Article

Field Methods
1-18

DOI: 10.1177/1525822X20908575
journals.sagepub.com/home/fmx

Exploring Mental Models behind Self-rated Health and Subjective Life Expectancy through Web Probing

Sunghee Lee[1], Colleen McClain[1], Dorothée Behr[2], and Katharina Meitinger[3]

Abstract

Self-rated health (SRH) and subjective life expectancy (SLE) are widely used for understanding health and predicting mortality. However, what these items measure remains unclear, due to the lack of conceptual frameworks. We administered a web survey across the United States, Great Britain, Germany, Spain, and Mexico. The questionnaire included SRH and SLE, each immediately followed by a question that probed respondents' thought processes. We examined the relationship between SRH and SLE, the response difficulty, and attributes that respondents considered for forming responses. Overall, SRH and SLE were moderately related, eliciting different information and varying in difficulty. Compared to SLE, SRH was

[1] Institute for Social Research, University of Michigan, Ann Arbor, MI, USA
[2] GESIS—Leibniz-Institute for the Social Sciences, Mannheim, Germany
[3] Utrecht University, Utrecht, the Netherlands

Corresponding Author:
Sunghee Lee, Institute for Social Research, University of Michigan, 426 Thompson St., Ann Arbor, MI 48104, USA.
Email: sungheel@umich.edu

perceived as easier but covered a narrower information spectrum. While illness and health behaviors were dominant attributes of SRH responses, family longevity history, life situations, and lack of control were additionally considered for SLE. When combined, SRH and SLE may capture a fuller range of attributes germane to health and mortality.

Population aging has made mortality prediction an important scientific and policy topic for which two survey questions are often considered: self-rated health (SRH) and subjective life expectancy (SLE). However, with individuals lacking access to information about their own mortality, it is difficult to gather useful data. Moreover, few researchers have systematically examined these questions in tandem, instead conducting studies that rely on unverified aspects of mortality prediction based on these questions. Because of this, the field lacks a nuanced understanding of what SRH and SLE measure.

This study focuses on improving our understanding about SRH and SLE, necessary for establishing their measurement frameworks. SRH asks respondents to rate their health and is by far the most popular health-related survey question (Fienberg et al. 1985). Shown to be a strong predictor of mortality, health outcomes, and care utilization (Idler and Benyamini 1997), SRH is recommended by health organizations (Hennessy et al. 1994). SLE, on the other hand, asks respondents to estimate their expectations about their own longevity and is administered in numerous aging-related surveys, such as the Health and Retirement Study in the United States (Perozek 2008). Like SRH, responses to SLE are shown to predict mortality (Siegel et al. 2003).

Measurement of SRH and SLE

Despite their wide use, neither SRH nor SLE is a "designed" question to elicit certain information on the topic of health or mortality. SRH was born out of fieldwork. Because health surveys may be lengthy and include questions on sensitive behaviors, interviewers asked some variations of SRH as an icebreaker (Elinson 1994). As such, SRH is simple to administer and conversational, often using verbally labeled Likert-type response scales—for instance, "excellent-very good-good-fair-poor," although variations exist (Lee 2014). For example, SRH in EQ-5D uses a 0–100 numeric scale (see appendices 2, 4, and 6 of Devlin and Brooks 2017).

In the 1960s, a particular question format asking about future expectations was born out of consumer behavior research (Juster 1964). By

applying this format to the concept of life and death, SLE was created (Lee et al. 2018; Manski 1990). SLE typically asks: "What is the percent chance that you will live to be [future age determined by respondents' current age] or more?" and uses a 0–100 numeric response scale. Variations to SLE wording exist, such as: "To what age do you expect to live?" (e.g., Mirowsky and Ross 2000) or "What is the percent chance that you will die by [future age determined by respondents' current age]?" (e.g., Kerry and Embretson 2018), as well as to its response scale, such as a four-point verbal scale from "very likely" to "very unlikely" (e.g., van Doorn and Kasl 1998).

Empirical studies have shown correlates of SRH and SLE, from socioeconomic status (e.g., education) to objective health status (e.g., chronic diseases) and to health behaviors (e.g., exercise), and to subtle nuances within each. For example, SLE responses are influenced by the health status of family members but more so by the health status of family members of the same sex as the respondents than others (Zick et al. 2014). For SRH, not only current but also expected future health status is related to one's mortality (Ferraro and Wilkinson 2015). Although these findings lead to a hypothesis that respondents incorporate different attributes when answering SRH than when answering SLE, they are based on relationships observed in collected data and not on actual survey response processes or theories. At the same time, respondents' incorporating subtle information germane to mortality that is unascertainable through objective measures is hypothesized to make SRH and SLE strong predictors of mortality (Perozek 2008; Stone et al. 2000). Surprisingly, the extant literature provides little methodological clarity regarding the type of information that SRH and SLE prompt respondents to consider, how the information is related to the underlying construct of mortality, how these questions are similar or different between SRH and SLE, or how best to ask them (Bailis et al. 2003; Griffin et al. 2013; Jylhä 2011). This makes formulating hypotheses about measurement mechanisms of SRH and SLE challenging.

