 of 2 locations: preceding and following question items on specific chronic conditions. Respondents were randomly assigned to 1 of 2 versions of the locations by the split-half method. Although no order effect was present in the English interviews, the authors found a significant and large effect with Spanish interviews: Self-rated health appeared much worse when asked before chronic conditions than when asked after them. This order effect was larger among females than males. Order effects for self-rated health differ by interview language; inferences about the health status of Spanish-speaking populations (and potentially Latinos) depend on question order. If maintaining comparability is important, the authors' finding contradicts current recommendations, as inserting the self-rated health question before specific questions led to larger differences in health status between English and Spanish speakers.

data collection; emigrants and immigrants; health status; health surveys; Hispanic Americans; minority health; questionnaires; research

Self-rated general health status is a strong predictor of current and subsequent mortality and morbidity even after accounting for sociodemographic and medical risk factors (1–13). It is typically assessed in a single question with a 5-point scale. Although subjective and not readily verifiable by external measures, self-rated health is sensitive and comprehensive in capturing the full spectrum of health conditions and adds an additional dimension to assessing health status than objective measures alone (14, 15), possibly reflecting undetected and unreported symptoms (16).

Because of its clinically proven utility, self-rated health is recommended by the World Health Organization (17) and is frequently used in public health research. Self-rated health is included in a wide range of data collection surveys, from health surveys (e.g., US National Health Interview Survey, US Behavioral Risk Factor Surveillance System, Canadian Community Health Survey, and Health Survey of England)

to social surveys (e.g., 2007 International Social Survey Program conducted across 41 countries) and even labor surveys (e.g., US Current Population Survey).

Despite its effectiveness and popularity, self-rated general health may be sensitive to question order, as self-reports and general questions are well known to be subject to a context effect due to question wording, format, and order (18–21). It is especially true if the general health question is preceded by specific and objective health questions, as those questions may influence respondents' answers to the general health item. Crossley and Kennedy (22) studied the reliability of the general health item by asking it twice—before and after a set of health-related questions. They found that the specific response categories chosen by respondents changed substantially but that overall estimates were not affected by alternating the location. Bowling and Windsor (23) showed that this item could potentially be subject to an order effect,

as the general health status was more positively reported when it was placed after objective health condition measures. Although the effect was small, this led the authors to suggest that the item should be placed at the beginning rather than at the end, echoing the recommendations of Keller and Ware (24), who found that general or summary questions may be influenced by preceding items if they are related. Study findings about the order effect from these studies, however, may not be free from potential confounders. Some studies asked the general health item twice. In this case, respondents may have been conditioned by their answers to the first question when answering it the second time.

As non-English-speaking and linguistically isolated populations in the United States have grown (25), so has the need to include these populations in public health research and surveillance. The self-rated general health item has been translated into many different languages as part of the validated health survey instruments SF-36 and SF-12 (multi-purpose, short-form health surveys with 36 and 12 questions, respectively), but potential order effects for this item have not been investigated in languages other than English. Moreover, because the location recommendation for the general health question was based on English language interviews, it is uncertain whether it will hold in other languages. In addition, because this item is used in health disparity research, and health disparities are highly associated with language capabilities (26–29), it is important to understand the order effect of the general health question in a multilingual survey setting. The results may inform the development of questionnaires that are more culturally and linguistically comparable as more studies collect data in languages other than just English (30).

MATERIALS AND METHODS

A randomized experiment was conducted with a subset of respondents in the 2007 California Health Interview Survey. This survey is a random-digit-dial telephone survey of California households conducted every other year since 2001 by the Center for Health Policy Research, University of California, Los Angeles. In order to represent California's racially and linguistically diverse population, the California Health Interview Survey is administered in 5 languages: English, Spanish, Chinese (Mandarin and Cantonese dialects), Korean, and Vietnamese. Non-English questionnaires are developed through cultural adaptation and multiple forward translation methods with a referee to judge the quality of the translation (31). Initial household contact was made primarily by English-only and English-Spanish bilingual interviewers. When an interviewer was unable to continue the contact because of a language problem, the case was assigned for follow-up with an interviewer able to speak the appropriate language.

Self-rated health has been the first question asked in the California Health Interview Survey following introductory demographics; it is followed by a series of items to assess the chronic health status concerning asthma, diabetes, hypertension, and heart disease. This experiment compared the results of placing the question about self-rated health after instead of before the chronic health questions (hereafter termed "after" and "before" conditions). The English

self-rated health question read, "In general, would you say your health is excellent, very good, good, fair, or poor?" The Spanish version read, "En general, ¿diría usted que su salud es excelente, muy buena, buena, regular, o mala?"

