

RESEARCH SYNTHESIS

RESEARCH IN AND PROSPECTS FOR THE MEASUREMENT OF HEALTH USING SELF-RATED HEALTH

DANA GARBARSKI*

Abstract Self-rated health (SRH)—for example, “in general would you say your health is excellent, very good, good, fair, or poor?”—is the most widely used measure of health across a range of survey research studies. This paper synthesizes extant research and provides a framework for future research on the measurement of health using SRH, focusing on four interrelated topics: the factors that influence respondents’ health ratings, the survey measurement features of SRH, how SRH answers are analyzed, and the stated purpose of SRH as a proxy for more objective health or as a perception of health.

Extant research on the health, psychological, and social factors influencing respondents’ SRH answers is reviewed, as is research concerned with the survey measurement features of SRH that influence how respondents rate their health. The synthesis proposes a framework for future research that focuses on further explicating the factors that underlie respondents’ SRH answers and improving features of SRH measurement and analysis in ways that are consistent with the various goals of the researchers who both collect and analyze the data.

Introduction

Self-rated health (SRH) is the most widely used measure of health in medical, social, and behavioral science research using survey data. SRH is also used as

DANA GARBARSKI is an assistant professor in the Department of Sociology at Loyola University Chicago, Chicago, IL, USA. The author thanks Jennifer Dykema and Nora Cate Schaeffer for comments on previous versions of this paper as well as years of collaboration; Helena Dagadu, Kristen Olson, the editors, and the anonymous reviewers for helpful feedback on previous versions of this paper; and Ashley Baber and Bill Byrnes for research assistance. The opinions expressed herein are those of the author. This work was supported by funding from the Eunice Kennedy Shriver National Institute of Child Health and Human Development [T32 HD049302 to Gloria Sarto] and from core funding to the Center for Demography and Ecology [P2C HD047873] and Center for Demography of Health and Aging [P30 AG017266] at the University of Wisconsin–Madison. *Address correspondence to Dana Garbarski, Department of Sociology, Loyola University Chicago, 1032W. Sheridan Road, 440 Coffey Hall, Chicago, IL 60660, USA; e-mail: dgarbarski@luc.edu.

doi:doi:10.1093/poq/nfw033

© The Author 2016. Published by Oxford University Press on behalf of the American Association for Public Opinion Research. All rights reserved. For permissions, please e-mail: journals.permissions@oup.com

a summary indicator to monitor the health of populations (OECD 2015) and patients in clinical settings (Mavaddat et al. 2014). Its popularity stems in part from its association with multiple domains of health and subsequent mortality (Idler and Benyamini 1997; Jylhä 2009; DeSalvo et al. 2006). As Idler and Benyamini (1997, 31) noted two decades ago, “a very long list of variables is required to explain the effect of one brief four- or five-point scale item.” Overall, the measurement of SRH, its use in analysis, and understanding how respondents rate their health when SRH is asked have important implications for research, policy, and clinical practice.

The lack of specificity in how the SRH question is asked gives rise to an inherent tension as the measure’s most important benefit and drawback: Respondents can make a comprehensive summary of their health, but researchers have little control over what respondents consider when rating their health. Furthermore, while researchers agree that SRH answers are ratings of respondents’ health that integrate and summarize across the health domains salient to respondents, at least two conceptualizations of SRH exist in terms of how researchers *use* respondents’ SRH answers: as proxies for respondents’ more objective health characteristics and as perceptions of health. In research concerned with the causes and consequences of health, SRH is overwhelmingly and implicitly used as a proxy for more objective health,¹ emphasizing the “health” in SRH. Yet, others have argued that SRH should be conceptualized by researchers as perceptions of health, constructed through psychological filters that are themselves part of the assessment rather than a nuisance factor to be controlled—that is, emphasizing the “self-rated” (Herrmann et al. 1994; Jylhä 2009, 2010; Huisman and Deeg 2010). While these two conceptualizations of SRH answers are complementary and highlight the “dual nature” of the item (Jylhä 2009), they emphasize different facets of SRH, with varying implications for understanding what underlies SRH answers and how SRH is—or should be—measured in surveys and analyzed by researchers.

This paper synthesizes extant research on the health, psychological, social, and survey measurement factors that influence respondents’ ratings of their health when SRH is asked at a single time point and integrates this research into a conceptual model. This synthesis proposes a framework for future research, focusing on further explicating the factors that underlie respondents’ SRH answers and improving features of the measurement and analysis of SRH. The tension between the stated purposes of SRH (as a proxy for more objective health or perception of health) is addressed throughout, as this tension often has implications for the understanding, measurement, and analysis of SRH.

1. Objective health can include several measures, from self-reports of health behaviors and conditions to assessments from medical professionals to biomarkers of stress to mortality. Each of these measures is more objective than self-rated health because they each specify a particular domain of health. However, each of these more objective measures is also subject to errors in measurement. Thus, “objective health” is qualified with “more.”

Review of the Factors That Influence Respondents’ SRH Answers

Figure 1 summarizes extant research on the factors that influence respondents’ SRH answers. Examples of the most commonly studied factors are grouped based on the broader dimensions that influence SRH answers: health, psychological, social, and survey measurement. The model highlights the interplay among the factors that influence respondents’ SRH answers.

HEALTH

Existing studies seeking to describe what underlies respondents’ SRH answers focus on the different types of health factors respondents presumably consider when rating their health, such as health conditions, health behaviors, physical functioning, health care, and health knowledge. The health dimension (figure 1) contains several factors under the heading *respondent’s health factors* to signal that respondents consider their own health factors in forming an assessment of their health. The heading *health environment* incorporates the observation that respondents make judgments about their own health based on what they observe about health in the world around them: Whether and how one considers certain health factors as part of their health assessment

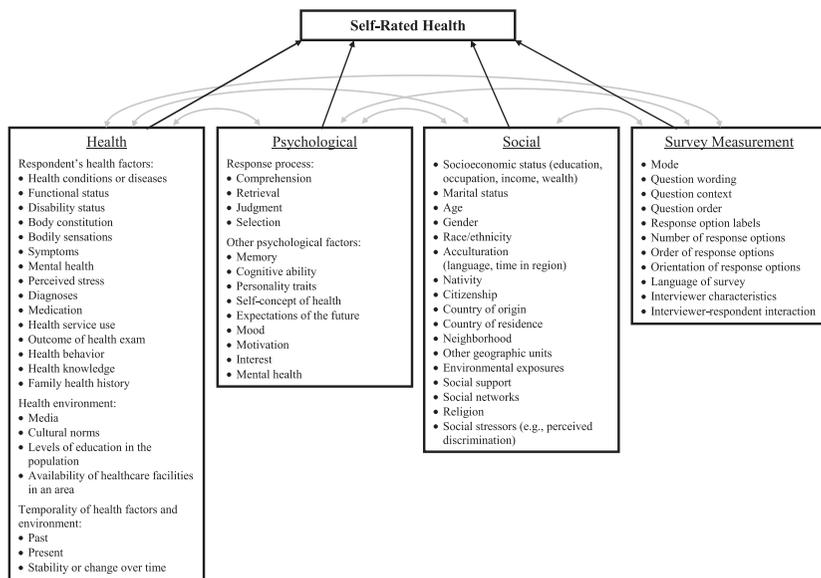


Figure 1. Model of the Factors Influencing Respondents’ Self-Rated Health Answers.

