Using Verbal and Paralinguistic Behaviors to Explain Variation in Responses to Self-Reported Health Questions

Dana Garbarski[1,2], Nora Cate Schaeffer[1,3], Jennifer Dykema[3]
[1]Department of Sociology, University of Wisconsin-Madison, 8128 William H. Sewell Social Science Building, 1180 Observatory Drive, Madison, WI 53706
[2] Department of Population Health Sciences, University of Wisconsin-Madison, 707 WARF Building, 610 N. Walnut St., Madison, WI 53726
[3]University of Wisconsin Survey Center, 1800 University Ave., Madison, WI 53726

Abstract

The self-reported health item summarizes a wide range of information on health status across several domains of health, and is widely used to measure health because it predicts mortality well (Idler and Benyamini, 1997). We examine whether the characteristics of the interviewer-respondent interaction during the administration of the self-reported health item reflect in part actual health status and thus may be useful indirect measures of health status when limited data on health are available.

We find that the odds of several characteristics of the interaction are significantly lower for those with better self-reported health. We also find that reporting a larger number of health conditions is associated with higher odds of each of the characteristics of the interaction we examine, and that many of these relationships remain significant after controlling for the respondent's self-reported health and cognitive ability. Finally, we find some evidence that the characteristics of the interaction may be able to explain variation in health status among respondents with the same self-reported health, in that most of the characteristics of the interaction we examine are associated with the number of health conditions reported among respondents whose self-reported health is "excellent."

Overall, we find promising evidence that the characteristics of the interaction might reveal important information about respondents' health status not captured by the self-reported health question. We recommend further investigations of whether researchers using self-reported health to measure health status would benefit from coding the characteristics of the interaction as an additional, indirect measure of health status.

Key Words: self-reported health, interaction coding, interviewer-respondent interaction, cognitive ability, health, response latency, uncertainty, speech tokens, disfluencies of speech, paradigmatic question-answer sequence

1. Introduction

The self-reported health item summarizes a wide range of information on health status across several domains of health, and is widely used to measure health because it predicts mortality well (Idler and Benyamini, 1997). Researchers have demonstrated that self-reported health is related to several domains of life including illnesses, symptoms of

undiagnosed diseases, judgments about the severity of illness, family history, dynamic health trajectory, behaviors, and the presence or absence of resources for good health (Idler and Benyamini, 1997). In sum, "a very long list of variables is required to explain the effect of one brief 4- or 5-point scale item…" (Idler and Benyamini, 1997: 31).

1.1 Studies of Dimensions of Health Associated with Self-Reported Health

There are two broad sets of studies that investigate what dimensions of health respondents think about when they answer the self-reported health item: studies that investigate the associations between self-reported health and other measures of health, and those that use open and closed followup probes to ask respondents what they were thinking when they answered the self-reported health item.

Studies that investigate covariates of self-reported health ask about self-reported health along with other health and well-being items, investigate which of the covariates are more strongly associated with self-reported health, and infer that these are weighted more heavily by respondents as they construct an answer. These studies have found the following: current health experience is more strongly associated with self-reported health than is prior health experience (Benyamini, Leventhal, Leventhal, 1999); indicators of positive health are as important in determining future and current self-reported health as negative indicators (Benyamini, Idler, Leventhal, Leventhal, 2000); and men's self-reported health is associated with serious, life threatening diseases while women's self-reported health is associated with both life-threatening and non-life-threatening diseases (Benyamini, Leventhal, Leventhal, 2000). In addition, the dimensions of health that respondents rate as important for their self-reported health vary by the answer category (e.g., "excellent," "very good," "good," "fair" or "poor") they select. For example, respondents who selected "poor" or "fair" rated current disease highly while those choosing "excellent," "very good" and "good" rated risk factors and positive indicators highly; respondents in all categories rated overall functioning and vitality factors high (Benyamini, Leventhal and Leventhal, 2003).

Studies that follow the self-reported health item with open or closed followup probes seek to ascertain how respondents construct an answer. Groves, Fultz, and Martin (1992) found using an open probe that respondents who reported using external cues (absence or presence of illness, health service usage, and outcome of physical exam) were slightly less likely than those who used internal cues (feelings, physical performance/ability, affect) to report "excellent," "fair," or "poor" health. For the closed probes, those who reported that they considered their health in recent time were less likely to report "excellent" and more likely to report "good" or "fair" than those who reported that they considered their health over the past few years. Krause and Jay (1994) reported that overall there was not a significant relationship between self-reported health and the content of the respondent's report to the open-ended followup probes. However, interesting patterns emerged in the data: respondents who reported that they compared their health to others were more likely to report "excellent," while respondents whose health referents were physical functioning, health problems, or health behaviors were more likely to report "good." In cognitive testing of the self-reported health item, Canfield and colleagues (2003) reported that respondents considered a variety of situational factors in constructing an answer including reporting "good" health despite a long list of serious health conditions, weighing how well a condition was controlled by medication, comparing themselves with others their own age or with similar medical conditions, or considering a prior question about physical limitations.