In order to minimize seasonal effects that could have confounded the experiment, we subsampled 1,191 English-speaking and 824 Spanish-speaking adults interviewed during the same time period. By using the split-half method, 574 English and 406 Spanish interviews were randomly assigned to the control group, and 617 English and 418 Spanish interviews were randomized to the experimental group. As the subjects were randomized, there were no significant differences in demographic characteristics (e.g., age, gender, educational attainment), as well as chronic conditions (e.g., asthma, diabetes, hypertension, heart problems), between the 2 location groups within each language.

The estimates of general health status were compared by language and question order. As this study is based on experimental data and because generalization about the population is not of interest, we used unweighted analyses and performed Student *t* tests for determining significance. (Note that we also conducted weighted analyses, which did not change any of the conclusions.) In order to eliminate age and gender differences between languages, we also examined age- and gender-adjusted estimates. The adjustment was done by using cross-classified distributions of age (18–35, 36–50, 51–64, and 65 years or more) and gender (male and female) of the total English sample in the experiment. The effects of language and question order on the response to the general health status question were also examined in a multivariate logistic regression controlling for the effects of ever diagnosed with chronic health conditions and demographic characteristics (age, gender, and educational attainment). We used SAS, version 9.1, software (SAS Institute, Inc., Cary, North Carolina) to analyze the data.

RESULTS

Table 1 shows the distribution of health status response by interview language and question order. Although there was little to no order effect in English interviews, there were large differences between the 2 question locations for Spanish-language interviews. Compared with the "before" condition, the response rates of "very good" and "good" health condition increased by 4.3% and 5.6%, respectively, when general health was asked about after the chronic conditions, while the estimates for "fair" and "poor" health decreased significantly (–7.5% and –4.5%, respectively).

When health status was dichotomized by combining "fair" and "poor" health into 1 category, the question-order effect in Spanish was even more apparent. The fair/poor estimate decreased from 45.3% to 33.3% when the item was placed after that on chronic conditions. This implies that nearly 1 out of 2 adults interviewed in Spanish reported fair/poor health when the general health question was asked before the chronic conditions, while only 1 out of 3 reported fair/poor health when the item was asked after those conditions. We also examined the fair/poor health estimates among Latinos interviewed in English. The order effect in this group was in the same direction

Table 1. Distribution of Health Status by Interview Language and Question Order, California Health Interview Survey, 2007

General Health	English Interviews					Spanish Interviews				
	Before (n = 574)		After (n = 617)		Before – After	Before (n = 406)		After (n = 418)		Before – After
	%	SE	%	SE		%	SE	%	SE	
Excellent	21.8	1.7	22.2	1.7	–0.4	6.7	1.2	8.9	1.4	–2.2
Very good	36.1	2.0	32.9	1.9	3.2	8.1	1.4	12.4	1.6	–4.3*
Good	26.5	1.8	27.1	1.8	–0.6	39.9	2.4	45.5	2.4	–5.6
Fair	12.2	1.4	13.5	1.4	–1.3	37.9	2.4	30.4	2.3	7.5*
Poor	3.5	0.8	4.4	0.8	–0.9	7.4	1.3	2.9	0.8	4.5**
Fair/poor										
Unadjusted	15.7	1.5	17.8	1.5	–2.2	45.3	2.5	33.2	2.3	12.1***
Age and gender adjusted	16.1	1.6	17.6	1.5	–1.5	54.5	2.8	38.1	2.6	16.4***

Abbreviation: SE, standard error.

* $P < 0.05$; ** $P < 0.01$; *** $P < 0.001$.

as that of the Spanish-language interviews and was significant (19.8% ($n = 121$) vs. 10.5% ($n = 133$) in reporting fair/poor health, with $P = 0.038$).

When age and gender adjustments were used, the order effect in Spanish appeared even larger, with 16.4 percentage points fewer people reporting fair/poor health under the “after” condition. For English, there was no change after the age and gender adjustment. Regardless of controlling for age and gender, the degree of potential disparity in responses about general health between English- and Spanish-speaking adults could differ dramatically depending on the location of the general health item in the questionnaire: Compared with English speakers, Spanish-speaking adults might report fair/poor health status 2 (“after” condition) or 3 (“before” condition) times as frequently.