(or as positive or negative contributions) is influenced by features of the health environment such as the media, cultural norms, levels of education in the population, and availability of health care facilities in an area (Sen 1993, 2002). The heading *temporality of health factors and environment* signals that these aspects can refer to the past, present, and stability or changes over time (Schulster 1994; Idler and Benyamini 1997).

Existing studies have taken both quantitative and qualitative approaches to examine the range of health factors respondents consider when rating their health. Quantitative studies examine the association between SRH answers and various health factors. Cross-sectional studies that examine the association between SRH and other concurrent health factors have shown a range of negative and positive physical and mental health factors associated with SRH, such as health conditions (in the past, chronic, or recent), medication use, functioning, activity, health behaviors, and negative and positive affect (Benyamini, Leventhal, and Leventhal 1999; Benyamini et al. 2000; Singh-Manoux et al. 2006). In these studies, the inference is made that the health factors that are more strongly associated with SRH answers are weighed more heavily by respondents when constructing their answers; some studies have respondents explicitly rate how important a series of health factors was to the respondent when they were rating their health (Benyamini, Leventhal, and Leventhal 1999, 2003).

Longitudinal studies that examine the association between SRH and future health outcomes have shown that SRH predicts changes in functional ability (Idler and Kasl 1995) and subsequent mortality (Idler and Benyamini 1997; Benyamini, Leventhal, and Leventhal 1999). In these and other longitudinal studies, subsequent morbidity and mortality are used as criteria to examine the predictive validity of SRH answers across a range of respondents' social and health characteristics, exploring whether, when, and for whom SRH predicts subsequent morbidity and mortality (further addressed below under Social). For example, SRH is a good predictor of subsequent mortality among respondents with experiential knowledge of an illness (having a circulatory system disease) but not for respondents with no diagnosed disorder, supporting the notion that respondents with experiential knowledge of their health are better able to rate their health in a way that is predictive of subsequent mortality (Idler et al. 2004).

Qualitative studies ascertain the types of health factors respondents take into account by asking respondents to describe what they were thinking about when they rated their health. Some of these studies use cognitive interviewing techniques, in which respondents rate their health then answer probes that elicit descriptions of what they considered (Groves, Fultz, and Martin 1992; Krause and Jay 1994; Canfield et al. 2003; Kaplan and Baron-Epel 2003; Miller et al. 2005; Garbarski et al. 2015). Other studies use semi-structured interviewing protocols (Manderbacka 1998; Simon et al. 2005). The results of the qualitative studies describe a range of health factors respondents reported

considering while rating their health, differences across groups (e.g., gender, race/ethnicity, age, and education) in whether and how often certain health factors were reported, and how respondents integrate multiple and sometimes disparate domains of health when formulating a rating.

Overall, quantitative and qualitative studies have uncovered the health factors respondents consider when answering SRH through a *process of association*: either quantifying the association between SRH answers and other health factors or asking respondents which health factors they considered when formulating their answer. The main difference between these types of studies is that the health factors considered are reported by respondents for the latter (although they are coded and categorized by the researchers), whereas in the former, the health-related factors are selected by researchers *a priori* and presented to respondents in a survey.

PSYCHOLOGICAL

Respondents do not simply formulate a mental list of the health factors that pertain to them. Rather, several psychological factors—cognitive and affective—may influence how those health factors are retrieved and synthesized to formulate an assessment of one's health (figure 1).

The literature on cognitive aspects of survey methodology (CASM) highlights the response process respondents undertake to answer survey questions: comprehension of the question, retrieval of relevant information from memory to answer the question, use of retrieved information to make judgments, and selection and reporting of an answer (Tourangeau, Rips, and Rasinski 2000; see also Jylhä [2009]; Lee [2014] for integration of this model with respect to SRH). The subjectivity of this response process becomes apparent when thinking through the cognitive steps, strategies, and heuristics respondents may use to rate their health. These include cognitive processes such as defining what constitutes, for example, “health” or “excellent health”; drawing on aspects of health that are most salient in memory (Knäuper and Turner 2003); sampling from the pool of health-relevant considerations (Tourangeau, Rips, and Rasinski 2000; Lee 2014); comparisons to one's past health or others (which of course depends on the social aspect of whom one is around) (Idler 1994; Schulster 1994; Singer 1994); evaluating the importance of health-relevant information; weighting or ranking health information; integrating information into a summary; and anchoring to the first response option then adjusting to arrive at the final answer (Tversky and Kahneman 1974).

Furthermore, how the process of rating one's health unfolds depends on *other psychological factors*. As noted above, one's memory has implications for the health factors that are considered. One's cognitive ability is related to SRH answers, although much of this association can be explained by adult socioeconomic outcomes such as education and income (Schnittker 2005a). In addition, dispositional or personality characteristics are included in the

model: negative affect, optimism, need for cognition, need to evaluate, and features of psychological well-being such as autonomy, environmental mastery, personal growth, positive relations with others, purpose in life, and self-acceptance. These are often invoked as sources of methodological artifacts in subjective ratings, although some studies examine the association between personality characteristics and concurrent or subsequent SRH as a substantive research question (Benyamini, Leventhal, and Leventhal 1999; Aiken-Morgan et al. 2014; Letzring, Edmonds, and Hampson 2014; Ryff, Radler, and Friedman 2015). Also included in the model are one's self-concept of health, which includes enduring and established beliefs about one's health identity (Idler 1994; Schulster 1994; Schaeffer 2000; Bailis, Segall, and Chipperfield 2003; Lee 2014; Brenner and DeLamater 2016); one's expectations about the future, which can refer to life chances in general or expectations about health in particular (Schulster 1994; Hitlin, Erickson, and Brown 2015; Hitlin and Johnson 2015); and other psychological factors that may affect respondents' performance in the survey interview, such as their mood, motivation to complete the task, or interest in the survey. The model also includes factors that may be considered the respondent's health factors, such as mental health and perceived stress; here, the model demonstrates that some factors do not fit neatly within just one dimension.

Thus, what underlies SRH answers includes the health factors respondents consider when rating their health, the psychological processes respondents go through to choose and synthesize that information and formulate a response to the question, and the psychological characteristics of respondents that may influence the process of rating their health. In addition, the conceptual model highlights the interplay among the various factors that influence SRH answers. For example, the stages of the response process may be truncated, elongated, or otherwise influenced by the other factors in the model. Similarly, health factors may influence the psychological states that frame evaluations of health, such as changes in one's health catalyzing changes in the internal standards of measurement, values, or definition that one uses to rate their health as described in models of response shift (Sprangers and Schwartz 1999; Schwartz et al. 2007).

