

	<b>Groningen Symposium on Language and Social Interaction</b>	
--	---	--

*Thursday, 19 January 2017, Marie Lokezaal, [Harmoniebuilding](#)  
University of Groningen, Oude Kijk in 't Jatstraat 26, Groningen*

Thu 19/1	<b>Keynote Nora Cate Schaeffer (University of Wisconsin)</b>	
10.30-11.30	<i>Interaction before and during the survey interview: Insights from conversation analysis</i>	

In their studies of interaction, conversation analysts have described the sequential organization of talk and the formation of actions in a wide variety of contexts. The standardized survey interview seems very different from other talk because the interview itself is highly structured and much of the talk is guided by the interviewer. Nevertheless, conversational practices play an important role in the interaction between the interviewer and sample member both during recruitment and during the interview. This paper gives a selective review of some contributions of conversation analysis to our understanding of how the practices of conversation intersect with the practices of standardization before and during this interview. For example, Schegloff's classic analysis of the identification/recognition sequence that opens phone calls has implications for how interviewers structure their first turn. Similarly, the placement of a question - or repair - from the sample member, may predict whether or not a contact leads to an interview. Within the interview, interviewers use skills from the interactional substrate to diagnose comprehension problems and repair answers, while respondents routinely add conversational elements to their answers.

Thu 19/1	Wil Dijkstra (Professor Emeritus VU University Amsterdam)	
11.30-12.30	<i>SEQUENCE VIEWER REVIEWED</i>	

Sequence Viewer is a program for the analysis of social interactions. The program is especially suited for the analysis of interviewer-respondent interactions. In this case we usually speak of “Question-Answer sequences” (or simply *sequence*): all verbal acts of interviewer and respondent, starting with the interviewer posing a question from a questionnaire, and ending at the moment the interviewer poses the next question. In Sequence Viewer such verbal acts are called *events*.

Sequence Viewer is primarily directed towards the quantitative analysis of such sequences of events, although the program also offers quite a few tools for more qualitative analyses. Quantitative analysis requires that events are coded. Sequence Viewer offers the possibility of what is called a multidimensional coding scheme. That is, each event (or verbal act in interviewer-respondent interaction) is coded on a number of different aspects, or *code variables*. For example, one code variable might be used to designate the speaker; a second code variable to describe the type of exchange (e.g. a question, an answer, a request for repetition, and so on); a third code variable to indicate whether the verbal act is directly relevant to the question from the questionnaire (e.g. posing that question by the interviewer, or selecting a response alternative by the interviewer), or is less relevant (e.g. when the respondent deliberately motivates the answer), or completely irrelevant; a fourth code variable might be the adequacy (is the question posed as scripted; equals the answer one of the response alternatives); and so on, up to nine code variables.

Such a multidimensional coding scheme has a number of advantages. First, it is easy to account for infrequently occurring verbal acts, without inventing new codes. For example, if the respondent asks the interviewer about her opinion, this can be coded as RQ (R for Respondent, Q for Question). Second, analyses can be performed on a selection of code variables. For example, suppose one wants to study turn taking in a group discussion, one could look at the sequence of the code variable designating the speaker only. Third, the number of different codes for each separate code variable can be relatively small, thus making the coding task easier.

In addition to code variables, each verbal act or event, can also have a number of *event variables*: numerical values describing other characteristics of the event; for example the number of words. Two important event variables are the onset and offset times of each event. These two event variables are especially important because next to the transcribed interactions, also the original sound or video files can be linked to the verbal utterances. This makes it possible to calculate a number of other event variables that may be of interest; for example the speech rate of verbal utterances, latency times, pitch, and so on.

A third type of variables are so-called sequence variables, describing characteristics of the sequence as a whole; for example gender and age of interviewer and respondent, question number, the eventual answer, the length of a sequence (the number of verbal acts), the number of requests for elucidation by the respondent, etcetera.

Analyses usually consist of finding characteristics of coded sequences, and assigning such characteristics to event or sequence variables. For example, one can calculate how much a so-called paradigmatic sequence (that is, a straightforward sequence like posing the question as scripted, providing an adequate answer, and acknowledgement of the answer by the interviewer) resemble the actual sequences. Such resemblance can be assigned to a sequence variable, and related to other sequence variables, for example the type or length of the question.