In studies which use probes to ask respondents what they were thinking when they answered the self-reported health item, respondents may report only the factors that are most salient to them at the moment or the factors they think are most important to report in an interview. Also, studies that investigate possible correlates of self-reported health do not systematically reveal how respondents conceptualize their health. We present evidence that post-survey coding of the characteristics of the interviewer-respondent interaction occurring during the administration of the self-reported health item provides an indirect way to capture information on respondents' health status beyond their answer to the self-reported health item.

1.2 Model of the Response Process and Characteristics of the Interaction

Information-processing models of the response process view the process of constructing as answer to a survey question as having four stages: comprehension of the item, retrieval of relevant information to answer the item from memory, use of retrieved information to make judgments, and selection and reporting of an answer to the item (Tourangeau, Rips, and Rasinski, 2000). During these four stages the respondent uses various cognitive tools to answer a survey item, but the process can vary: stages can overlap, or the respondent can backtrack to an earlier stage and skip or truncate one or more of the subsequent stages (Tourangeau, Rips, and Rasinski, 2000). Which stages a respondent uses and how depends on several factors, such as the wording of the item, how accurate the respondent wants to be, and the respondent's cognitive ability.

The circumstances of the respondent's life and experience also affect the response process and the difficulty of constructing an answer. In the case of the self-reported health item, the respondent's actual health can affect each of the four stages. For example, poor health may make it hard to comprehend the item or may influence the respondent's interpretation of "health." The complexity of the respondent's actual health status may require that the respondent combine disparate types of information, such as disease and functioning, which can make it more difficult to retrieve relevant information, formulate judgments about health, and map an answer onto the response categories.

Some behaviors of the interviewer and respondent during the question-answer process -- such as response latencies, expressions of uncertainty by respondents, and probing of respondents by interviewers -- have been shown to be associated with inaccurate or unreliable answers, the difficulty of the task, and cognitive ability (Draisma and Dijkstra, 2004; Dykema, Lepkowski and Blixt, 1997; Fowler and Cannell, 1996; Hess, Singer, and Bushery, 1999; Knäuper et al., 1997; Mathiowetz, 1999; Schaeffer and Dykema, 2004; Schaeffer et al., 2008; van der Zouwen and Smit, 2004). These characteristics of the interaction can also be viewed as by-products of the information processing that occurs when a respondent answers a survey item (Fowler and Cannell, 1996; Holbrook, Cho, and Johnson, 2006). Using factor analyses, Holbrook, Cho, and Johnson (2006) concluded that certain respondent behaviors occurring during the administration of an item were associated with distinct factors that could be broadly classified as comprehension difficulty (e.g., requests for clarification) and mapping difficulty (e.g., inadequate or imprecise responses).

1.3 Conceptual Model of Self-Reported Health

Our conceptual model is depicted in Figure 1. The diagram indicates that the response process is affected by both the respondent's cognitive ability and their actual health,

which may also affect each other. The response process and actual health affect the respondent's answer to the self-reported health question, and the response process and cognitive ability affect the characteristics of the interaction. Finally, self-reported health and the characteristics of the interaction are associated because they are co-produced by the respondent after the item is posed by the interviewer.