Cognitive probing presents possibilities for assessing survey response processes, but few studies have utilized it in this context. Recent work has made use of open-ended probes on web surveys as a complement to resource-intensive in-person cognitive interviews for examining measurement error (Behr et al. 2017). Web probing may prove to be a fruitful route to examine the measurement mechanisms behind SRH and SLE.

Response Burden of SRH and SLE

SRH and SLE may pose different types of response difficulties. For SRH, although health is a topic frequently discussed in everyday life, there is no

standardized meaning (Bailis et al. 2003; Larson 1999). Respondents need to infer the meanings of response categories and differentiate adjacent categories (e.g., "fair" vs. "poor"). The major difficulties concerning SRH lie in understanding the meaning of health and mapping perceived health status to response categories (Lee and Schwarz 2014). The concept that SLE attempts to measure (i.e., life/death) is clear-cut. However, people often have insufficient information for predicting their own future (Jylhä 2011). Furthermore, its 101-point probability response scale, used infrequently in surveys, is known to have measurement problems even for straightforward behaviors like voting (de Bresser and van Soest 2019). Thus, with SLE, respondents are faced with an unfamiliar numeracy task for which they have insufficient information, making SLE particularly difficult to answer. Given this, it is not surprising to see reports about response heaping and high item nonresponse rates on SLE (Bruine de Bruin and Carman 2018; Lee et al. 2018; Lee et al. 2017; Lee and Smith 2016), both of which are signs of question difficulty (Couper et al. 2006).

Influences on Measurement of SRH and SLE: Culture and Psychosocial Traits

The concept of health and how one views one's own mortality may vary across cultures (Larson 1999; McCarthy et al. 2004). Moreover, survey response processes are also shown to vary across cultures (Johnson et al. 1997). With country being a proxy for culture (Hong 2009), it becomes tempting to designate SRH and SLE measurement mechanisms to country. Unfortunately, there is no literature that may guide such a nexus.

Nonetheless, the effect of culture-specific psychosocial traits on SLE response burden has recently emerged in the literature. A country-level analysis of older adults in 11 countries reported a negative relationship between future time orientation and SLE item missing rates (Lee et al. 2017). In an individual-level analysis of older adults, SLE item missing rates were associated positively with religiosity and present time orientation but negatively with future optimism and sense of control (Lee et al. 2018; Lee et al. 2017; Lee and Smith 2016). As persons with low optimism, low sense of control, and/or higher religiosity are reported to be less future oriented (Abeles 1991; Carter et al. 2012; Marko and Savickas 1998) and likely to have insufficient information for future prediction (Bergadaa 1990), SLE may be perceived as being more difficult by these individuals than the counterparts. It warrants extending this methodological analysis to younger persons.

Research Questions

This study attempts to address the validity of SRH and SLE, the cornerstone of measurement, as well as response burden. Thus, we pursue research that should have accompanied the development of these measures at the beginning of their use. Specifically, we examine SRH using the five-point verbal response scale and an alternative 101-point numeric scale, as well as SLE using the 101-point probability scale for the following questions.

First, on validity:

1. Concurrent validity: How comparable are responses to SRH and SLE with one another and with actuarial life expectancy (ALE) at the country level? Specifically, are countries ranked similarly across these measures?
2. Convergent validity: How are responses to SRH and SLE related to one another? Does response scale matter in this relationship?
3. Construct validity: Is the information captured by SRH and SLE similar or different?

Second, on potential respondent burden:

1. Does question difficulty vary between SRH and SLE and by aspects of their administration, specifically response scale?
2. How certain are respondents about their answers to SLE? Do psychosocial traits pertinent to time orientation play a role?

Method

Data

Our data came from a web survey with 2,689 respondents aged 18–65, sampled from opt-in panels in five countries: the United States, Great Britain, Germany, Spain, and Mexico. Despite lacking population representation, opt-in web samples are used increasingly as a cost-effective means of experimentation (e.g., Lee et al. 2016). As the focus of this study is examining response processes, population-level generalizability is not a major concern, unless factors that influence self-selection into opt-in panels also affect response processes. The sample was balanced on age (18–30, 31–50 vs. 51–65 years old), sex (male vs. female), and education (<tertiary/postsecondary education vs. ≥some tertiary/postsecondary education) across countries. Supplementary Material 1 (https://t.ly/xyx0p) provides the sample composition.

The survey was fielded in June 2014 by Respondi (www.respondi.com). The questionnaire was translated from English to German and Spanish and included questions on political and social attitudes as well as items specific to the experiments described shortly. Supplementary Material 2 (https://t.ly/OvGxA) provides exact wording of questions in the experiments by country. We supplemented the survey data with country-level ALE at birth (i.e., the number of years a person is expected to live) by the World Health Organization (2017) to examine the concurrent validity of SRH and SLE.