Sequence Viewer offers a number of elementary statistical analyses like analysis of variance and multiple regression to investigate the relations between sequence variables or event variables. In case interesting relations appear to exist, data can be exported to for example SPSS in order to apply more sophisticated analyses.

Next to code, event and sequence variables, Sequence Viewer has the possibility to assign keys to parts of transcripts, or to time periods. The latter option is especially useful if video recordings of interactions are available; for example to assign particular keys to non-verbal behavior, like gestures, frowns, or looking to the other person, in order to investigate whether such behaviors coincide with particular verbal utterances..

In my presentation I will demonstrate how Sequence Viewer can be used to obtain more insight in what is going on in a survey-interview. For more information, see:

<http://www.sequenceviewer.nl/>

Thu 19/1 12.30-14.00	Wil Dijkstra & Yfke Ongena Q&A Sequence Viewer, including lunch	
-------------------------	--	--

<p>Thu 19/1 14.00-14.30</p>	<p align="center"><b>Jessica Broome (JessicaBroomeResearch)</b> <i>How Telephone Interviewers' Vocal Characteristics and Responsiveness Impact their Success.</i></p>
---------------------------------	---

Telephone survey interviewers vary widely in their success at persuading potential respondents to participate. This persuasive act can be viewed in two stages: first, the initial impression the interviewer makes on potential respondents; and, assuming there is success in this stage, the interviewer's ability to respond appropriately to concerns expressed by potential respondents.

My study looks at both of these stages in a corpus of audio-recorded telephone survey introductions for which the call outcome (agreement or refusal to participate in a survey) is known. Initial impressions of interviewers were assessed by asking web survey respondents to listen to the initial seconds of an interviewer's recorded introduction (typically "Hello, this is \_\_\_\_ and I'm calling from [INSTITUTION] about our study on \_\_\_\_.") and to rate the interviewer on several personal and vocal characteristics, including "professional," "competent," "confident," "enthusiastic," "friendly," "pleasant to listen to," "natural-sounding," and "scripted." I found no relationship between any of these positive characteristics and the outcome of the call; the only characteristic found to have a relationship with call outcome was scriptedness, where interviewers rated as less scripted were more likely to have their invitations accepted.

Interviewers' responsiveness to concerns expressed by sample members was assessed through analysis of the interaction in complete survey introductions (as opposed to the initial seconds that were the focus of the first part). These introductions were transcribed and codes assigned indicating the presence of a specific concern (for example, "I don't have time.") and the interviewer's response to the concern (for example, "We can call you back at a more convenient time."). I found that interviewers who offer relevant responses to concerns expressed by sample members were more successful.

Results from both stages of research will be discussed, as well as practical implications for survey operations.

Thu 19/1 14.30-15.00	<b>Jose Benki (University of Michigan)</b> <i>Prosodic properties of the telephone survey invitation and survey participation</i>
-------------------------	--

During the typically rapid participation decision in a telephone survey interview, potential respondents have very little information aside from what they hear over the phone. Given reliable long-term differences in participation rates across interviewers, even when the linguistic content of the invitation is controlled, it stands to reason that paralinguistic speech properties of the interviewer such as intonation, rhythm, and phrasing play an important role in the participation decision. This presentation analyzes the relationship between invitation outcome (agree-to-participate, scheduled-callback, and refusal) and prosodic attributes of interviewers during the invitation, in a corpus of 1380 audio-recorded contacts, sampled from five studies conducted by the University of Michigan Survey Research Center. The analysis focuses on four prosodic attributes of the interviewer invitation: the intonation pattern in the initial greeting, speech rate, pausing, and pitch. Agreement is positively associated with initial greetings with pitch rises, faster speech rates by male interviewers, and more pausing. However, no conclusive pattern is apparent in the relationship between participation and overall interviewer pitch during the invitation.

