Dissertation Defense: Matt Jans

CAN SPEECH CUES AND VOICE QUALITIES PREDICT ITEM NONRESPONSE IN INCOME QUESTIONS

October 30, 2009

12:00 noon

368 ISR and 1208 LeFrak

Co-Chairs: Frederick G. Conrad and James M. Lepkowski

Income nonresponse is a significant problem in survey data, with rates as high as 50%, yet we know little about why it occurs. It is plausible that the *way* respondents answer survey questions (their voice and speech characteristics, and their question-answering behavior) can predict whether they will provide income data, and will reflect the psychological states that produce this decision. Five questions from 185 recorded interviews conducted by the Surveys of Consumers were selected. One was the annual household income question. Exchanges between interviewers and respondents were transcribed and coded for respondent speech and question answering behavior. Voice pitch was extracted mechanically using the Praat software. These speech, voice, and question answering behaviors are used as verbal paradata; characteristics of the survey process that are not captured by default. Verbal paradata are hypothesized to reflect respondents' affective and cognitive states, which then predict income nonresponse. It was hypothesized that indicators of respondent affect (e.g., pitch) and cognitive difficulty (e.g., disfluency) would be affected by sensitive and complex questions differently, and would predict whether respondents provide income in dollar amount, brackets, or not at all. Results show that verbal paradata can distinguish between income nonrespondents and respondents, even when only using verbal paradata that occur before the income question. Income nonrespondents have lower affective involvement and express more negativity before the income question. Bracketed respondents express more signs of cognitive difficulty. Income nonresponse is predicted by behavior before the income question, while bracketed response is predicted by indicators on the income question itself. Further, question characteristics affected respondent paradata, but largely in unpredicted ways. There is evidence for a conversationality mechanism through which respondents *reduce* verbal paradata when questions are demanding, rather than increasing it as signs of trouble. The results have implications for theory of income nonresponse, specifically the role of question characteristics and respondent paradata in understanding what subjective psychological states respondents are experiences when they answer survey questions, and how those states predict whether income is reported. There are also potential extensions to interviewer training and design interventions that could produce more complete income data.