Experiments

For SRH, we randomized the response scale (five-point vs. 101-point) to examine whether response scale influences response burden, the type of information respondents use, and SRH's covariation with SLE and ALE. On SLE, we varied the format of probing question (open- vs. close-ended) described shortly. We implemented these experiments independently and assigned a random two-thirds of the sample to five-point SRH and another random two-thirds to the open-ended SLE probing question. Detailed descriptions of experimental features are provided in Supplementary Material 3 (https://t.ly/W61Nk). Our experiment module included three psychosocial measures: future optimism/time orientation, sense of control, and religiosity, which were shown to affect methodological as well as substantive aspects of SLE (Kerry and Embretson 2018; Lee et al. 2018; Lee et al. 2017).

Response Probing

Immediately after SRH and SLE, respectively, we probed respondents about their thought processes. All respondents for SRH and a random two-thirds for SLE were asked an open-ended question: "Please explain why you chose [answer on SLE/SRH]." The close-ended format given to one-third of SLE respondents asked: "Which of the options below best represents how you think about that answer? (1) I am very sure about the chance; (2) I am pretty sure about the chance; (3) I actually have no idea about the chance; and (4) No one can know the chance."

Variables of Interest and Analysis

We examined various aspects of SRH and SLE as outlined below. In all analyses, we made comparisons across countries. It should be noted that country-level comparisons were done for exploratory purposes due to the

absence of adequate theories. Ancillary comparisons by age, sex, and education are included in Supplementary Material 4 (https://t.ly/XNRW5).

Response Distribution

If SRH and SLE are to predict mortality: (1) Countries should rank similarly across SRH, SLE, and ALE; and (2) they should be related to each other, regardless of the response format. The five-point SRH was coded by combining "excellent–very good–good" response categories as "positive" health and the remaining categories as "negative" health. The 101-point SRH and SLE measures were used as continuous variables. We ranked countries on these variables along with ALE. Comparisons were made through *F*-tests. Note that, to address negative skewness, we transformed 101-point SRH and SLE using log and square root transformation, respectively, with reflection of the largest value.

Further, we examined the relationship using Spearman's rank-order correlation coefficients for 101-point SRH and SLE and point-biserial correlation coefficients for the binary SRH and SLE and tested the differences using Fisher's *Z*-transformation.

Probing Responses

We based our coding on a scheme developed for SRH in the U.S. context (Groves et al. 1992) and made adaptations to fit the expansion of countries and question types (see Supplementary Material 5 for coding details: https://t.ly/xyxwp). Specifically, we utilized the following nine codes for probing responses: (1) BMI-related health behaviors (e.g., exercise), (2) other health behaviors (e.g., substance use), (3) health service utilization (e.g., doctors' visits), (4) physical illness (e.g., specific ailments), (5) other illness (e.g., mental health), (6) general health comments (e.g., "my health is good"), (7) family history (e.g., parental longevity), (8) life situations, including demographics (e.g., "I am young"), and (9) other reasons (e.g., feelings). For SLE, we added two more: (10) lacking control (e.g., "up to God") and (11) mortality being up to chance (e.g., "50–50"). We coded all responses, allowing multiple codes per respondent. For example, a response to SLE probing that read, "No one can say for sure, but I don't smoke and I'm not overweight," was assigned with three codes: 10, 2, and 1. We excluded cases where respondents left the probing answer blank, said, "I don't know," or inserted uncodable responses (e.g., smiley faces), which occurred at 3.8% (SRH) and 7.9% (SLE). Overall, 65.7% and 54.7% of

SRH and SLE respondents, respectively, were given one code; the remainders were given multiple codes.

We compared the number of reported attributes by country in linear models and examined content of specific attributes. Because the results were virtually the same between all codes combined or the code given to the attribute reported the first, we used the content of the first code for a clearer illustration. Because SRH response formats (five-point and 101-point) did not affect the number and content of reported attributes, our results combined both formats.

There were four coders: two bilingual English–Spanish speakers coding English and Spanish cases; a native German speaker with high proficiency in Spanish coding German and Spanish cases, and a native German speaker coding only German cases. Coders included survey methodologists and trained undergraduate students. For testing coder reliability, random subsets of English responses were double-coded: (1) 147 SRH responses by two bilingual English–Spanish coders; (2) 29 SRH responses by one German and two bilingual English–Spanish and coders; and (3) 27 SLE responses by one bilingual English–Spanish and one German coder. Cohen's κ ranged from 0.540 to 0.839 for SRH and 0.717 for SLE.