SOCIAL

The third dimension in the conceptual model is *social*, which includes several social factors that are associated with health, associated with differences in evaluative frameworks across social groups,² or both (figure 1). Evaluative frameworks can be broadly defined as the process through which respondents rate their health; thus, *differences in evaluative frameworks*

2. "Groups" in this context refers to groups defined by social characteristics, including, for example, race, ethnicity, language spoken, gender, indicators of socioeconomic status, marital status, age, and geographic location.

describe differences across groups in, for example, the types and scope of health factors considered, the reference groups used in comparisons, definitions of health, interpretation and use of the response scale, and interpretation of other aspects of the survey question (Jylhä 2009). Aspects of this phenomenon have also been referred to as “reporting differences” (Burgard and Chen 2014), “reporting heterogeneity” (Lindeboom and van Doorslaer 2004; Dowd and Todd 2011; Dowd 2012), or “differential item functioning” (Grol-Prokopczyk, Freese, and Hauser 2011). Evidence for differences in evaluative frameworks across social groups is found in studies that highlight differences in SRH across groups among individuals that are otherwise similarly situated with respect to health factors, or differences in health outcomes (including mortality) across groups among individuals with similar SRH ratings. These differences have been found across groups defined by race and ethnicity (Ferraro and Kelley-Moore 2001; Assari, Lankarani, and Burgard 2016); indicators of acculturation such as language spoken at home, language spoken in the interview, and time in the United States (Finch et al. 2002; Bzostek, Goldman, and Pebley 2007); gender (Benyamini, Leventhal, and Leventhal 2000; Benjamins et al. 2004); indicators of socioeconomic status such as education, occupation, and income (Dowd and Zajacova 2007, 2010; Layes, Asada, and Kephart 2012); marital status (Zheng and Thomas 2013); age (Idler 1993; Lindeboom and van Doorslaer 2004; Schnitker 2005b); and features of geographic location (Jürges 2007; Quesnel-Vallée 2007; Bjornstrom and Kuhl 2014).

In this quantitative research, differences in evaluative frameworks are assessed indirectly by examining how the association between SRH and health factors or outcomes varies across groups, in conjunction with the interpretation that such differences across groups may arise because evaluative frameworks are similar within groups and differ across groups. The assessment of differences in evaluative frameworks could be made somewhat more directly in qualitative studies in which respondents are asked why they rated their health as they did, but only if respondents discuss these factors (e.g., their definition of health, their interpretation of the response options, the health factors they considered) and researchers then examine differences across groups (Garbarski et al. 2015). The notion of the evaluative frameworks respondents use to evaluate and rate their health moves beyond *which* health factors respondents consider to *how* those health factors and the SRH question itself are experienced, conceptualized, interpreted, and integrated, and adds that the process of “how” may vary systematically across groups defined by social characteristics (see also Jylhä [2009, 2010]), represented in the conceptual model by the interplay among the first three dimensions.

By considering the first three dimensions of the conceptual model, we can begin to see how the stated purpose of SRH has implications for the analysis of respondents’ SRH answers. If systematic differences in evaluative frameworks across groups exist, using SRH to examine health disparities (differences in

health across groups) is problematic if the purpose of SRH is to be a proxy for more objective health (that is, emphasizing the “health” in SRH). In this instance, the goal is an understanding of group differences in objective health factors, and differences in evaluative frameworks can be viewed as compromising group comparisons, thus reducing the validity of SRH for this purpose.³ For example, if blacks and Latinos in the United States are more likely to select response options at the extremes of the response scale than are non-Hispanic whites (Warnecke et al. 1997; Johnson, Shavitt, and Holbrook 2011), estimates of differences or similarities in SRH across these two groups will conflate the more objective health factors that underlie respondents’ assessment of their health and group differences in this particular type of response style (as well as other differences in evaluative frameworks). However, differences across groups in evaluative frameworks are not necessarily problematic if the stated purpose of the research is to compare perceptions of health (emphasizing the “self-rated” in SRH). “Perception” speaks to both the health factors being considered and the ways in which those pieces of information are integrated to answer the question. Thus, the frameworks used to evaluate and rate one’s health are integral to the target concept, and differences in evaluative frameworks across groups do not necessarily reduce the validity of SRH for this purpose.

SURVEY MEASUREMENT

Another dimension that shapes respondents’ SRH answers highlights that how surveys measure—observe and record—the concept of health has implications for how respondents rate their health and the analytic potential of these ratings. Here, I define *survey measurement* as the interrelated features of the design and implementation of surveys—such as mode, question context, question order, response option labels, response option order, and so forth—that may influence respondents’ answers to survey questions. It is important to understand whether and how these features of survey measurement influence SRH answers because survey measurement is often outside the control of researchers who are analyzing previously collected data. In addition, features of survey measurement often vary across studies that researchers would like to compare with each other. In other words, potential implications exist for cross-survey (across studies and over time) comparability in estimates of SRH answers. The survey measurement features that have received the most attention with respect to SRH are the order of response options and the context established by question order. Less attention has been paid to the impact on SRH of survey measurement features such as mode, question wording, labels for response options, number of response options

3. I take the stance that no single objective measure of “true” health could serve as a criterion for SRH (Herrmann et al. 1994; Jylhä 2009, 2010; Huisman and Deeg 2010). “Validity” with respect to SRH thus depends on its stated purpose.

used, response scale orientation in visual modes, language of the interview, and interviewer characteristics and interviewer-respondent interaction in interviewer-administered modes.

Most surveys order the SRH response options from the positive to the negative end of the scale (e.g., “excellent” to “poor”) regardless of mode. Garbarski, Schaeffer, and Dykema (2015a, 2016) find in a web survey experiment that mean SRH is higher (e.g., better health) and the proportion in “fair” or “poor” health lower when the response options are ordered from positive to negative. One interpretation is that ordering the SRH response options from “poor” to “excellent” appears to increase the likelihood that respondents consider some of the less desirable response options when assessing their health rather than choosing the first response option perceived to be an acceptable answer (Krosnick and Alwin 1987; Krosnick 1991; Garbarski, Schaeffer, and Dykema 2015a, 2016). Further, the concurrent validity of SRH is better—the association between SRH answers and medical plan visits is significant—when the SRH response options are ordered from negative to positive (Means et al. 1989), although replication is needed with other criteria and larger samples. While beginning with the less desirable end of the scale has been suggested to expand the range of response options that respondents consider (Sudman and Bradburn 1982), more research is needed—examining the mode of administration and whether ordering response options from negative to positive violates conversational norms (Holbrook et al. 2000)—in order to support a recommendation to do so for SRH (Garbarski, Schaeffer, and Dykema 2015a).

Common wisdom dictates that SRH, as a general health question, be asked before questions about more specific aspects of health so that respondents’ answers to these domain-specific health items do not affect their SRH answers. Thus, the position of SRH in the survey instrument relative to other health items may be consequential for respondents’ answers and the validity of SRH. Two mechanisms survey researchers use to describe the effects of question context established by question order are assimilation effects and contrast effects. Assimilation effects would lead to *larger* associations between SRH and other domain-specific health items when SRH is administered after those health items compared to when SRH is administered before (Schwarz, Strack, and Mai 1991; Schwarz and Bless 1992; Tourangeau, Rips, and Rasinski 2000; Garbarski, Schaeffer, and Dykema 2015a). In general, a change in the order of questions could lead to an increased association between the items if the change leads to a reduction in random error even though the underlying association is unaffected or leads to an increase in the underlying association. Several interrelated processes could elicit these effects (Garbarski, Schaeffer, and Dykema 2015a): 1) if the order of questions conveys to respondents that SRH should summarize the information they provided in their answers to the preceding specific health items (Schwarz, Strack, and Mai 1991; Schwarz and Bless 1992); 2) if the preceding specific health questions provide a similar

definition of health or the response option scale for respondents that they use when rating their health (Hopkins and King 2010; Lee and Schwarz 2014); or 3) if the preceding specific health questions initiate for respondents a similar memory structure of beliefs, evaluations, and feelings about health that become salient when answering SRH (Tourangeau, Rips, and Rasinski 2000; Lee 2014). Contrast effects would lead to *smaller* associations between SRH and other health items when SRH is administered after those items compared to before; this could occur if respondents infer that SRH is asking about something different from the preceding specific health items (Schwarz, Strack, and Mai 1991; Toureangeau, Rasinski, and Bradburn 1991; Schwarz and Bless 1992; Tourangeau, Rips and Rasinski 2000).