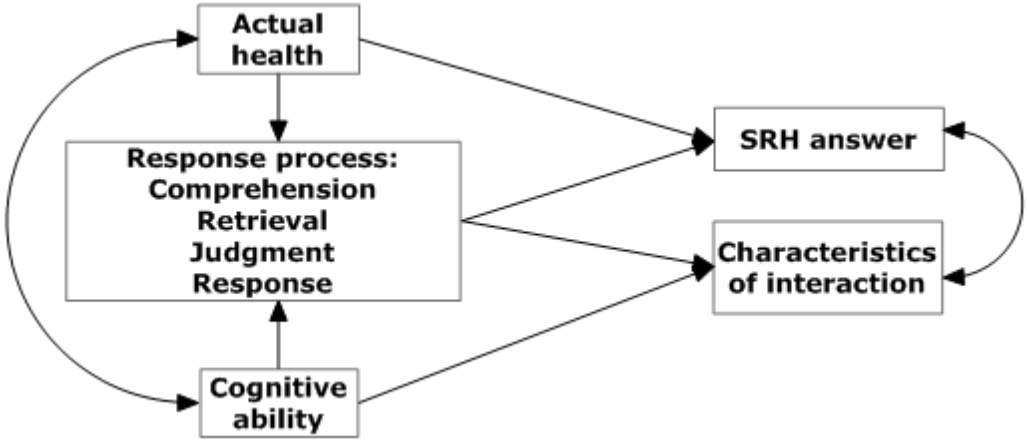

Figure 1: Conceptual model

Prior research has shown that some characteristics of the interaction are associated with cognitive ability as measured by performance on cognitive assessments (e.g., Schaeffer et al. 2008). For self-reported health, our model suggests that this association could include both the general effect of cognitive ability on the characteristics of the interaction, as well as the effect that cognitive ability has on the response processing of this particular question. We also expect that actual health will influence characteristics of the interaction in three ways: by implicating a specific answer to the self-reported health item, by the effects of actual health on cognitive ability, and by the impact of the respondent's actual health status on the processing of the question (e.g., difficulties in combining information about disease and functioning or difficulties in mapping a resulting judgment onto one of the offered categories). Because we expect that the response process is affected in part by both cognitive ability and the respondent's actual health, we consider whether the characteristics of the interaction may in part reflect actual health status and, as a result, may provide useful proxy indicators of health status when limited data on health are collected. For example, when the self-reported health question is the only measure of health, adding information about the interaction during which the answer is produced might provide either additional information about actual health or information about measurement error in the answer to the self-reported health question. A similar argument is made by Mathiowetz (1998, 1999), who found that response behavior was related to actual health care utilization (number of outpatient visits) and reporting accuracy (verified by records), with exact reporters providing significantly lower mean number of visits and more accurate reports of the number of visits than qualified reporters or reporters whose initial response required interviewers to use scripted probes.

1.4 Hypotheses

We examine a distinct set of interactional characteristics that are indicators of potential problems in the question-answer process, including use of tokens or expressions of uncertainty by the respondent, pre-emptive behaviors by the interviewer, whether the

question-answer sequence is paradigmatic, the number of exchanges in the question-answer sequence, and response latencies (see later). These problematic behaviors are indicators of potential problems in the question-answer process. We predict that:

> Hypothesis 1: The problematic behaviors that accompany answers to the self-reported health question will vary with the responses to the self-reported health question because these behaviors are co-produced with the answer. In particular, we expect that the levels of these problematic behaviors will be lower when respondents report higher levels of self-reported health because healthier respondents have a simpler response task and are less likely to have cognitive abilities impaired by poor health.

> Hypothesis 2: The characteristics of the interaction will be related to an indicator of actual health status, the number of health conditions reported by the respondent. We further expect this association to remain after controlling for self-reported health and a measure of cognitive ability.

> Hypothesis 3: The characteristics of the interaction will be associated with the number of health conditions among respondents who select the same response category for the self-reported health question.

2. Methods

2.1 Data

We analyze interviewer-respondent interaction produced during the administration of the self-reported health item and several other questions about physical and mental health from digitally recorded interviews of older adults in the 2004 wave of the Wisconsin Longitudinal Study (WLS), a 1/3 random sample of the Wisconsin high school class of 1957. From a random subset of the total sample, we drew a sample of 100 interviewers and stratified respondents within interviewer according to the respondent's IQ (high, medium, and low and measured while the respondent was in high school). We then attempted to sample 5 respondents for each interviewer. Our analysis includes 355 coded cases.

We examine the interaction during the administration of the self-reported health question, beginning with the interviewer's reading of the question and ending with the respondent's final answer. We coded the interaction using an elaborate scheme that segmented the interaction into a series of events and assigned codes to very specific behaviors, including response latencies, disfluencies of both parties, and task behaviors from the interviewer, such as probing respondents. We used Sequence Viewer (Wil Dijkstra, http://www.sequenceviewer.nl/) to code each event.

2.2 Measures

We examine six features of the interaction that may indicate problems in cognitive processing, arising either from the respondent's actual health or cognitive ability.

Tokens such as "um," "uh" or "well" are sometimes labeled disfluencies, and they have been interpreted as indicating disruption in the speaker's cognitive processing (e.g., Bortfeld et al., 2001). Some tokens have also been described as continuers, because they appear to respond to a prior utterance in a way that signals to the speaker to continue

(Schegloff, 1981). We include an indicator for whether the respondent uttered any tokens (0=none, 1=1 or more).

Respondents express various kinds of uncertainty in many ways, and we use an indicator of whether the respondent exhibited behaviors that may express uncertainty (0=none, 1=1 or more behaviors). The behaviors we count include hypothetical response options (e.g., the respondent says "pretty good" even though that category is not offered), ranges (e.g., "good to very good"), reports (e.g., "my mental health is ok, my physical health is not"), and mitigating phrases offered as answers or parts of answers (e.g., "I guess excellent" or "just," "maybe," "about," "put," or "I'd say").