Thu 19/1	<b>Yfke Ongena (University of Groningen)</b>	
15.30-16.00	<i>Multivariate coding of Event History Calendar Interviews</i>	

Event History Calendar (EHC) interviews have been developed to improve respondent's recall of retrospective information, and are becoming a more and more popular replacement of conventional (standardized) interviews. Contrary to conventional interviews, EHC interviews give interviewers freedom to probe in their own words. The EHC-specific cross-referencing probes (e.g., "did you move when you got married?") are assumed to help respondents recall events by relating them to other events. Conversational interviewing styles in EHC interviews may also provoke narrative styles that may stimulate recall of autobiographical information. However, the extent to which interviewers should be allowed such freedom is subject of an ongoing debate on standardized versus conversational interviewing. This debate was primarily focused on whether interviewers should be allowed to clarify questions, or not.

The quality of interviewer probing can best be studied with behavior coding, i.e., the systematic study of interviewer-respondent interaction in survey interviews (Cannell, Lawson, & Hausser, 1975; Ongena & Dijkstra, 2006). As Cannell already noted, more attention should be paid to "the actual dynamics" of personal interviewer-respondent interaction. Since the 1970's numerous studies analyzing behavior in conventional interviews have been conducted. Belli's (2004) study was the first to explore verbal behaviors in Calendar interviews, and to compare these with verbal behaviors in conventional interviews. However, Belli et al. did not explore the effects of specific interviewer behavior on subsequent respondent behavior. Such sequential analysis of interviewer-respondent interaction analysis allows researchers to test the hypothesized relations between behaviors (Ongena, 2010).

A new multivariate behavior coding scheme was developed, comprising a combination of an existing multivariate coding scheme for standardized interviews (Dijkstra, 1999; Dijkstra, 2002, Ongena 2005) and an existing univariate coding scheme that was designed for Calendar interviews (Belli et al., 2004).

The new scheme, like Dijkstra's scheme, is a full, multivariate coding scheme, using the utterances as a unit of coding (Ongena & Dijkstra, 2006). Sequential analysis of interviewer-respondent interaction revealed the effects of different types of probes on subsequent respondent behaviour. For example, respondents are more than three times as likely to give a don't know answer after 'Do-you-remember' probes than after direct probes.

Thu 19/1	<b>Live stream: Bob Belli (University of Nebraska, Lincoln)</b>	
16.00-16.30	<b>Using Behavior Coding to Understand Respondent Retrieval Strategies that Inform the Structure of Autobiographical Knowledge</b>	

Currently, there are three predominate models regarding the structure of autobiographical knowledge, the basic-systems model, transition theory, and the self-memory system (SMS) model. The basic-systems model focuses on specific episodes as the foundation on which autobiographical knowledge is built, and de-emphasizes the role of more generic memories. Both transition theory and the SMS model highlight the importance of generic memories and their organization. With both, the structure of autobiographical knowledge has been theorized to consist temporally of periods of stability and transitions between these periods. In transition theory, emphasis is placed on transitions as the points that are most memorable and most important in the reconstructive remembering of one's past. In the self-memory system (SMS) model, this temporal organization is complemented by a thematic hierarchical structure. The survey interviewing of respondents to collect retrospective reports on their life course provides an opportunity to examine how participants navigate through their autobiographical memory structures, and to determine the adequacy of models of autobiographical knowledge. In both calendar and conventional interviews, behavior coding reveals that respondents use parallel, timing, duration, and sequential retrieval strategies. The prevalence of parallel retrieval strategies, which represent the associative interrelationship across life course themes as seen by events that occurred contemporaneously or nearly so, challenges the SMS model which fails to account for the occurrence of parallel associations. The content of parallel associations is nearly evenly split between general and transitional events, supporting the importance of transitions in autographical memory in accordance with transition theory, but also raising challenges to transition theory in that general events are also used as retrieval cues at a rate equivalent to the use of transitions. In addition, that these generic memories are used as cues for the remembering of generic events in other themes challenges the basic-systems model which conceives of autobiographical remembering as focusing on specific episodes. Within calendar interviews, associations in respondents' memories (both parallel and sequential) demonstrate complex interactions with interviewer verbal behaviors during generative retrieval, and there are sequences of interviewer-respondent interactions that are most predictive of respondents using these retrieval strategies in their generation of responses. Taken together, results not only raise challenges regarding existing theories of autobiographical memory, but demonstrate how associations are implemented as retrieval strategies in reconstructing one's past, and have implications for interviewer training.