A few studies examine how the placement of SRH with respect to specific questions about health affects the distribution of SRH answers—with varying results—but do not examine associations between SRH and the other health items, and thus do not indicate whether placing SRH after other health items elicits assimilation, contrast, or no effects (Crossley and Kennedy 2002; Bowling and Windsor 2008; Lee and Grant 2009). However, consistent with assimilation effects, Garbarski, Schaeffer, and Dykema (2015a) demonstrate that the associations between SRH answers and other health items are larger when SRH is administered after the health items than when SRH is administered before.

Preceding SRH with other health items may diminish differences in how SRH is interpreted by providing a common referent for respondents, leading to more comparable estimates of SRH across groups. This approach has been used to examine what I call the *Latino SRH paradox*: Latinos in the United States tend to be as healthy as or healthier than non-Hispanic whites across a range of health indicators (despite their comparably lower socioeconomic status; this is often referred to as the Hispanic/Latino health paradox [Markides and Eschbach 2005])—yet their SRH is often worse. The work of Lee and colleagues demonstrates that question order effects (differences in estimates of SRH depending on whether SRH precedes or follows other health items) tend to exist for Latinos (but not non-Hispanic whites) and Spanish speakers (but not English speakers) (Lee and Grant 2009; Lee and Schwarz 2014; Lee, Schwarz, and Goldstein 2014). The direction of these question order effects is important as well: Preceding SRH with other health items for all respondents produces estimates of SRH that are more similar across these groups (Latinos and non-Hispanic whites, Spanish and English speakers)—a potentially desirable outcome in terms of the Latino SRH paradox. Furthermore, preceding SRH with health items appears to increase the predictive validity of SRH answers with respect to subsequent mortality for Spanish speakers (Lee and Schwarz 2014). Thus, if SRH is to be used as a proxy for more objective health, preceding SRH with other health items appears to diminish differences in how SRH is interpreted (at least for Spanish-speaking and Latino respondents in the United States compared to other groups) and increase the predictive

validity of SRH answers (at least with respect to mortality) for Spanish speakers in the United States.

However, placing SRH before other domain-specific health items is advisable for several reasons (Garbarski, Schaeffer, and Dykema 2015a). First, the content of the other health items that precede SRH will vary across studies, with implications for cross-survey comparability of SRH answers. For example, in the study by Garbarski, Schaeffer, and Dykema (2015a), the health items preceding SRH included a range of health behaviors, conditions, and limitations. A study with different health questions preceding SRH (Lee, Schwarz, and Goldstein 2014) showed a different pattern of results. Although several factors could contribute to the discrepant results of the two studies (such as the mode and target population), it raises the question of whether the different health contexts elicited by the health items produced different associations between SRH answers and an index of current health risks constructed from these other health items.⁴ Second, the distribution of health conditions, behaviors, and limitations varies across study populations, which can also reduce cross-survey comparability of estimates of SRH answers if the health items preceding SRH are relevant for one population but not another, such that the preceding health questions “define health” in different ways across the two populations. Third, these first two issues and the apparent assimilation effects—stronger associations between SRH answers and other health items when SRH is administered after these items compared to before—have implications for many types of multivariate analysis in which SRH and other health items from the survey are modeled simultaneously, such as potentially increasing multicollinearity when SRH and other health items are independent variables in a model or attenuating the effects of other independent variables when SRH is the dependent variable and the other health items are independent variables.

These suggestions are relevant regardless of whether SRH is conceptualized as a proxy for more objective health or as a perception of health. If SRH is considered a proxy for more objective health, preceding SRH with health items may reduce the comparability of SRH estimates and multivariate analyses across studies because of variation across studies in the preceding health items used and their relevance for a given study population. Furthermore, if the goal is for SRH to capture perceptions of health, it may be problematic to precede SRH with other health items, as influencing the health factors respondents

4. In the California Health Interview Survey used by Lee and colleagues (2014), SRH is asked after questions about (1) chronic health conditions or (2) mental health assessment and service utilization questions. This study shows that those with one (English- and Spanish-speaking) or two (Spanish-speaking) current comorbidities (an index constructed from the other health items) have a higher proportion of positive health ratings when SRH is presented *last compared to first*. However, Garbarski, Schaeffer, and Dykema (2015a) find that respondents with four or more current health risks (constructed from the other health items) have significantly higher (better) mean SRH when SRH is presented *first compared to last*.

consider when rating their health may diminish the between-group differences that are precisely of interest.

A related issue concerns cross-survey (across groups and over time) comparability of SRH beyond the issue of question order. In theory, the features of survey measurement—mode, question wording, context, response option order, and so forth—should be standardized to the extent possible regardless of the stated purpose of SRH, as the standardized measurement of SRH is desirable for cross-survey comparability regardless of whether SRH is being used as a proxy for more objective health or as a perception of health. In practice, this will not be the case given the myriad permutations of design and implementation features in surveys containing SRH.

SUMMARY

Overall, the conceptual model of the factors that influence respondents' SRH answers highlights the complex interplay among the health, psychological, social, and survey measurement factors. The implications of this interplay depend on the stated purpose of SRH. If SRH is conceptualized as a proxy for more objective health (that is, emphasizing the “health” in SRH), then controlling for or removing the influence of the psychological, social, and survey measurement factors is needed, particularly when making cross-group or cross-survey comparisons of SRH. However, if the purpose of SRH is to capture respondents' perceptions of health (that is, emphasizing the “self-rated”), then each factor is part of the “SRH package” as something to be unpacked and understood rather than controlled.

Future Research Directions to Improve the Understanding, Measurement, and Analysis of SRH

This section draws on the synthesis of the extant research to outline future research directions that focus on 1) further explicating the factors that underlie respondents' SRH answers and 2) improving features of the measurement and analysis of SRH.

HOW TO MEASURE GROUP DIFFERENCES IN HEALTH WITH SRH

The conceptual model highlights an individual-level model in terms of which factors influence respondents' SRH answers. However, extant studies demonstrate systematic differences across groups in the association between health factors and SRH, indicating potential differences in evaluative frameworks across groups (that is, differences across groups in how respondents rate their health). These findings raise an important question as to “where and how” SRH can be used to examine differences in health across specific subgroups (Jylhä 2009; Dowd 2012).

As noted in the preceding sections, controlling for differences in evaluative frameworks is not desirable if the stated purpose of SRH is to capture perceptions of health. If SRH is being used as a proxy for more objective health, however, differences in evaluative frameworks are problematic for cross-group comparisons. “Where and how” SRH can be used to examine group differences in health when SRH is used as a proxy for more objective health has been addressed with methods that attempt to make SRH “more objective” by controlling for some aspect of subjectivity. These methods include anchoring vignettes (King et al. 2004; Grol-Prokopczyk, Freese, and Hauser 2011; Grol-Prokopczyk et al. 2015; Bzostek et al. 2016; Xu and Xie forthcoming), controlling for more objective health measures in analysis, attempting to isolate particular parts of evaluative frameworks and their differences across groups (Layes, Asada, and Kephart 2012; Hardy, Acciai, and Reyes 2014; Altman, Van Hook, and Hillemeier 2016), and influencing the definition of health respondents use by preceding the presentation of SRH with specific health items (Lee and Schwarz 2014). Each method seeks to standardize SRH for cross-group comparison (either by standardizing the researcher’s interpretation of SRH answers using other measures as calibration or by standardizing respondents’ interpretations of SRH when the item is administered), and each has particular assumptions, strengths, and weaknesses. Unknown and uncontrolled sources of differences in evaluative frameworks still remain with these methods, and future research that continues to refine these methods and compare them to one another—particularly within a given study—is warranted.