Question Difficulty

We examined the difficulty through response time, item nonresponse rates, and response heaping. For response time (in seconds), we top-coded outliers above the 95th percentile for a given question within each country and compared the response time of SRH by response format and SLE by response status (i.e., item response vs. nonresponse). We examined how item nonresponse rates, format-specific SRH response times, and response-status-specific SLE response times varied by country through interaction terms in linear models. Interaction analyses used log-transformed response times as dependent variables to address their positive skewness. For 101-point SRH and SLE, we examined responses heaping at multiples of 10s and 25s (e.g., 20 or 75) and compared these distributions across countries by contrasting country-level means through F-tests in linear models.

SLE Response Uncertainty

There were multiple ways a respondent could express uncertainty when answering SLE in our study: (1) nonresponse to SLE itself, (2) nonresponse to SLE probing, (3) indicating inability to know answers to SLE on the

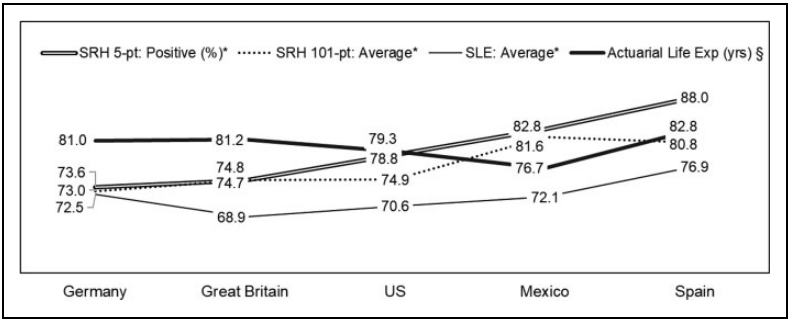

Figure 1. Distribution of self-rated health, subjective life expectancy, and actuarial life expectancy (ALE) by country. *Significant difference across countries at $p < .05$. §ALE at birth published by the World Health Organization (2017:figure 3.2)

open-ended probing by reporting attributes related to lacking control or longevity being up to chance, or (4) choosing "no one can know" or "no idea" responses in the close-ended probing. If any of the four indicators was positive for a given respondent, the respondent was coded as expressing uncertainty.

We modeled response uncertainty as a function of three psychosocial measures (future optimism/time orientation, lacking sense of control and religiosity) in logistic regression, while controlling for country, age, sex, and education. To ascertain whether psychosocial measures played a role in explaining uncertainty, we compared the goodness of fit of the model with and without these measures through a likelihood-ratio test.

Results

Response Distribution of SRH and SLE

Figure 1 includes positive health rates from five-point SRH, average scores of 101-point SRH and of SLE along with ALE (in years) by country. Given that our sample was not meant to represent the population, SRH and SLE cannot be expected to approximate ALE. Rather, if SRH and SLE convey similar information, their respective relationship to ALE should be similar. County-level estimates of SRH and SLE differed. On five-point SRH, Spain ranked the highest with 88.0% respondents reporting positive health, followed by Mexico (82.8%), the United States (78.8%), Great Britain (74.8%), and Germany (73.0%). On 101-point SRH and SLE, rankings were not identical but similar to five-point SRH, with Mexico and Spain at the

Table 1. Correlation Coefficients between Binary Positive Self-rated Health (SRH) and Subjective Life Expectancy (SLE) and between 101-point SRH and SLE by Country.

	Total	Germany	Great Britain	United States	Mexico	Spain
Positive SRH and SLE	$n = 1{,}455$	$n = 301$	$n = 295$	$n = 281$	$n = 315$	$n = 263$
Point-biserial correlation coefficients	.321	.339	.409	.285	.296	.239
101-Point SRH and SLE	$n = 714$	$n = 127$	$n = 137$	$n = 156$	$n = 139$	$n = 155$
Spearman's rank-order correlation coefficients	.343	.448	.350	.381	.299	.232

top. Rankings based on SRH and SLE did not correspond to ALE, however. On ALE, Mexico was ranked the lowest (76.7 years) and the United States the second lowest (79.3 years). Germany, Great Britain, and Spain were similar around 81–82 years.

The relationship between SRH and SLE in Table 1 was modest regardless the SRH response format: point-biserial correlation coefficients between five-point SRH and SLE ranged from 0.239 (Spain) to 0.409 (Great Britain); and Spearman's rank-order correlation coefficients between 101-point SRH and SLE from 0.232 (Spain) to 0.448 (Germany). No significant difference was observed across countries in these correlation estimates.

Probing Responses

Table 2 summarizes the number and content of attributes probed for SRH and SLE. On average, respondents reported fewer attributes on SRH (1.40) than SLE (1.55). Proportions of respondents who reported 0, 1, and 2 or more attributes were 3.8%, 61.9%, and 34.3% for SRH and 8.9%, 45.8%, and 45.2% for SLE. The counts differed by countries marginally significantly for SRH and significantly for SLE, with respondents in Mexico and Spain reporting more attributes on both questions than by those in the remaining countries.