Interviewers often intervene in order to obtain a codable answer from a respondent before the respondent has answered, and such pre-emptive behaviors by interviewers may indicate a respondent is having difficulty in answering (see Schaeffer and Maynard, 2002). For example, if the respondent hesitates, the interviewer might probe with the response categories in order to obtain a codable answer from the respondent (0=no pre-emptive behavior from the interviewer, 1=any such behavior).

Paradigmatic sequences may be described as interactions in which the interviewer's administration of the item is directly followed by the respondent's answer (with or without a preceding pause). The sequence may end there or the interviewer may say "okay" or repeat the respondent's answer before the sequence ends (Schaeffer and Maynard 1996). Any sequence that does not follow this pattern is non-paradigmatic (0=paradigmatic, 1=non-paradigmatic). The number of exchanges between the interviewer and respondent that occur in a question-answer sequence has also been found to be associated with lower data quality (Schaeffer and Dykema, 2004; see also van der Zouwen and Smit, 2004), and so we include a dummy variable for whether the interaction has more than one exchange or not (0=no, 1=yes).

Response latency, the time in seconds from the end of the interviewer's reading of the question until the respondent's first complete codable answer, has been used as an indicator of cognitive processing time, and in some cases longer latencies are associated with lower (or higher) data quality (Ehlen, Schober, and Conrad, 2007; Schaeffer and Dykema, 2004; Schaeffer et al. 2008). For this analysis, the response latency ends when the respondent provides one of the response options (excellent, very good, good, fair or poor).

The main independent variable is the answer to the self-reported health question. Of the 355 respondents in this analysis, 102 reported "excellent" health, 125 reported "very good," 95 reported "good," 24 reported "fair," and 9 reported "poor." "Fair" and "poor" health were combined into one category for this analysis because of the low number of respondents in each category. Self-reported health is coded so that a higher value means better health (fair/poor=1 to excellent=4).

In some analyses we control for actual health status, indicated by the number of health conditions reported by the respondent in the 2004 wave of the WLS survey. The seven health conditions include arthritis, diabetes, high blood pressure, heart conditions (current or in past), cancer, stroke, and high blood sugar. Respondents' reports ranged from 0 to 5 health conditions, with a mean of 1.37 conditions (1.09 SD). Our measure of cognitive ability is the respondent's IQ measured in high school.

2.3 Analysis

In Section 3.1, we regress each characteristic of the interaction on self-reported health to investigate Hypothesis 1, the hypothesis that the characteristics of the interaction vary across the self-reported health answer categories. Because most of the characteristics of the interaction are dichotomous variables, we use logistic regression and report the odds ratios (the exception is response latency, where we perform OLS regression and report the coefficient). In Section 3.2, we regress each of the characteristics of the interaction on the number of health conditions reported by the respondent to investigate Hypothesis 2, the hypothesis that the characteristics of the interaction are associated with actual health status. We then perform the same regressions while controlling for the respondent's self-reported health, and again controlling for the respondent's self-reported health and IQ, in order to determine whether the characteristics of the interaction predict actual health status beyond their association with the respondent's self-reported health and cognitive ability. In Section 3.3, we examine the correlations between actual health status and characteristics of the interaction to investigate Hypothesis 3, the hypothesis that the characteristics of the interaction are associated with variation in actual health status among respondents with the same self-reported health.

3. Results

3.1 Characteristics of the Interaction by Self-Reported Health

First we investigate whether the behaviors that accompany self-reported health vary by the self-reported health answer given. Because the answer to the self-reported health question and the characteristics of the interaction are co-produced, we expect that the behaviors that accompany self-reported health answers vary by the self-reported health answer given. We predict that there will be fewer potentially problematic behaviors when respondents report better health.

Figure 2 shows that in general, the proportion of question-answer sequences with each characteristic is lower for respondents with better self-reported health. The exceptions are having any uncertainty indicators and a non-paradigmatic sequence, for which the proportions are lower for "fair/poor" compared to "good." Figure 3 shows that the means and standard deviations of response latencies are smaller for respondents with better self-reported health.