We tested the effect of question order for predicting fair/poor health when controlling for demographic characteristics and chronic conditions using stratified logistic regression by interview language. We fit the same model to the

English and Spanish speaker samples separately. Not surprisingly, the data showed that diagnoses of all chronic conditions were significantly associated with reporting fair/poor health for both interview language groups except for asthma among Spanish speakers (Table 2). Age and male gender were not significant predictors of fair/poor health for the English-language interview group but were significant for the Spanish interview group, where older and female Spanish-speaking people were more likely to report fair/poor health than did younger and male Spanish-speaking people. The results indicated that the effect of question order was significant only for the Spanish interview group: Placing the general health item before chronic conditions produced higher response rates for fair/poor health.

We further examined the effect of question order on self-rated general health by age and gender in Figure 1. First, the estimate of fair/poor health was examined by language, question location, and age. There were no significant differences in reported health status by question location among those

Table 2. Stratified Logistic Regressions Predicting Fair/Poor Health With Demographic Characteristics, Chronic Conditions, and Question Order by Interview Language, California Health Interview Survey, 2007

Predictors	English Interviews		Spanish Interviews	
	Odds Ratio	P Value	Odds Ratio	P Value
Intercept	0.127	0.000	0.244	0.000
Age, years	0.998	0.697	1.013	0.032
Male	1.099	0.618	0.659	0.008
Some college or more	0.538	0.001	0.504	0.004
Asthma, ever	2.303	0.000	1.610	0.105
Diabetes, ever	3.998	0.000	3.196	0.000
Hypertension, ever	2.077	0.001	2.114	0.000
Heart diseases, ever	3.076	0.000	2.272	0.010
Question placed before chronic conditions	1.039	0.837	1.696	0.001

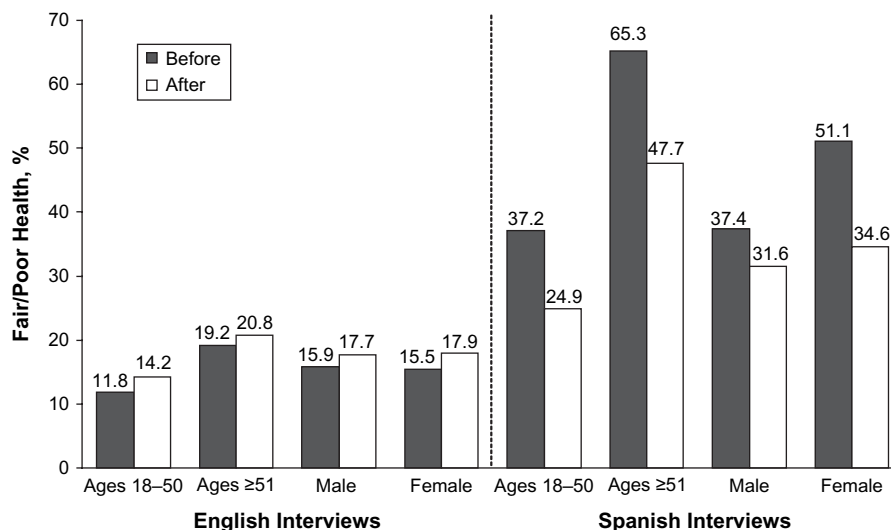


Figure 1. Percentage of fair/poor health report by interview language, question order, age, and gender, California Health Interview Survey, 2007. Significant differences by question order are shown for ages 18–50 years, ages ≥51 years, and females interviewed in Spanish ($P < 0.01$).

aged 18–50 years or 51 or more years for the English-language interviews. Spanish-language interviews, however, showed significant differences for both age groups in the same direction, with reports by older Spanish speakers showing a slightly larger effect (12.3 percentage point difference among those aged 18–50 years and 17.6 percentage point difference among those aged 51 or more years). When examining by gender, a striking trend emerged. In English, there was no difference in question order effect by gender: Males and females reported fair/poor health at almost identical proportions. However, when the Spanish interview sample was examined, the question-order effect was not significant among male but was large and significant for females. When the self-rated health question was asked after the chronic health condition questions, reports of fair/poor health decreased by 16.5 percentage points among Spanish-speaking females.