The first codes probed for SRH were dominated by three attributes: illness (23.0% physical illness; 28.7% nonphysical illness), comments related to general health (17.7%), and health behaviors (10.7% BMI-related; 3.2% other). While the dominance of these three attributes

Table 2. Counts and Content of Reported Attributes for Self-rated Health (SRH) and Subjective Life Expectancy (SLE) by Country.

	SRH						SLE					
	Total	Germany	Great Britain	United States	Mexico	Spain	Total	Germany	Great Britain	United States	Mexico	Spain
Attribute counts[a]	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count	Count
(*n*)	2,627	538	533	540	537	524	1779	344	370	356	359	350
Total	3,738	740	727	736	783	752	2,754	513	574	476	631	560
Average	1.40	1.38	1.36	1.36	1.46	1.44	1.55	1.49	1.55	1.34	1.76	1.60
F-test	$F = 1.97$; $df = 4$ ($p = .097$)						$F = 9.46$; $df = 4$ ($p < .001$)					
Attribute content[b]	%	%	%	%	%	%	%	%	%	%	%	%
(*n*)	2,570	501	501	516	533	519	1,638	309	328	313	349	339
Health behavior: BMI-related	10.7	7.0	14.4	15.5	9.4	7.5	3.6	3.2	4.0	2.9	6.9	0.9
Health behavior: Non-BMI-related	3.2	2.4	2.8	3.9	4.1	2.9	9.4	8.1	5.2	6.7	18.6	7.7
Health service utilization	3.5	4.0	1.4	2.5	5.1	4.4	0.4	1.0	0.6	0.3	.	.
Illness: Physical	23.0	23.0	29.9	20.3	24.8	17.0	3.1	1.9	4.0	4.2	4.9	0.3
Illness: Nonphysical	28.7	31.3	24.4	22.1	28.9	36.8	4.9	5.5	4.9	5.1	5.2	4.1
General health comments	17.7	10.4	18.8	26.2	15.0	18.3	9.7	6.1	9.8	11.8	12.0	8.6
Family history	0.5	0.2	0.0	0.0	0.4	2.1	15.6	16.5	19.2	19.2	12.9	10.9
Life situation	6.8	9.2	5.4	4.8	7.1	7.3	12.0	13.9	11.0	9.6	7.7	18.0
Lack of control	na	na	na	na	na	na	17.9	20.4	16.5	13.4	17.8	21.5
Up to chance	na	na	na	na	na	na	7.3	8.1	7.6	8.9	5.2	6.8
Other	5.8	12.6	3.0	4.7	5.3	3.7	16.1	15.2	17.4	17.9	8.9	21.2
χ^2 test	$\chi^2 = 221.52$; $df = 40$ ($p < .001$)						$\chi^2 = 149.47$; $df = 40$ ($p < .001$)					

[a]Based on all codes assigned to open-ended probing responses.
[b]Based on the first code assigned to open-ended probing responses.

combined was similar across countries, the contribution of each attribute differed significantly. For example, in Spain, nonphysical illness was reported most frequently (36.8%), while it was general health in the United States (26.2%).

SLE drew on a wide range of attributes. Illness, health behaviors, and general health comments combined accounted for 30.7% of SLE, a rate much lower than for SRH (83.3%). Family history and life situations were mentioned 15.6% and 12.0% of the time. Another 17.9% of the respondents expressed their lack of control and 7.3% up to chance. The "other" code (e.g., longevity desires) was given to 16.1% of the responses. SLE attributes diverged across countries. For example, BMI-related and other health behaviors in combination were reported by 25.5% in Mexico but only by 8.6–11.3% in other countries. The other attributes were reported by 8.9% of respondents in Mexico but by 15.2–21.2% in other countries.

Question Difficulty

There was virtually no nonresponse to SRH regardless of response format. SLE, on the other hand, produced a large nonresponse rate at 19.3% with little variation across countries. On 101-point SRH, 68.1% of the answers heaped at multiples of 10s and 25s similarly across countries. However, the proportion of response 50 differed significantly across countries (from 1.6% [Spain] to 7.4% [Germany]; $p < .01$), as well as that of 100 (from 2.9% [Great Britain] to 13.4% [Mexico]; $p < .001$). Overall, 85.8% of the SLE respondents chose a heaping response with different rates across countries (from 80.1% [United States] to 88.8% [Spain]; $p = .004$). The most frequently chosen SLE response was 50 similarly across countries from 12.9% (Spain) to 18.3% (Great Britain). Respondents also reported 100 frequently on SLE (11.8%), with a significant difference across countries (from 8.3% [Great Britain] to 14.3% [Germany]; $p = .025$).