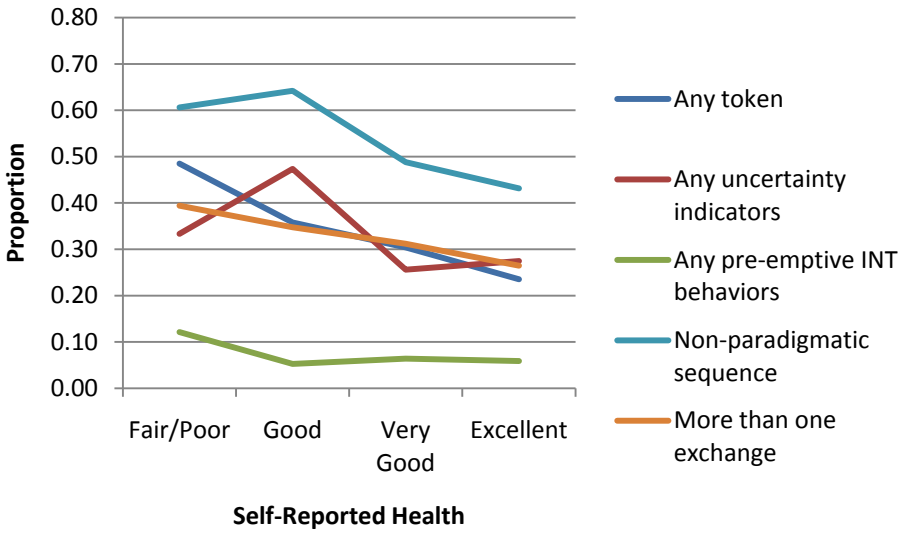

Figure 2: Proportion of question-answer sequences with behavior by answer to the self-reported health question

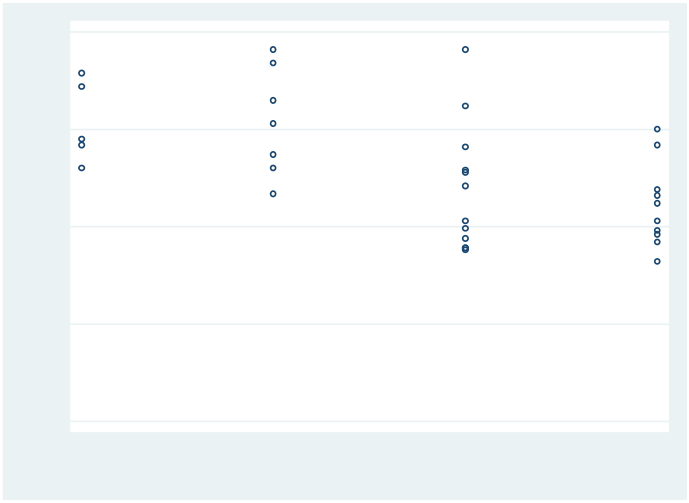

Figure 3: Mean-standard deviation plot of response latency by self-reported health answer categories

Bivariate regression analyses demonstrate that the odds of each characteristic of the interaction significantly decrease for those with better self-reported health, with the exceptions of more than one exchange and any pre-emptive interviewer behaviors (see Table 1). Likelihood-ratio tests (not shown) were conducted to determine whether self-reported health can be treated as an interval measure with evenly spaced categories or whether this leads to a loss of information about the association between the independent and dependent variables. The likelihood-ratio tests compared models where self-reported health is an ordinal variable with models where self-reported health is included as an ordinal variable as well as dummy variables for each answer category (omitting two) (Long and Freese, 2001). The likelihood-ratio tests show that the self-reported health

answer categories can be treated as evenly spaced for all regressions except when the dependent variable is any uncertainty indicators. Subsequent logistic regressions of any uncertainty indicators on dummy variables for the self-reported health answer categories show that the odds of having any uncertainty indicators are significantly higher when self-reported health is "good" compared to "excellent" and "good" compared to "very good," while all other differences among self-reported health categories are not significant. In other words, self-reported health has a linear relationship with all of the characteristics of the interaction except for any uncertainty indicators, where the relationship appears to be curvilinear.

Table 1. Regressions of Characteristics of the Interaction on Self-Reported Health

	Odds Ratio	95% CI
Any token [vs. none]	0.71**	0.56-0.90
Any uncertainty indicators [vs. none]	0.76*	0.60-0.96
Any pre-emptive INT behaviors [vs. none]	0.85	0.55-1.32
Non-paradigmatic sequence [vs. paradigmatic]	0.72**	0.57-0.90
More than one exchange [vs. one exchange]	0.82	0.65-1.04
	Coefficient	95% CI
Response latency (seconds)	-0.52***	-0.78--0.26

***p<.001, **p<.01, *p<.05, †p<.1

3.2 Predicting Characteristics of the Interaction during Self-Reported Health from Actual Health

In the previous section we showed that there are some differences in the characteristics of the interaction across categories of self-reported health. Next, we examine whether the respondent's actual health predicts the characteristics of the interaction during the self-reported health question, and whether the association remains when controlling for the answer ultimately produced. If this relationship is demonstrated, then researchers might improve their measurement of self-reported health by augmenting the answer to that question with measures of characteristics of the interaction during which the answer is produced. We also control for IQ as a measure of cognitive ability, because we expect the characteristics of the interaction to reflect cognitive ability as well as actual health.