DISCUSSION

This study examined question order as an aspect of survey instrument design to address cultural equivalence among a linguistically diverse population. Question order is only 1 aspect of data collection methods and has received scant attention relative to other issues, such as questionnaire translation. This study showed some potentially large reporting differences due to the effect of question order on estimates of general health. This finding demonstrates that “cultural equivalence” is more than translational equivalence and that efforts to improve the comparability of measures across a diverse population need to consider data collection from a more holistic vantage point.

Assumptions about questionnaire design conventions in 1 language may not necessarily hold in other languages. In our study, the absence of a question-order effect in English in-

terviews was not echoed in Spanish interviews. The conventional placement of self-rated general health before objective health-related question items may be a valid practice to minimize order effect for English-language data collection. However, experimental data from the 2007 California Health Interview Survey showed that this practice did not minimize, but rather profoundly increased, the order effect for Spanish speakers (and Latinos to a certain extent) when answering the general health question. It is quite reasonable that some people in linguistic and cultural minority groups may not be familiar with the task of evaluating their general health when asked without any cues or a frame of reference. The objective health condition questions, when preceding the general health question, may serve as stimuli for these respondents, helping them to summarize and contextualize their subsequent report of general health status. For instance, Finch et al. (32) found limited predictability of self-reported health for mortality risk but an increased association between the 2 with an increased acculturation level among Latino populations in the United States. This finding may be related to the potential difficulty we hypothesize that Latinos encounter when evaluating their general health. Alternatively, some cultural factors, such as aspects considered when assessing one’s health or the level of willingness to provide personal information (33), apply differently for Latinos, depending on the sequence of the questions, unlike non-Latino English-speaking populations.

Our study also found that a slight variation in questionnaire design may change inferences dramatically. When placing the general health question after chronic condition questions, the fair/poor health estimate decreased in Spanish interviews, and the difference in fair/poor health rates between English and Spanish interviews decreased substantially and significantly. Therefore, depending on the location of the item, inferences not only about the general

health of the Spanish-speaking population and potentially Latinos more generally but also about health disparities could fluctuate considerably.

Given the importance and utility of the self-rated general health item and growing interest in understanding minority populations, we suggest 2 areas for further research. First, we do not know what types of order effect may exist in languages other than English and Spanish. Our findings may be generalizable only to English- and Spanish-speaking Californians. Although Spanish speakers account for a large proportion of the non-English-speaking population, the finding may not apply to other non-English speakers. Second, survey interview language contains more information than just language. It may reflect one's race, ethnicity, culture, attributes affected by these, and interactions among them. To disentangle these elements, there is a need for carefully designed experimental studies, prepared with in-depth knowledge about the characteristics of the study population and with longitudinal components.

ACKNOWLEDGMENTS

Author affiliations: Center for Health Policy Research, School of Public Health, University of California, Los Angeles, California (Sunghee Lee, David Grant); and Department of Biostatistics, School of Public Health, University of California, Los Angeles, California (Sunghee Lee).

This work was partially funded by the California Department of Public Health, the California Department of Health Care Services, the California Endowment, and the National Cancer Institute.

The authors thank the California Health Interview Survey for generous support.

Conflict of interest: none declared.