Respondents spent twice as long to respond 101-point than five-point SRH (13.2 vs. 6.3 seconds; $p < .001$). This difference was not the same across countries, resulting in a significant interaction between country and SRH response scale on response time ($p < .001$). As in Figure 2A, 101-point SRH took 13.5 seconds longer than five-point SRH in Mexico, compared to 4- to 6-second differences for other countries. On SLE, respondents spent 16.9 seconds, three seconds longer than nonrespondents did ($p < .001$). However, as in Figure 2B, there was virtually no difference in this time between SLE respondents and nonrespondents in Mexico, while

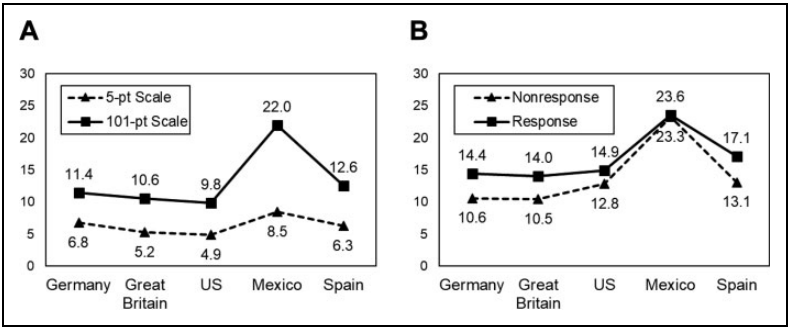

Figure 2. Response time in seconds on 5-point self-rated health (SRH), 101-point SRH, subjective life expectancy (SLE) nonresponse, and SLE response by country. (A) Response time on SRH by response scale and country. (B) Response time on SLE by response status and country.

respondents spent 2–4 seconds longer than nonrespondents in other countries, a significant interaction between country and SLE response status on response time ($p < .001$).

SLE Response Uncertainty

Response uncertainty on SLE was expressed by 37.7% of the respondents who were given the open-ended probing question and by 47.9% of those given the close-ended probing. Multivariate models in Table 3 (model 1 without and model 2 with psychosocial traits) examined SLE response uncertainty. Including psychosocial traits improved the model fit significantly ($\Delta 2\text{Log } L = 23$, $\Delta df = 3$; $p < .001$). Specifically, high future optimism was associated with lowered response uncertainty significantly, whereas high religiosity with increased response uncertainty. Sense of control showed no association with response uncertainty.

Discussion

While SRH and SLE both predict mortality, our study suggests that they elicit different information from respondents and impose varying levels of burden. Respondents incorporated a broader spectrum of attributes on SLE than SRH: Answers to SRH were formed focusing on illness and health behaviors, while additional attributes, such as family history, were considered for SLE. A moderate relationship between SRH and SLE further evidences this. It may be that SRH and SLE, when combined together,

Table 3. Logistic Regression of Uncertainty about Subjective Life Expectancy (SLE) Response.[a]

	Dependent Variable: Uncertainty about SLE Response					
	Model 1: Without Psychosocial Measures			Model 2: With Psychosocial Measures		
	OR	95% CI	*p* Value	*OR*	95% CI	*p* Value
Independent variables						
Future optimism/time orientation	na	na	na	0.86	[0.79–0.93]	.000
Lacking sense of control	na	na	na	1.02	[0.95–1.11]	.568
Religiosity	na	na	na	1.08	[1.01–1.16]	.023
Control variables						
Country (Ref: United States)						
Germany	1.12	[0.88–1.43]	.232	1.20	[0.93–1.55]	.150
Mexico	0.78	[0.61–1.00]	.001	0.84	[0.65–1.09]	.004
Spain	1.22	[0.96–1.56]	.021	1.35	[1.04–1.75]	.006
Great Britain	1.03	[0.81–1.32]	.882	1.04	[0.81–1.33]	.690
Age in years (Ref: ≥60)						
18–29	0.94	[0.71–1.26]	.130	0.96	[0.72–1.27]	.160
30–39	1.07	[0.78–1.45]	.869	1.09	[0.80–1.49]	.711
40–49	1.12	[0.82–1.53]	.428	1.12	[0.82–1.53]	.480
50–59	1.13	[0.84–1.52]	.318	1.12	[0.84–1.51]	.403
Sex: Female vs. male	1.28	[1.09–1.49]	.002	1.23	[1.05–1.44]	.011
Education: <Tertiary vs. ≥tertiary	1.60	[1.37–1.88]	<.001	1.55	[1.32–1.81]	<.001
Model fit	–2 Log $L = 3{,}577$ ($df = 10$)			–2 Log $L = 3{,}554$ ($df = 13$)		

[a]Uncertainty is constructed using responses to SLE and SLE probing questions.

capture a fuller range of information germane to mortality. Developing a new question that blends the measurement mechanisms of SRH and SLE while maintaining the response burden at a minimum will improve our ability to predict mortality more accurately through survey questions.

One may question whether using the 101-point numeric scale is more advantageous than the five-point verbal scale for SRH as the numeric scale eliminates connotations in verbal response categories. Our analysis provides no support. Nearly 70% of the responses to 101-point SRH heaped at multiples of 10s and 25s. Response time was twice longer for 101-point SRH, while respondents used similar information between the two response formats. Thus, we recommend using the five-point rather than the 101-point scale, given the latter's potential burden.