Furthermore, the very notion of “health disparities” emphasizes differences between groups while glossing over within-group heterogeneity. For example, while between-group differences in cultural orientations exist, not all Latinos in the United States have “collectivist” orientations and not all non-Hispanic whites have “individualist” orientations (Lee, Schwarz, and Goldstein 2014). Future research is needed that goes beyond the question of intergroup comparability and develops conceptual and empirical models of SRH that are nuanced with respect to intersecting systems of oppression and identity (Crenshaw 1989; Choo and Ferree 2010) and within-group variability.

MORE DIRECT OBSERVATION OF THE SRH RESPONSE PROCESS

Existing research documents the health factors respondents consider when answering SRH by either quantifying the association between SRH and other health measures or asking respondents which health factors they considered when formulating their answer. Future research should extend this line of inquiry, examining both *which* health factors respondents take into account when rating their health as well as *how* they take these factors into account: the processes and pathways through which health factors and the SRH question itself are experienced, conceptualized, interpreted, and integrated to formulate a health rating. By focusing on respondents’ explanations and accounts

of how they formulate their answer in addition to the health factors considered, such studies will be able to characterize more completely the processes through which respondents rate their health. In particular, studies are needed that focus on more direct observation of the SRH response process as it occurs. These studies could examine explanations and accounts of how respondents arrived at their answers through cognitive interviewing (Garbarski et al. 2015) or by observing features of the interviewer–respondent interaction (Garbarski, Schaeffer, and Dykema 2011).

FEATURES OF SURVEY MEASUREMENT OF SRH

The overall picture of how survey measurement features impact SRH answers is incomplete, and little is known about how combinations of survey measurement features *jointly* impact the distribution of SRH answers and their association with covariates. This has implications for improving the measurement of SRH, as the “best” version likely varies across populations and the stated purpose of SRH. For example, while a particular combination of survey measurement factors may yield an optimal prediction of subsequent mortality (a criterion for SRH when it is a proxy for more objective health), this combination may not yield optimal measurement of respondents’ perceptions of their health—and the effects of these combinations of survey measurement factors on SRH answers may vary across groups. Furthermore, many researchers are working with measures of SRH in which they had little or no control over how the data were collected. The best course of action for researchers in this position is to work within the existing structure of their survey data, at least noting the particular features of the measurement of SRH in their study when reporting results to allow for an assessment of how their particular SRH measure and attendant results compare with other studies.

Here, I outline suggestions for future research concerning the survey measurement of SRH and note that the implications of a proposed study may vary across populations and stated purpose of SRH. While recent research has focused on the effects of question order and response option order on SRH answers, these studies should be replicated across different populations. In addition, future research should examine the effects on SRH answers of 1) the number of response options; 2) response option labels; 3) mode of administration; and 4) question wording.

First, the number of response options used varies depending on the survey. For example, the General Social Survey uses a version with “excellent, good, fair, or poor,” and the Gallup Panel qualifies “fair” with “only” (Shim, Shin, and Johnson 2013); many other surveys also include “very good” between “good” and “excellent.” Including more response options may make sense in order to capture finer gradations in health, particularly positive options such as “very good” to obtain more discrimination in populations that are healthier or

that view health optimistically, but this claim warrants empirical investigation (Smith 2005).

Second, the response option labels for SRH vary across studies, with implications for making comparisons across studies. Many US studies use some version of a scale ranging from “excellent” to “poor,” while the World Health Organization and many European studies use response options that range from “very good” to “very bad”; still other researchers are examining different response options altogether, such as a rating thermometer (Perneger et al. 2013). It is unclear which version is preferable (which likely depends on stated purpose) and how to make comparisons across surveys that use different versions of the scale (Eriksson, Undén, and Elofsson 2001; Jürges, Avendano, and Mackenbach 2008). A further consideration is the language of the survey, which may lead to variation across groups being compared if different translations do not correspond to similar meanings for response options nor similar distances between the response options (Bzostek, Goldman, and Pebley 2007; Sanchez and Vargas 2016; Viruell-Fuentes et al. 2011). For both the response option labels and the language of administration, more research is needed to produce recommendations for best practices in surveys.

Third, little attention has been paid to differences in SRH answers across modes, holding constant other features of survey measurement. One notable exception is a recent experiment using the Pew Research Center’s American Trends Panel, showing that respondents interviewed over the phone were more likely to report “excellent” (and less likely to report “very good” or “good”) health compared to web respondents (Pew Research Center 2015). In addition, in the Gallup Panel 2008 Health Survey, web respondents had better SRH than mail respondents (after controlling for relevant covariates associated with Internet access) (Shim, Shin, and Johnson 2013). Further research is needed that compares the distribution of SRH and its association with covariates across other modes and across survey measurement features within a mode (e.g., horizontal or vertical presentation of response options in a self-administered mode [Garbarski, Schaeffer, and Dykema 2015b]), and that also incorporates respondents’ characteristics such as cognitive ability and socioeconomic status.

Fourth, more research should examine whether the distribution of SRH and its association with covariates are influenced by differences in question wording—for example, invoking a particular reference period or group: “In general would you say your health is,” “Overall would you say your health is,” “During the last four weeks would you say your health is,” and “Compared to others your age and sex would you say your health is” (Singer 1994; DeSalvo et al. 2006). The latter two versions define reference periods and groups, which may standardize the comparisons used in assessing health and so be useful for certain purposes, but also may render these measures incomparable to each other and the former, more general versions.

ADDITIONAL DIRECTIONS FOR FUTURE RESEARCH

In online [appendix A](#), I propose two additional directions for future research that extend the conceptual model beyond respondents' ratings of their health at one point in time to ratings of general health made by others and the measurement of SRH over time.

Conclusion

Over forty years ago, [Maddox and Douglass \(1973\)](#) succinctly described the paradox of SRH as a ubiquitous, cost-effective, and potentially valid measure of health in surveys: "These ratings clearly measures something more—and something less—than objective medical ratings" (92). This paradox has implications for the understanding, measurement, and analysis of health in survey research using SRH. This paper reviews existing research on the health, psychological, social, and survey measurement factors that influence respondents' SRH answers. The synthesis of extant research lays a framework for future research on SRH that focuses on further explicating the factors that underlie respondents' SRH answers and improving features of the measurement and analysis of SRH. These future research areas include where and how to examine group differences in SRH, a more direct observation of the SRH response process, examining additional features of survey measurement with respect to SRH, examining ratings of general health made by others, and examining the measurement of SRH over time. While much insightful work has been done, future research can do much more to delineate what underlies respondents' SRH answers and to improve the measurement and analysis of SRH to fit the various goals of the researchers who both collect and analyze the data.