When we predict characteristics of the interaction during the self-reported health question from the number of health conditions, we find that there are higher odds of each of the characteristics of the interaction and longer response latencies for those with more health conditions (see Table 2). When self-reported health is added to the model, the relationship between the health conditions reported and each characteristic of the interaction is attenuated, but still significant or marginally significant (except for any tokens). Controlling for IQ does not lead to any changes in interpretation. This analysis lends evidence to the hypothesis that these characteristics of the interaction are associated with actual health (indexed by the number of health conditions reported), and picking up something about actual health that self-reported health alone does not**.**

Table 2. Regression of Characteristics of Interaction on the Number of Health Conditions

	Bivariate		Controlling for SRH		Controlling for SRH, IQ	
	Odds Ratio	95% CI	Odds Ratio	95% CI	Odds Ratio	95% CI
Any token [vs. none]	1.24*	1.01 -1.52	1.11	0.89 -1.40	1.11	0.88 -1.39
Any uncertainty indicators [vs. none]	1.43***	1.16 -1.76	1.37**	1.09 -1.72	1.36**	1.08 -1.72
Any pre-emptive INT behaviors [vs. none]	1.42†	0.99 -2.04	1.44†	0.95 -2.17	1.43†	0.95 -2.17
Non-paradigmatic sequence [vs. paradigmatic]	1.50***	1.22 -1.84	1.41**	1.13 -1.77	1.41**	1.12 -1.77
More than one exchange [vs. one exchange]	1.29*	1.05 -1.58	1.25†	0.99 -1.56	1.24†	0.99 -1.56
	Coefficient	95% CI	Coefficient	95% CI	Coefficient	95% CI
Response latency (seconds)	0.40***	0.17 -0.63	0.25†	-0.01 -0.50	0.24†	-0.01 -0.50

***p<.001, **p<.01, *p<.05, †p<.1

3.3 Association between Characteristics of the Interaction and Number of Health Conditions within Category of Self-Reported Health

Another way to determine whether characteristics of the interaction have predictive power beyond that of self-reported health is to determine whether the characteristics of the interaction are associated with variation in actual health (indexed by the number of health conditions) among respondents who have the same self-reported health. We have already seen that between categories of the self-reported health question there are differences in characteristics of the interaction (Figure 2), and Figure 4 shows that within categories of the self-reported health question there is variation in the number of health conditions. It is plausible to expect that the characteristics of the interaction may capture some of the variation in health status among respondents with the same level of self-reported health.

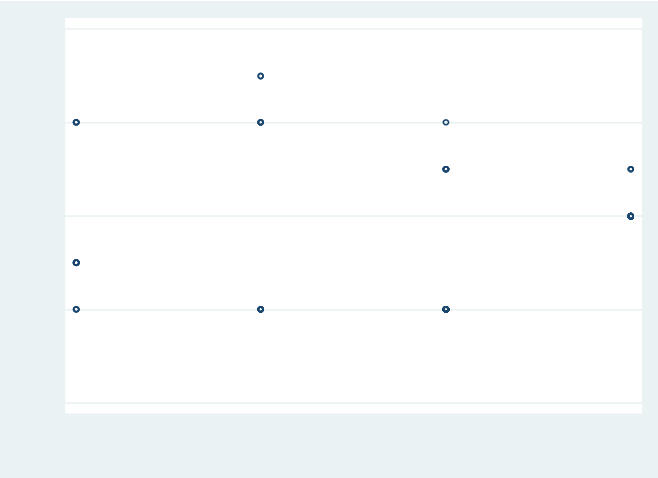

Figure 4: Mean-SD plot of number of health conditions by self-reported health answer categories

Table 3 shows that among respondents who report "excellent" health, the characteristics of the interaction (except for the use of any tokens) are significantly or marginally significantly associated with the number of health conditions reported in the expected direction. Even though all of these respondents answered "excellent," the characteristics of the interaction may reveal some variation in health status (in this case the number of health conditions reported) among these respondents. A few of the characteristics of the interaction are associated with the number of health conditions reported among respondents who answered "very good" and "good" as well.

Table 3. Correlations of Characteristics of the Interaction with the Number of Health Conditions by Self-Reported Health

	Excellent (N=102)	Very Good (N=125)	Good (N=95)	Fair/Poor (N=33)
Any token [vs. none]	0.08	0.17†	0.00	-0.18
Any uncertainty indicators [vs. none]	0.17†	0.03	0.19*	0.15
Any pre-emptive INT behaviors [vs. none]	0.24*	0.02	0.06	0.09
Non-paradigmatic sequence [vs. paradigmatic]	0.17†	0.25**	0.11	-0.01
More than one exchange [vs. one exchange]	0.17†	0.19*	-0.06	0.17
Response latency (seconds)	0.18†	0.04	0.11	0.11

***p<.001, **p<.01, *p<.05, †p<.1

4. Discussion

Answers to the self-reported health question and the characteristics of the interaction are co-produced, and we expected more paradigmatic sequences with fewer potentially problematic behaviors when respondents reported better self-reported health. This is what we found for the use of tokens and uncertainty expressions by the respondent, the presence of a paradigmatic question-answer sequence, and response latency.