16.30-17.00	<b>Discussion: Peter Miller (US Census) &amp; Tom Koole (University of Groningen)</b>	
-------------	---	--

Friday, 20 January 2017 Marie Lokezaal, [Harmoniebuilding](#)  
University of Groningen, Oude Kijk in 't Jatstraat 26, Groningen

9.30-10.30

Opening & Keynote Fred Conrad (University of Michigan)

It has been observed for several decades that standardized interviewing, the prevailing approach to collecting survey data in the social sciences and government research, may not always lead to uniform interpretation (e.g., Suchman & Jordan, 1990). If some respondents interpret questions differently than intended by the question authors, the accuracy of individual responses and possibly of the resulting population estimates may be compromised. Standardized interviewers read the question as worded and then use only nondirective (i.e., largely content-free) probing to address comprehension problems. To address this potential weakness with standardized methods, an alternative approach to survey interviewing has been proposed that encourages interviewers to clarify the survey concepts, using whatever words they judge will be most effective, when they determine there is misalignment between the respondents' interpretation and how the question was intended (e.g., Schober & Conrad, 1997; Conrad & Schober, 2000). The logic behind this approach is that successful everyday communication often involves back and forth between speaker and listener to assure they are on the same page - at least sufficiently to accomplish their current conversational task. This process of "conversational grounding" (e.g., Clark, 1996) is at the heart of the proposed alternative interviewing technique and has led to its being called "conversational interviewing." In contrast, standardized interviewers cannot ground question meaning because this could involve substantive wording that might differ between respondents.

At least ten studies have been conducted to evaluate the pros and cons of standardized and conversational interviewing. This paper reviews and synthesizes several of these studies describing what has been learned and what is still unknown. Among the clear findings are that conversational interviewing can dramatically improve response accuracy for factual questions when there is ambiguity about the meaning of key concepts in the questions. But the improvement requires additional interviewing time as clarification necessarily involves additional words; this is the case not only in interviewer-administered interviews but also in automated interviews carried out by animated virtual interviewers. Across studies, the improvement in response accuracy is greater when the interviewer can provide clarification both when respondents ask for it and when the interviewer judges the respondent needs it, even without an explicit request. This is also true in web questionnaires with clickable definitions and that can clarify concepts when respondents are slow to answer. Respondents seem to be sensitive to whether interviewers are able to provide clarification in this way, using more disfluent speech and, in face-to-face interviews, averting the interviewers' gaze more often than do respondents in standardized interviews. Conversational interviewers with greater interpersonal sensitivity are more efficient, whether clarifying concepts about factual or opinion questions. And these benefits accrue without increasing interviewer variance. One question currently being investigated is whether conversational interviewing can help improve quality in other ways such as reducing acquiescence and straightlining by better communicating the meaning of response scale values. More also remains to be learned about the practical tradeoffs involved in administering conversational interviews in production surveys, but the evidence suggests they are worth exploring seriously.

Fri 20/1	<b>Jolene Smyth (University of Nebraska, Lincoln)</b>	
11.00-11.30	How do Mismatches Affect Interviewer/Respondent Interactions in the Question/Answer Process?	

When writing survey questions, researchers often focus heavily on the wording of question stems, and overlook the need to ensure that the question stem matches the response options provided. As a result, survey questions are often written with mismatches such as ordinal or nominal response options with yes/no question stems, closed-ended response options with open-ended question stems, or questions where the question stem does not make the response task clear. Previous research on mail and telephone surveys shows that mismatches impact item nonresponse rates, substantive responses, and response time (Smyth & Olson 2016). Although substantive answers are different, it is unclear whether the answers obtained in the mismatch condition are worse or better than those obtained in the matched condition.

In this paper we first examine the interviewer-respondent interaction for questions with matched versus mismatched response options in a telephone survey experiment. Of particular interest is whether rates of interviewers reading questions exactly as worded or with changed wording and rates of standardized and nonstandardized probing, verification, clarification, and feedback behaviors varied across the matched and mismatched versions. Likewise, on the respondent side, we will compare rates of adequate versus inadequate initial answers and requests for clarification and feedback. For both interviewers and respondents, we will also compare rates of interruptions and laughter across the conditions. These analyses will help reveal possible mechanisms for the differences in answers and response time for questions with matched versus mismatched response options.