REFERENCES

- Bergner M, Rothman ML. Health status measures: an overview and guide for selection. *Annu Rev Public Health*. 1987;8:191–210.
- Idler EL, Angel RJ. Self-rated health and mortality in the NHANES-I Epidemiologic Follow-up Study. *Am J Public Health*. 1990;80(4):446–452.
- Pijls LT, Feskens EJ, Kromhout D. Self-rated health, mortality, and chronic diseases in elderly men: the Zutphen Study, 1985–1990. *Am J Epidemiol*. 1993;138(10):840–848.
- Hays JC, Schoenfeld DE, Blazer DG. Determinants of poor self-rated health in late life. *Am J Geriatr Psychiatry*. 1996;4(3):188–196.
- Idler EL, Benyamini Y. Self-rated health and mortality: a review of twenty-seven community studies. *J Health Soc Behav*. 1997;38(1):21–37.
- Benyamini Y, Idler EL. Community studies reporting association between self-rated health and mortality: additional studies, 1995 to 1998. *Res Aging*. 1999;21(3):392–401.
- McGee DL, Liao Y, Cao G, et al. Self-reported health status and mortality in a multiethnic US cohort. *Am J Epidemiol*. 1999;149(1):41–46.
- Idler EL, Russell LB, Davis D. Survival, functional limitations, and self-rated health in the NHANES I Epidemiologic Follow-up Study, 1992. First National Health and Nutrition Examination Survey. *Am J Epidemiol*. 2000;152(9):874–883.
- Franks P, Gold MR, Fiscella K. Sociodemographics, self-rated health, and mortality in the U.S. *Soc Sci Med*. 2003;56(12):2505–2514.
- Huang GH, Palta M, Allen C, et al. Self-rated health among young people with type 1 diabetes in relation to risk factors in a longitudinal study. *Am J Epidemiol*. 2004;159(4):364–372.
- DeSalvo KB, Fan VS, McDonell MB, et al. Predicting mortality and healthcare utilization with a single question. *Health Serv Res*. 2005;40(4):1234–1246.
- Vuorisalmi M, Lintonen T, Jylhä M. Global self-rated health data from a longitudinal study predicted mortality better than comparative self-rated health in old age. *J Clin Epidemiol*. 2005;58(7):680–687.
- Jylhä M, Volpato S, Guralnik JM. Self-rated health showed a graded association with frequently used biomarkers in a large population sample. *J Clin Epidemiol*. 2006;59(5):465–471.
- Mossey JM, Shapiro E. Self-rated health: a predictor of mortality among the elderly. *Am J Public Health*. 1982;72(8):800–808.
- Maddox GL. Some correlates of differences in self-assessment of health status among the elderly. *J Gerontol*. 1962;17(2):180–185.
- Verbrugge LM, Ascione FJ. Exploring the iceberg: common symptoms and how people care for them. *Med Care*. 1987;25(6):539–569.
- De Bruin A, Picavet HJS, Nossikov A. *Health Interview Surveys. Towards International Harmonization of Methods and Instruments*. Regional Publications European Series No. 58. Geneva, Switzerland: World Health Organization; 1996.
- Payne SL. *The Art of Asking Questions*. Princeton, NJ: Princeton University Press; 1951.
- Schwarz N, Strack F, Mai H. Assimilation and contrast effects in part-whole question sequences: a conversational logic analysis. *Public Opin Q*. 1991;55(1):3–23.
- Tourangeau R, Rasinski K, Bradburn N. Measuring happiness in surveys: a test of the subtraction hypothesis. *Public Opin Q*. 1991;55(2):255–266.
- Schwarz N. Self-reports: how the questions shape the answers. *Am Psychol*. 1999;54(2):93–105.
- Crossley TF, Kennedy S. The reliability of self-assessed health status. *J Health Econ*. 2002;21(4):643–658.
- Bowling A, Windsor J. The effects of questions order and response-choice on self-rated health status in the English Longitudinal Study of Aging (ELSA). *J Epidemiol Community Health*. 2008;62(1):81–85.
- Keller SD, Ware JE. Questions and answers about SF-36 and SF-12. *Med Outcomes Trust Bull*. 1996;4:3.
- Bureau of the Census, US Department of Commerce. Language use and English-speaking ability: 2000. Census 2000 brief. Washington, DC: Bureau of the Census; 2003. (<http://www.census.gov/prod/2003pubs/c2kbr-29.pdf>). (Accessed June 2, 2008).
- Carrasquillo O, Orav JE, Brennan TA, et al. Impact of language barriers on patient satisfaction in an emergency department. *J Gen Intern Med*. 1999;14(2):82–87.
- Jacobs EA, Lauderdale DS, Meltzer D, et al. Impact of interpreter services on delivery of health care to limited-English-proficient patients. *J Gen Intern Med*. 2001;16(7):468–474.

28. Institute of Medicine. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care*. Washington, DC: The National Academy Press; 2003.
29. Yeo S. Language barriers and access to care. *Annu Rev Nurs Res*. 2004;22(May):59–73.
30. Lee S, Nguyen HA, Jawad M, et al. Linguistic minorities in a health survey. *Public Opin Q*. 2008;72(3):470–486.
31. Ponce NA, Lavarreda SA, Yen W, et al. The California Health Interview Survey 2001: translation of a major survey for California's multiethnic population. *Public Health Rep*. 2004; 119(4):388–395.
32. Finch BK, Hummer RA, Reindl M, et al. Validity of self-rated health among Latino(a)s. *Am J Epidemiol*. 2002;155(8): 755–759.
33. Vuorisalmi M, Pietilä I, Pohjolainen P, et al. Comparison of self-rated health in older people of St. Petersburg, Russia, and Tampere, Finland: how sensitive is SRH to cross-cultural factors? *Eur J Ageing*. 2008;5(4):327–334.