Furthermore, these differences in behavior appear to be related to a measure of actual health status. A larger number of health conditions reported by the respondent is associated with higher odds of each of the characteristics of the interaction. Many of these relationships remain significant, although some are attenuated, after controlling for self-reported health, indicating that the characteristics of the interaction explain some of the variation in health status beyond that explained by self-reported health alone. Thus, the characteristics of the interaction might increase the predictive power of self-reported health, and surveys that use self-reported health as the only indicator of health might benefit from augmenting this answer by coding the characteristics of the interaction.

We also found that the characteristics of the interaction might capture some of the variation in health status among respondents with the same self-reported health answer. Most of the characteristics of the interaction we examined are associated with the number of health conditions reported by the respondent among respondents who reported "excellent" health, indicating that characteristics of the interaction may be able explain some of the variation in health status among those with the same self-reported health.

Overall, we find promising evidence that the characteristics of the interaction might reveal important information about the respondent's health status. When limited information about health is collected for a survey, survey researchers using self-reported health as a measure of health might benefit from coding the characteristics of the interaction as an indirect measure of health status. Future analyses will explore whether using characteristics of the interaction as an additional control variable leads to improvements in estimation.

Acknowledgements

This research uses data from the Wisconsin Longitudinal Study (WLS) of the University of Wisconsin-Madison. Since 1991, the WLS has been supported principally by the National Institute on Aging (AG-9775 and AG-21079), with additional support from the Vilas Estate Trust, the National Science Foundation, the Spencer Foundation, and the Graduate School of the University of Wisconsin-Madison. A public use file of data from the Wisconsin Longitudinal Study is available from the Wisconsin Longitudinal Study, University of Wisconsin-Madison, 1180 Observatory Drive, Madison, Wisconsin 53706 and at http://www.ssc.wisc.edu/wlsresearch/data/. The opinions expressed herein are those of the authors.

This research was supported by National Institute on Aging grant R01 AG0123456, and by core grants to the Center for Demography and Ecology at the University of Wisconsin-Madison (R24 HD047873) and to the Center for Demography of Health and Aging at the University of Wisconsin-Madison (P30 AG017266).

References

Benyamini, Yael, Ellen L. Idler, Howard Leventhal, and Elaine A. Leventhal. 2000. "Positive Affect and Function as Influences on Self-Assessments of Health Expanding Our View Beyond Illness and Disability." *Journals of Gerontology Series B: Psychological Sciences and Social Sciences* 55(2): 107-116.

Benyamini, Yael, Elaine A. Leventhal, and Howard Leventhal. 1999. "Self-Assessments of Health: What Do People Know That Predicts Their Mortality?" *Research on Aging* 21(3): 477.

Benyamini, Yael, Elaine A. Leventhal, and Howard Leventhal. 2000. "Gender Differences in Processing Information for Making Self-Assessments of Health." *Psychosomatic Medicine* 62: 354-364.

Benyamini, Yael, Elaine A. Leventhal, and Howard Leventhal. 2003. "Elderly People's Ratings of the Importance of Health-Related Factors to Their Self-Assessments of Health." *Social Science & Medicine* 56(8): 1661-1667.

Bortfeld, Heather, Silvia D. Leon, Jonathan E. Bloom, Michael F. Schober, and Susan E. Brennan. 2001. "Disfluency Rates in Conversation: Effects of Age, Relationship, Topic, Role, and Gender." *Language and Speech* 44:123-49.

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. Downloaded from: http://wwwn.cdc.gov/QBANK/report%5CCanfield_NCHS_2003AdultNHISReport.pdf

Draisma, Stasja and Wil Dijkstra. 2004. "Response Latency and (Para) Linguistic Expressions as Indicators of Response Error." Pp. 131-48 in *Methods for Testing and Evaluating Survey Questionnaires*, edited by Stanley Presser, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, and Eleanor Singer. New York: Springer-Verlag.

Dykema, Jennifer, James M. Lepkowski, and Steven Blixt. 1997. "The Effect of Interviewer and Respondent Behavior on Data Quality: Analysis of Interaction Coding in a Validation Study." Pp. 287-310 in *Survey Measurement and Process Quality*, edited by Lars Lyberg, Paul Biemer, Martin Collins, Edith de Leeuw, Cathryn Dippo, Norbert Schwarz, and Dennis Trewin. New York: Wiley-Interscience.