We also examine whether answers to the survey questions themselves are associated with behaviors during the interview, and whether this effect varies across the experimental conditions. For example, we asked “To what extent do you or others help define the objectives of your job” with the response options of “I mostly define the objectives,” “Others mostly define the objectives,” and “About equal” in the matched version and “Do you help define the objectives of your job?” in the mismatched version. We found that the matched version had a significantly longer response time ( $p < .001$ ) than the mismatched version. Additionally, 19.5% of respondents report that they mostly define the objectives in the matched version, compared to 52.5% in the mismatched version ( $p < .0001$ ). We will examine whether these differences can be explained by interviewer or respondent behaviors.

We use the Work and Leisure Today II survey, which was conducted by telephone in summer 2015 ( $n=911$ , AAPOR RR3=7.84%). Recordings of the interviews were transcribed and behavior coded using Sequence Viewer 6.1 (Dijkstra 2016) at the conversational turn level, allowing a detailed examination of the interviewer-respondent interaction. We discuss implications for questionnaire design and interviewer training.

Fri 20/1	<b>Marieke Haan (Utrecht University)</b>	
11.30-12.00	<i>How can interviewer-respondent interaction analysis help to improve surveys: A study focusing on interviewer deviations in question reading.</i>	

In this paper we show that by looking specifically at types of interviewer's rewording, many more suggestions for improvement of question wording can be revealed than in mainstream behavior coding studies focusing only on differences between major and minor deviations.

In the first study we analyze CATI interviewer-respondent interactions of the 2006 Nebraska Annual Social Indicators Survey. This study shows that interviewers add, substitute, and omit words of items. More specifically, it is found that interviewers sometimes spontaneously add numbers to response options when administering the question. Due to this change in question reading, respondents appeared to be better able to formulate an answer. In the second study, effects of explicitly numbered response options on response behavior are tested. An experiment is conducted using the questionnaire of the European Social Survey round 5 in a CATI experiment. It is found that more mismatch answers are given when numbers are not read by the interviewer. However, no significant differences are found between the experimental conditions. For both conditions it is found that respondents answer the questions more often in terms of words than using a number. A possible explanation for the differences between the two studies is that in study 1 the numbers were spontaneously added to questions with (relatively difficult) nominal response options, whereas in study 2 the numbers were added to questions with (relatively easy) ordinal response options. In addition, the analyses of study 2 show that not all interviewers followed the instructions of reading the question with numbers very well. So also the manipulation was not always successful. In conclusion, even minor interviewer deviations can influence survey data quality. Studying these deviations in the pre-testing phase or the evaluation phase of a survey can assist in improving a standardized questionnaire.

Fri 20/1	<b>Mike Huiskes (University of Groningen)</b>	
12.00-12.30	The sequential unfolding of requests: when a simple yes or no is not enough.	

This paper addresses the design of turns in which a speaker responds to directive sequence-initiating actions such as requests, proposals or offers. Responses to these actions regularly consist out of two different actions in two separate turn slots. In extract 1, for example, R answers to I's question with "ja is goed"

**extract 1. *Ja is goed as an acceptance of a request for cooperation***

(...)

I: en ik zou u hier graag een aantal vragen over willen stellen.  
*and I would like to ask you a number of questions about (this).*

R: → ja is goed  
*yes is good*  
*yes is okay*

I: ok. nou hartstikke fijn dat u mee wilt werken  
*okay. well awfully nice that you want to cooperate*

Following Schegloff, the bipartite composition of the response turn reflects the double duty nature of so called 'double barreled' first pair-parts of adjacency-pair sequences (Schegloff 2007: 73-78). I is not only asking a question making a yes/no answer a relevant next, her utterance is also heard as a request for cooperation to the survey. I attends to both action levels of the FPP turn, with both *ja* and *is goed*.

In this talk we will argue that both elements (both *ja* and *is goed*) are necessary to accept these double barreled request: if either element is missing, I will pursue a full response by redoing the request for cooperation in his/her next utterance. This not only shows that an orientation to conversational norms plays an important role in the daily practice of standardized survey interviews, but also that standardized survey interviews (by virtue of their standardized questions) provide us with natural experiments to study these normative conversational practices in real time.