Supplementary Data

Supplementary data are available online at <http://poq.oxfordjournals.org/>

References

- Aiken-Morgan, Adrienne T., Jacqueline Bichsel, Jyoti Savla, Christopher L. Edwards, and Keith E. Whitfield. 2014. "Associations between Self-Rated Health and Personality." *Ethnicity and Disease* 24:418–22.
- Altman, Claire E., Jennifer Van Hook, and Marianne Hillemeier. 2016. "What Does Self-Rated Health Mean? Changes and Variations in the Association of Obesity with Objective and Subjective Components of Self-Rated Health." *Journal of Health and Social Behavior* 57:39–58.
- Assari, Shervin, Maryam Moghani Lankarani, and Sarah Burgard. 2016. "Black–White Difference in Long-Term Predictive Power of Self-Rated Health on All-Cause Mortality in United States." *Annals of Epidemiology* 26:106–14.
- Bailis, Daniel S., Alexander Segall, and Judith G. Chipperfield. 2003. "Two Views of Self-Rated General Health Status." *Social Science & Medicine* 56:203–17.

- Benjamins, Maureen Reindl, Robert A. Hummer, Isaac W. Eberstein, and Charles B. Nam. 2004. "Self-Reported Health and Adult Mortality Risk: An Analysis of Cause-Specific Mortality." *Social Science & Medicine* 59:1297–306.
- Benyamini, Yael, Ellen L. Idler, Howard Leventhal, and Elaine A. Leventhal. 2000. "Positive Affect and Function as Influences on Self-Assessments of Health." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 55:107–16.
- Benyamini, Yael, Elaine A. Leventhal, and Howard Leventhal. 1999. "Self-Assessments of Health." *Research on Aging* 21:477–500.
- . 2000. "Gender Differences in Processing Information for Making Self-Assessments of Health." *Psychosomatic Medicine* 62:354–64.
- . 2003. "Elderly People's Ratings of the Importance of Health-Related Factors to Their Self-Assessments of Health." *Social Science & Medicine* 56:1661–67.
- Bjornstrom, Eileen E. S., and Danielle C. Kuhl. 2014. "A Different Look at the Epidemiological Paradox: Self-Rated Health, Perceived Social Cohesion, and Neighborhood Immigrant Context." *Social Science & Medicine* 120:118–25.
- Bowling, Ann, and John Windsor. 2008. "The Effects of Question Order and Response-Choice on Self-Rated Health Status in the English Longitudinal Study of Ageing (ELSA)." *Journal of Epidemiology and Community Health* 62:81–85.
- Brenner, Philip S., and John DeLamater. 2016. "Measurement Directiveness as a Cause of Response Bias." *Sociological Methods & Research* 45:348–71.
- Burgard, Sarah A., and Patricia V. Chen. 2014. "Challenges of Health Measurement in Studies of Health Disparities." *Social Science & Medicine* 106:143–50.
- Bzostek, Sharon, Noreen Goldman, and Anne Pebley. 2007. "Why Do Hispanics in the USA Report Poor Health?" *Social Science & Medicine* 65:990–1003.
- Bzostek, Sharon, Narayan Sastry, Noreen Goldman, Anne Pebley, and Denise Duffy. 2016. "Using Vignettes to Rethink Latino–White Disparities in Self-Rated Health." *Social Science & Medicine* 149:46–65.
- Canfield, Beth, Kristen Miller, Paul Beatty, Karen Whitaker, Alfredo Calvillo, and Barbara Wilson. 2003. *Adult Questions on the Health Interview Survey—Results of Cognitive Testing Interviews Conducted April–May 2003*. Hyattsville, MD: National Center for Health Statistics, Cognitive Methods Staff, 1–41.
- Choo, Hae Yeon, and Myra Marx Ferree. 2010. "Practicing Intersectionality in Sociological Research: A Critical Analysis of Inclusions, Interactions, and Institutions in the Study of Inequalities." *Sociological Theory* 28:129–49.
- Crenshaw, Kimberle. 1989. "Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics." *University of Chicago Legal Forum* 139–67.
- Crossley, Thomas F., and Steven Kennedy. 2002. "The Reliability of Self-Assessed Health Status." *Journal of Health Economics* 21:643–58.
- DeSalvo, Karen B., Nicole Bloser, Kristi Reynolds, Jiang He, and Paul Muntner. 2006. "Mortality Prediction with a Single General Self-Rated Health Question." *Journal of General Internal Medicine* 21:267–75.
- Dowd, Jennifer Beam. 2012. "Whiners, Deniers, and Self-Rated Health: What Are the Implications for Measuring Health Inequalities? A Commentary on Layes et al." *Social Science & Medicine* 75:10–13.
- Dowd, Jennifer Beam, and Megan Todd. 2011. "Does Self-Reported Health Bias the Measurement of Health Inequalities in US Adults? Evidence Using Anchoring Vignettes from the Health and Retirement Study." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 66B:478–89.
- Dowd, Jennifer Beam, and Anna Zajacova. 2007. "Does the Predictive Power of Self-Rated Health for Subsequent Mortality Risk Vary by Socioeconomic Status in the Us?" *International Journal of Epidemiology* 36:1214–21.

- . 2010. "Does Self-Rated Health Mean the Same Thing across Socioeconomic Groups? Evidence from Biomarker Data." *Annals of Epidemiology* 20:743–49.
- Eriksson, Ingeborg, Anna-Lena Undén, and Stig Elofsson. 2001. "Self-Rated Health. Comparisons between Three Different Measures. Results from a Population Study." *International Journal of Epidemiology* 30:326–33.
- Ferraro, Kenneth F., and Jessica A. Kelley-Moore. 2001. "Self-Rated Health and Mortality among Black and White Adults Examining the Dynamic Evaluation Thesis." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 56:S195–205.
- Finch, Brian Karl, Robert A. Hummer, Maureen Reindl, and William A. Vega. 2002. "Validity of Self-Rated Health among Latino(a)s." *American Journal of Epidemiology* 155:755–59.
- Garbarski, Dana, Jennifer Dykema, Kenneth D. Croes, Tara Piché, and Dorothy Farrar Edwards. 2015. "How Respondents Report Their Health Status: Cognitive Interviews of Self-Rated Health across Race, Ethnicity, Gender, Age, and Socioeconomic Status." Paper presented at the Annual Meeting of the American Association for Public Opinion Research, May, Hollywood, FL, USA.
- Garbarski, Dana, Nora Cate Schaeffer, and Jennifer Dykema. 2011. "Are Interactional Behaviors Exhibited When the Self-Reported Health Question Is Asked Associated with Health Status?" *Social Science Research* 40:1025–36.
- . 2015a. "The Effects of Question Order and Response Option Order on Self-Rated Health." *Quality of Life Research* 24:1443–53.
- . 2015b. "Examining the Effects of Response Option Order and Scale Orientation on Self-Rated Health." Paper presented at the Annual Meeting of the Midwest Association for Public Opinion Research, Chicago, IL, USA.
- . 2016. "The Effect of Response Option Order on Self-Rated Health: A Replication Study." *Quality of Life Research* 27:2117–21.
- Grol-Prokopczyk, Hanna, Emese Verdes-Tennant, Mary McEniry, and Márton Ispány. 2015. "Promises and Pitfalls of Anchoring Vignettes in Health Survey Research." *Demography* 52:1703–28.
- Grol-Prokopczyk, Hanna, Jeremy Freese, and Robert M. Hauser. 2011. "Using Anchoring Vignettes to Assess Group Differences in General Self-Rated Health." *Journal of Health and Social Behavior* 52:246–61.
- Groves, Robert M., Nancy Fultz, and Elizabeth Martin. 1992. "Direct Questioning About Comprehension in a Survey Setting." In *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, edited by Judith M. Tanur, 49–61. New York: Russell Sage Foundation.
- Hardy, Melissa A, Francesco Acciai, and Adriana M. Reyes. 2014. "How Health Conditions Translate into Self-Ratings: A Comparative Study of Older Adults across Europe." *Journal of Health and Social Behavior* 55:320–41.
- Herrmann, Douglas J., Jennifer Madans, Ronald Wilson, and Floyd J. Fowler. 1994. "The National Center for Health Statistics Perspective: Panel Discussion." *Proceedings of the 1993 NCHS Conference on the Cognitive Aspects of Self-Reported Health Status*, 113–25. Hyattsville, MD: National Center for Health Statistics, Cognitive Methods Staff.
- Hitlin, Steven, Lance D. Erickson, and J. Scott Brown. 2015. "Agency and Mental Health: A Transition to Adulthood Paradox." *Society and Mental Health* 5:163–81.
- Hitlin, Steven, and Monica Kirkpatrick Johnson. 2015. "Reconceptualizing Agency within the Life Course: The Power of Looking Ahead." *American Journal of Sociology* 120:1429–72.
- Holbrook, Allyson L., Jon A. Krosnick, Richard T. Carson, and Robert Cameron Mitchell. 2000. "Violating Conversational Conventions Disrupts Cognitive Processing of Attitude Questions." *Journal of Experimental Social Psychology* 36:465–94.
- Hopkins, Daniel J., and Gary King. 2010. "Improving Anchoring Vignettes Designing Surveys to Correct Interpersonal Incomparability." *Public Opinion Quarterly* 74:201–22.
- Huisman, Martijn, and Dorly J. H. Deeg. 2010. "A Commentary on Marja Jylhä's 'What Is Self-Rated Health and Why Does It Predict Mortality? Towards a Unified Conceptual Model' (69:3, 2009, 307–16)." *Social Science & Medicine* 70:652–54.