Ehlen, Patrick, Michael F. Schober, and Frederick G. Conrad. 2007. "Modeling Speech Disfluency to Predict Conceptual Misalignment in Speech Survey Interfaces." *Discourse Processes,* 44(3):245-65.

Fowler Jr., Floyd J. and Charles F. Cannell. 1996. "Using Behavioral Coding to Identify Cognitive Problems with Survey Questions." Pp. 15-36 in *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, edited by Norbert Schwarz and Seymour Sudman. San Francisco, CA: Jossey-Bass.

Groves, Robert M., Nancy H. Fultz, and Elizabeth Martin. 1991. "Direct Questioning About Comprehension in a Survey Setting." Pp. 49-61 in *Questions About Questions: Inquiries into the Cognitive Bases of Surveys*, edited by Judith M. Tanur. New York: Russell Sage Foundation.

Hess, Jennifer, Eleanor Singer, and John M. Bushery. 1999. "Predicting Test-Retest Reliability from Behavior Coding." *International Journal of Public Opinion Research* 11(4):346-60.

Holbrook, Allyson, Young Ik Cho, and Timothy Johnson. 2006. "The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties." *Public Opinion Quarterly* 70(4): 565-95.

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(1): 21-37.

Knäuper, Barbel, Robert F. Belli, Daniel H. Hill, and A.R. Herzog. 1997. "Question Difficulty and Respondents' Cognitive Ability: The Effect on Data Quality." *Journal of Official Statistics* 13(2): 181-199.

Krause, Neal M. and Gina M. Jay. 1994. "What Do Global Self-Rated Health Items Measure?" *Medical Care* 32(9): 930-942.

Long, J. Scott and Jeremy Freese. 2001. *Regression Models for Categorical Dependent Variables Using Stata.* Stata press: College Station, TX.

Mathiowetz, Nancy A. 1998. "Respondent Expressions of Uncertainty: Data Source for Imputation." *Public Opinion Quarterly* 62(1): 47-56.

Mathiowetz, Nancy A. 1999. “Respondent Uncertainty as Indicator of Response Quality.” *International Journal of Public Opinion Research* 11(3): 289-296.

Schaeffer, Nora Cate and Jennifer Dykema. 2004. “A Multiple-Method Approach to Improving the Clarity of Closely Related Concepts: Distinguishing Legal and Physical Custody of Children.” Pp. 475-502 in *Methods for Testing and Evaluating Survey Questionnaires*, edited by Stanley Presser, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, and Eleanor Singer. New York: Springer-Verlag.

Schaeffer, Nora Cate, Jennifer Dykema, Dana Garbarski, and Douglas W. Maynard. 2008. "Verbal and Paralinguistic Behaviors in Cognitive Assessments in a Survey Interview." In *JSM Proceedings*, Statistical Computing Section. Alexandria, VA: American Statistical Association. 4344-4351.

Schaeffer, Nora Cate and Douglas W. Maynard. 1996. “From Paradigm to Prototype and Back Again: Interactive Aspects of Cognitive Processing in Survey Interviews.” Pp. 65-88 in *Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research*, eds. Norbert Schwarz and Seymour Sudman. San Francisco: Jossey-Bass.

Schaeffer, Nora Cate and Douglas W. Maynard. 2002. "Occasions for Intervention: Interactional Resources for Comprehension in Standardized Survey Interviews." Pp. 261-80 in *Standardization and Tacit Knowledge: Interaction and Practice in the Survey Interview*, edited by Douglas W. Maynard, Hanneke Houtkoop-Steenstra, Nora C. Schaeffer, and Johannes van der Zouwen. New York: Wiley.

Schegloff, Emanuel A. 1981. "Discourse as an Interactional Achievement: Some Uses of 'uh huh' and Other Things That Come Between Sentences." Pp. 71-93 in *Analyzing Discourse (Georgetown University Roundtable on Languages and Linguistics 1981)*, edited by Tannen, D. Washington D.C.: Georgetown University Press.

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response.* U.K.: Cambridge University Press.

van der Zouwen, Johannes and Johannes H. Smit. 2004. "Evaluating Survey Questions by Analyzing Patterns of Behavior Codes and Question-Answer Sequences: A Diagnostic Approach." Pp. 109-30 in *Methods for Testing and Evaluating Survey Questionnaires*, edited by Stanley Presser, Jennifer M. Rothgeb, Mick P. Couper, Judith T. Lessler, Elizabeth Martin, Jean Martin, and Eleanor Singer. New York: Springer-Verlag.

Wisconsin Longitudinal Study (WLS) [graduates]: 2004 Version 12.10. [machine-readable data file] / Hauser, Robert M; Hauser, Taissa S. [principal investigator(s)]. Madison, WI: University of Wisconsin-Madison, WLS. [distributor]; <http://www.ssc.wisc.edu/wlsresearch/documentation/>