Respondents appeared unsure about how to answer SLE: Almost 20% of respondents did not answer and another 25–30%, although they answered SLE, expressed uncertainty through probing questions. This was not the

case for SRH. Uncertainty around SLE was systematic: lower future optimism/time-orientation and higher religiosity associated with higher uncertainty. Further, relatively small differences in response time between those who gave an answer to SLE and those who did not may suggest that the source of such high nonresponse rates is not respondents' lacking effort but the question difficulty. Future work may test alternatives that may make SLE less burdensome (e.g., a four-point verbal response scale in van Doorn and Kasl 1998).

This study has a number of limitations. First, our opt-in sample does not offer generalizability. However, the focus of our study is to compare response processes between SRH and SLE through probing questions. There is no clear reason why a probability sample would produce different implications on them. Second, some coder reliability was passable at best. Nonetheless, the fact that the majority of coders showed high reliability bolsters our confidence in the overall findings. Future studies will build on the coding scheme to strengthen the reliability. Third, our study did not elaborate on the cross-national comparisons. This was done purposefully because the extant literature does not offer coherent theories to motivate hypotheses for such comparisons. Cross-national differences observed in this study, however, may stimulate exploring theoretical underpinnings and developing hypotheses. Particularly for SLE, time orientation was shown related to the response burden. As cultural backgrounds are believed to be a dominant factor in forming individuals' time orientation (Graham 1981), examining this further may open doors for cross-cultural investigations of SRH and SLE measurement. Hence, despite limitations, we believe that our study provides valuable insight into these two important survey questions and their relevance for predicting mortality.

Acknowledgments

We are grateful for contributions from Stefany Barba, Elias Chandarlis, and Verena Vad for coding open-ended responses.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: Data collection was funded by the

German Research Foundation (grant number: BR 908/5-1, PIs: Braun, Bandilla and Kaczmirek).

Supplemental Material

Supplemental material for this article is available online.

References

Abeles, R. 1991. Sense of control, quality of life, and frail older people. In *The concept and measure of quality of life in the frail elderly*, eds. J. Birren, J. Lubben, J. Rowe, and D. Deutschman, 297–314. San Diego, CA: Academic Press.

Bailis, D. S., A. Segall, and J. G. Chipperfield. 2003. Two views of self-rated general health status. *Social Science & Medicine* 56:203–17.

Behr, D., K. Meitinger, M. Braun, and L. Kaczmirek. 2017. *Web probing—Implementing probing techniques from cognitive interviewing in web surveys with the goal to assess the validity of survey questions. GESIS—Survey Guidelines*. Mannheim, Germany: GESIS—Leibniz-Institute for the Social Sciences.

Bergadaa, M. M. 1990. The role of time in the action of the consumer. *Journal of Consumer Research* 17:289.

de Bresser, J., and A. van Soest. 2019. The predictive power of subjective probabilities: Probabilistic and deterministic polling in the Dutch 2017 election. *Journal of the Royal Statistical Society. Series A: Statistics in Society* 182:443–66.

Bruine de Bruin, W., and K. G. Carman. 2018. Measuring subjective probabilities: The effect of response mode on the use of focal responses, validity, and respondents' evaluations. *Risk Analysis* 38:2128–43.

Carter, E. C., M. E. McCullough, J. Kim-Spoon, C. Corrales, and A. Blake. 2012. Religious people discount the future less. *Evolution and Human Behavior* 33: 224–31.

Couper, M. P., R. Tourangeau, F. G. Conrad, and E. Singer. 2006. Evaluating the effectiveness of visual analog scales. *Social Science Computer Review* 24: 227–45.

Devlin, N. J., and R. Brooks. 2017. EQ-5D and the EuroQol group: Past, present and future. *Applied Health Economics and Health Policy* 15:127–37.

Elinson, J. 1994. Introductory remarks. In *Proceedings of the 1993 NCHS Conference on the Cognitive Aspects of Self-Reported Health Status*, ed. S. Schechter, 2–5. Hyattsville, MD: National Center for Health Statistics.

Ferraro, K. F., and L. R. Wilkinson. 2015. Alternative measures of self-rated health for predicting mortality among older people: Is past or future orientation more important? *The Gerontologist* 55:836–44.

Fienberg, S. E., E. F. Loftus, and J. M. Tanur. 1985. Cognitive aspects of health survey methodology: An overview. *The Milbank Memorial Fund Quarterly. Health and Society* 63:547.

Graham, R. J. 1981. The role of perception of time in consumer research. *Journal of Consumer Research* 7:335.

Griffin, B., V. Loh, and B. Hesketh. 2013. A mental model of factors associated with subjective life expectancy. *Social Science & Medicine* 82:79–86.