- Idler, Ellen L. 1993. "Age Differences in Self-Assessments of Health: Age Changes, Cohort Differences, or Survivorship?" *Journal of Gerontology* 48:S289–300.
- . 1994. "Self-Ratings of Health: Mortality, Morbidity, and Meaning." *Proceedings of the 1993 NCHS Conference on the Cognitive Aspects of Self-Reported Health Status*, 113–25. Hyattsville, MD: National Center for Health Statistics, Cognitive Methods Staff.
- Idler, Ellen L., and Yael Benyamini. 1997. "Self-Rated Health and Mortality: A Review of Twenty-Seven Community Studies." *Journal of Health and Social Behavior* 38:21–37.
- Idler, Ellen L., and Stanislav V. Kasl. 1995. "Self-Ratings of Health: Do They Also Predict Change in Functional Ability?" *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 50B:S344–53.
- Idler, Ellen, Howard Leventhal, Julie McLaughlin, and Elaine Leventhal. 2004. "In Sickness but Not in Health: Self-Ratings, Identity, and Mortality." *Journal of Health and Social Behavior* 45:336–56.
- Johnson, Timothy P., Shavitt, Sharon, and Allyson L. Holbrook. 2011. "Survey response styles across cultures." In *Cross-Cultural Research Methods in Psychology*, eds D. Matsumoto and F. J. R. Van de Vijver, pp. 130–78. Cambridge: Cambridge University Press.
- Jürges, Hendrik. 2007. "True Health vs. Response Styles: Exploring Cross-Country Differences in Self-Reported Health." *Health Economics* 16:163–78.
- Jürges, Hendrik, Mauricio Avendano, and Johan P. Mackenbach. 2008. "Are Different Measures of Self-Rated Health Comparable? An Assessment in Five European Countries." *European Journal of Epidemiology* 23:773–81.
- Jylhä, Marja. 2009. "What Is Self-Rated Health and Why Does It Predict Mortality? Towards a Unified Conceptual Model." *Social Science & Medicine* 69:307–16.
- Jylhä, Marja. 2010. "Self-Rated Health between Psychology and Biology. A Response to Huisman and Deeg." *Social Science & Medicine* 70:655–57.
- Kaplan, Giora, and Orna Baron-Epel. 2003. "What Lies Behind the Subjective Evaluation of Health Status?" *Social Science & Medicine* 56:1669–76.
- King, Gary, Christopher J. L. Murray, Joshua A. Salomon, and Ajay Tandon. 2004. "Enhancing the Validity and Cross-Cultural Comparability of Measurement in Survey Research." *American Political Science Review* 98:191–207.
- Knäuper, Bärbel, and Patricia A. Turner. 2003. "Measuring Health: Improving the Validity of Health Assessments." *Quality of Life Research* 12:81–89.
- Krause, Neal M., and Gina M. Jay. 1994. "What Do Global Self-Rated Health Items Measure?" *Medical Care* 32:930–42.
- Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5:213–36.
- Krosnick, Jon A., and Duane F. Alwin. 1987. "An Evaluation of a Cognitive Theory of Response-Order Effects in Survey Measurement." *Public Opinion Quarterly* 51:201–19.
- Layes, Audrey, Yukiko Asada, and George Kephart. 2012. "Whiners and Deniers—What Does Self-Rated Health Measure?" *Social Science & Medicine* 75:1–9.
- Lee, Sunghye. 2014. "Self-Rated Health in Health Surveys." *Health Survey Methods* 193–216.
- Lee, Sunghye, and David Grant. 2009. "The Effect of Question Order on Self-Rated General Health Status in a Multilingual Survey Context." *American Journal of Epidemiology* 169:1525–30.
- Lee, Sunghye, and Norbert Schwarz. 2014. "Question Context and Priming Meaning of Health: Effect on Differences in Self-Rated Health between Hispanics and Non-Hispanic Whites." *American Journal of Public Health* 104:179–85.
- Lee, Sunghye, Norbert Schwarz, and Leanne Streja Goldstein. 2014. "Culture-Sensitive Question Order Effects of Self-Rated Health between Older Hispanic and Non-Hispanic Adults in the United States." *Journal of Aging and Health* 26:860–83.
- Letzring, Tera D., Grant W. Edmonds, and Sarah E. Hampson. 2014. "Personality Change at Mid-Life Is Associated with Changes in Self-Rated Health: Evidence from the Hawaii Personality and Health Cohort." *Personality and Individual Differences* 58:60–64.
- Lindeboom, Maarten, and Eddy van Doorslaer. 2004. "Cut-Point Shift and Index Shift in Self-Reported Health." *Journal of Health Economics* 23:1083–99.