Groves, R. M., N. H. Fultz, and E. Martin. 1992. Direct questioning about comprehension in a survey setting. In *Questions about questions: Inquiries into the cognitive bases of surveys*, eds. J. M. Tanur, 49–61. New York: Russell Sage Foundation.

Hennessy, C. H., D. G. Moriarty, M. M. Zack, P. A. Scherr, and R. Brackbill. 1994. Measuring health-related quality of life for public health surveillance. *Public Health Reports* 109:665–72.

Hong, Y.-Y. 2009. A dynamic constructivist approach to culture: Moving from describing culture to explaining culturele. In *Understanding culture: Theory, research, and application*, eds. Y. Wyer, R. S. Chiu, and C.-Y. Hong, 3–23. New York: Psychology Press.

Idler, E. L., and Y. Benyamini. 1997. Self-rated health and mortality: A review of twenty-seven community studies. *Journal of Health and Social Behavior* 38:21–37.

Johnson, T., D. O'Rourke, N. Chavez, S. Sudman, R. Warnecke, L. Lacey, and J. Horm. 1997. Social cognition and responses to survey questions among culturally diverse populations. In *Survey measurement and process quality*, eds. L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, and D. Trewin, 87–113. Hoboken, NJ: John Wiley & Sons.

Juster, F. T. 1964. *Anticipations and purchases*. Princeton, NJ: Princeton University Press.

Jylhä, M. 2011. Self-rated health and subjective survival probabilities as predictors of mortality. In *International handbook of adult mortality*, eds. R. G. Rogers and E. M. Crimmins, 329–44. New York: Springer Science + Business Media.

Kerry, M. J., and S. E. Embretson. 2018. An experimental evaluation of competing age-predictions of future time perspective between workplace and retirement domains. *Frontiers in Psychology* 8:2316.

Larson, J. S. 1999. The conceptualization of health. *Medical Care Research and Review* 56:123–36.

Lee, S. 2014. Self-rated health in health surveys. In *Handbook of health survey methods*, ed. T. P. Johnson, 193–216. Hoboken, NJ: Wiley.

Lee, S., F. Keusch, N. Schwarz, M. Liu, and Z. Tuba Suzer-Gurtekin. 2018. Cross-cultural comparability of response patterns of subjective probability questions. In *Advances in comparative survey method*, eds. T. P. Johnson, B. Pennell, I. A. L. Stoop, and B. Dorer, 455–75. Hoboken, NJ: John Wiley.

Lee, S., M. Liu, and M. Hu. 2017. Relationship between future time orientation and item nonresponse on subjective probability questions: A cross-cultural analysis. *Journal of Cross-Cultural Psychology* 48:698–717.

Lee, S., C. McClain, N. Webster, and S. Han. 2016. Question order sensitivity of subjective well-being measures: Focus on life satisfaction, self-rated health, and subjective life expectancy in survey instruments. *Quality of Life Research* 25: 2497–510.

Lee, S., and N. 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, S., and J. Smith. 2016. Methodological aspects of subjective life expectancy: Effects of culture-specific reporting heterogeneity among older adults in the United States. *Journals of Gerontology—Series B*: *Psychological Sciences and Social Sciences* 71:558–68.

Manski, C. F. 1990. The use of intentions data to predict behavior: A best-case analysis. *Journal of the American Statistical Association* 85:934–40.

Marko, K. Whan, and M. L. Savickas. 1998. Effectiveness of a career time perspective intervention. *Journal of Vocational Behavior* 52:106–19.

McCarthy, M. C., E. Ruiz, B. J. Gale, C. Karam, and N. Moore. 2004. The meaning of health: Perspectives of Anglo and Latino older women. *Health Care for Women International* 25:950–69.

Mirowsky, J., and C. E. Ross. 2000. Socioeconomic status and subjective life expectancy. *Social Psychology Quarterly* 63:133.

Perozek, M. 2008. Using subjective expectations to forecast longevity: Do survey respondents know something we don't know? *Demography* 45:95–113.

Siegel, M., E. H. Bradley, and S. V. Kasl. 2003. Self-rated life expectancy as a predictor of mortality: Evidence from the HRS and AHEAD surveys. *Gerontology* 49:265–71.

Stone, A. A., J. S. Turkkan, C. A. Bachrach, J. B. Jobe, H. S. Kurtzman, and V. S. Cain, Eds. 2000. *The science of self-report: Implications for research and practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

van Doorn, C., and S. V. Kasl. 1998. Can parental longevity and self-rated life expectancy predict mortality among older persons? Results from an Australian cohort. *The Journals of Gerontology—Series B: Psychological Sciences and Social Sciences* 53B: S28–34.

World Health Organization. 2017. *World health statistics 2016: Monitoring health for the SDGs, sustainable development goals*. Geneva: World Health Organization.

Zick, C. D., K. R. Smith, R. N. Mayer, and L. B. Taylor. 2014. Family, frailty, and fatal futures? Own-health and family-health predictors of subjective life expectancy. *Research on Aging* 36:244–66.