- Maddox, George L., and Elizabeth B. Douglass. 1973. "Self-Assessment of Health: A Longitudinal Study of Elderly Subjects." *Journal of Health and Social Behavior* 14:87–93.
- Manderbacka, Kristiina. 1998. "Examining What Self-Rated Health Question Is Understood to Mean by Respondents." *Scandinavian Journal of Public Health* 26:145–53.
- Markides, Kyriakos S., and Karl Eschbach. 2005. "Aging, Migration, and Mortality: Current Status of Research on the Hispanic Paradox." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 60(Special Issue 2):S68–75.
- Mavaddat, Nahal, Jose M. Valderas, Rianne van der Linde, Kay T. Khaw, and Ann L. Kinmonth. 2014. "Association of Self-Rated Health with Multimorbidity, Chronic Disease and Psychosocial Factors in a Large Middle-Aged and Older Cohort from General Practice: A Cross-Sectional Study." *BMC Family Practice* 15:185–95.
- Means, Barbara, Arti Nigam, Marlene Zarrow, Elizabeth F. Loftus, and Molla S. Donaldson. 1989. *Autobiographical Memory for Health-Related Events*. US Department of Health and Human Services, Public Health Service, Centers for Disease Control, National Center for Health Statistics.
- Miller, Kristen, Gordon Willis, Connie Eason, Lisa Moses, and Beth Canfield. 2005. "Interpreting the Results of Cross-Cultural Cognitive Interviews." In *Methodological Aspects in Cross-National Research*, edited by Jurgen H. P. Homeyer-Zlotnik and Janet Harkness, 79–92. Mannheim, Germany: ZUMA Nachrichten Spezial.
- Organization for Economic Co-Operation and Development. 2015. *Health at a Glance 2015*. OECD Publishing.
- Perneger, Thomas V., Angèle Gayet-Ageron, Delphine S. Courvoisier, Thomas Agoritsas, and Stéphane Cullati. 2013. "Self-Rated Health: Analysis of Distances and Transitions between Response Options." *Quality of Life Research* 22:2761–68.
- Pew Research Center. 2015. "From Telephone to the Web: The Challenge of Mode of Interview Effects in Public Opinion Polls." <http://www.pewresearch.org/2015/05/13/from-telephone-to-the-web-the-challenge-of-mode-of-interview-effects-in-public-opinion-polls/>.
- Quesnel-Vallée, Amélie. 2007. "Self-Rated Health: Caught in the Crossfire of the Quest for 'True' Health?" *International Journal of Epidemiology* 36:1161–64.
- Ryff, Carol D., Barry T. Radler, and Elliot M. Friedman. 2015. "Persistent Psychological Well-Being Predicts Improved Self-Rated Health over 9–10 Years: Longitudinal Evidence from MIDUS." *Health Psychology Open* 2(2).
- Sanchez, Gabriel R., and Edward D. Vargas. 2016. "Language Bias and Self-Rated Health Status among the Latino Population: Evidence of the Influence of Translation in a Wording Experiment." *Quality of Life Research* 25:1131–36.
- Schaeffer, Nora Cate. 2000. "Asking Questions About Threatening Topics: A Selective Overview." In *The Science of Self-Report: Implications for Research and Practice*, edited by Christine A. Bachrach, Arthur A. Stone, Jared B. Jobe, Howard S. Kurtzman, Virginia S. Cain, 105–22. Mahwah, NJ: Lawrence Erlbaum Associates.
- Schnittker, Jason. 2005a. "Cognitive Abilities and Self-Rated Health: Is There a Relationship? Is It Growing? Does It Explain Disparities?" *Social Science Research* 34:821–42.
- . 2005b. "When Mental Health Becomes Health: Age and the Shifting Meaning of Self-Evaluations of General Health." *Milbank Quarterly* 83:397–423.
- Schwartz, Carolyn E., Elena M. Andresen, Margaret A. Nosek, and Gloria L. Krahn. 2007. "Response Shift Theory: Important Implications for Measuring Quality of Life in People with Disability." *Archives of Physical Medicine and Rehabilitation* 88:529–36.
- Schwarz, Norbert, and Herbert Bless. 1992. "Scandals and the Public's Trust in Politicians: Assimilation and Contrast Effects." *Personality and Social Psychology Bulletin* 18:574–74.
- Schwarz, Norbert, Fritz Strack, and Hans-Peter Mai. 1991. "Assimilation and Contrast Effects in Part-Whole Question Sequences: A Conversational Logic Analysis." *Public Opinion Quarterly* 55:3–23.
- Schulster, Jerome R. 1994. "Health and Self: Paths for Exploring Cognitive Aspects Underlying the Self-Report of Health Status." *Proceedings of the 1993 NCHS Conference on the Cognitive*

- Aspects of Self-Reported Health Status*, 89–105. Hyattsville, MD: National Center for Health Statistics, Cognitive Methods Staff.
- Sen, Amartya. 1993. "Positional Objectivity." *Philosophy & Public Affairs* 22:126–45.
- . 2002. "Health: Perception Versus Observation: Self Reported Morbidity Has Severe Limitations and Can Be Extremely Misleading." *British Medical Journal* 324:860–61.
- Shim, Jae-Mahn, Eunjung Shin, and Timothy P. Johnson. 2013. "Self-Rated Health Assessed by Web versus Mail Modes in a Mixed Mode Survey: The Digital Divide Effect and the Genuine Survey Mode Effect." *Medical Care* 51:774–81.
- Simon, J. G., J. B. De Boer, I. M. A. Joung, Hans Bosma, and J. P. Mackenbach. 2005. "How Is Your Health in General? A Qualitative Study on Self-Assessed Health." *European Journal of Public Health* 15:200–208.
- Singer, Eleanor. 1994. "Self-Rated Health: How Are Judgments Made?" *Proceedings of the 1993 NCHS Conference on the Cognitive Aspects of Self-Reported Health Status*, 28–35. Hyattsville, MD: National Center for Health Statistics, Cognitive Methods Staff.
- Singh-Manoux, Archana, Pekka Martikainen, Jane Ferrie, Marie Zins, Michael Marmot, and Marcel Goldberg. 2006. "What Does Self-Rated Health Measure? Results from the British Whitehall II and French Gazel Cohort Studies." *Journal of Epidemiology and Community Health* 60:364–72.
- Smith, Tom W. 2005. "The Impact of Alternative Response Scales on Measuring Self-Ratings of Health." NORC/University of Chicago, GSS Methodological Report No. 103. <http://gss.norc.org/Documents/reports/methodological-reports/MR103.pdf>.
- Sprangers, Mirjam A. G., and Carolyn E. Schwartz. 1999. "Integrating Response Shift into Health-Related Quality of Life Research: A Theoretical Model." *Social Science & Medicine* 48:1507–15.
- Sudman, Seymour, and Norman M. Bradburn. 1982. *Asking Questions*. San Francisco: Jossey-Bass.
- Tourangeau, Roger, Kenneth A. Rasinski, and Norman M. Bradburn. 1991. "Measuring Happiness in Surveys: A Test of the Subtraction Hypothesis." *Public Opinion Quarterly* 55:255–66.
- Tourangeau, Roger, Lance J. Rips, and Kenneth A. Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press.
- Tversky, Amos, and Daniel Kahneman. 1974. "Judgment under Uncertainty: Heuristics and Biases." *Science* 185:1124–31.
- Viruell-Fuentes, Edna A., Jeffrey D. Morenoff, David R. Williams, and James S. House. 2011. "Language of Interview, Self-Rated Health, and the Other Latino Health Puzzle." *American Journal of Public Health* 101:1306–13.
- Warnecke, Richard B., Timothy P. Johnson, Noel Chávez, Seymour Sudman, Diane P. O'Rourke, Loretta Lacey, and John Horm. 1997. "Improving Question Wording in Surveys of Culturally Diverse Populations." *Annals of Epidemiology* 7:334–42.
- Xu, Hongwei, and Yu Xie. Forthcoming. "Assessing the Effectiveness of Anchoring Vignettes in Bias Reduction for Socioeconomic Disparities in Self-Rated Health among Chinese Adults." *Sociological Methodology*.
- Zheng, Hui, and Patricia A. Thomas. 2013. "Marital Status, Self-Rated Health, and Mortality Overestimation of Health or Diminishing Protection of Marriage?" *Journal of Health and Social Behavior* 54:128–43.