Fri 20/1	<b>Kristen Olson (University of Nebraska, Lincoln)</b>
13.30-14.00	<i>“During the LAST YEAR, did you...” : The effect of emphasis in CATI survey questions on data quality</i>

Questionnaire design texts commonly recommend emphasizing particularly important words with capitalization, underlining, boldface, or italics to promote their processing (e.g., Sudman and Bradburn 1982, p. 238; Dillman, Smyth and Christian, 2014, p. 174). In self-administered surveys, respondents can directly see the emphasis, but in an interviewer-administered survey, to be effective, emphasis has to be noticed by interviewers and then verbally communicated to respondents through changes in pitch, pauses, speech rate, or other audible signals. Although there have been examinations of the role of verbal emphasis on comprehension and cognitive processing in psychology and linguistics (e.g., Sanford, et al. 2006; Margolin 2013), to our knowledge, there are no survey methodological studies that evaluate the effects of visual emphasis of text in interviewer-administered questionnaires on survey measurement outcomes.

In this paper we report the results of two experiments in which telephone survey questions were presented to interviewers either with emphasis (i.e., full capitalization of important words) or without emphasis (i.e., no capitalization). The experiments were conducted in two dual-frame RDD surveys, the Work and Leisure Today 2 (WLT2) Survey (AAPOR RR3=7.8%) and the Gallup Consumer Marketplace survey (AAPOR RR3=6%). In this paper, we examine (1) whether interviewers

actually engage in verbal emphasis behaviors, (2) whether these emphasis behaviors, if enacted, change the subsequent interaction between the interviewer and respondent, and (3) whether the substantive answers to survey questions are affected by the use of emphasis. A random subset of approximately 900 WLT2 interviews (approximately 450 in each experimental treatment) were transcribed and behavior coded using the Sequence Viewer 6.1 software (Dijkstra, 2016). In addition, the Waveform analysis in Sequence Viewer was used to evaluate whether pitch (fundamental frequency) and intonation (variation in pitch) differed in the question reading for the version with emphasis compared to the version without emphasis.

Initial analyses of the WLT2 indicate that the questions with emphasis took about one second longer to ask and answer than the questions without emphasis ( $p < .0001$ ) and that some substantive answers also differed in the presence of emphasis (e.g., 10.8% with 1+ parking tickets without emphasis vs. 5.1% with emphasis,  $p = 0.02$ ). Additional analyses will also examine behavior codes (e.g., interviewers reading the question exactly as worded; respondents providing adequate answers; presence of paradigmatic sequences) and differences in vocal characteristics across the experimental versions. Implications for questionnaire design and interviewer training will be discussed.

Fri 20/1	<b>Geert Loosveldt (University of Leuven)</b>	
14.00-14.30	<i>Interaction Characteristics and Interviewer Variance</i>	

The evaluation of interviewer effects by means of interviewer variance analysis is a common practice in the assessment of survey data quality. Results of this type of analysis clearly demonstrate that interviewers can explain a significant proportion of variance in substantive variables. The interviewer's impact is not restricted to univariate distributions but can also be observed in the relationships between variables and the results of measurement models. Differences in interview skills and the fact that interviewers differentially deviate from the rules of standardized interviewing offer a general explanation for these interviewer effects. A more specific explanation looks at interviewer characteristics such as experience, age, gender and race. However, the explicit use of respondent-interviewer interaction characteristics as independent variables in models for the analysis of interviewer variance is scarce.

In the paper, we consider interaction characteristics as paradata useful to explain interviewer variance. We will integrate some interaction related paradata in the models that we use to evaluate interviewer variance. Three types of paradata will be used. They differ from each other in the extent to which they are directly related to the interviewer-respondent interaction. The first characteristic is the number of utterances during the interviewer-respondent interaction for a particular question. This characteristic is directly derived from the course of the interaction and is based on interaction coding. The second characteristic is the interview speed (the number of questions per minute) and is based on information collected by timers. The speed can be related to the whole interview or to a particular part of the interview. Speed can be considered as a more general indicator of the interaction process. One can consider both the number of utterances and speed as indicators of the 'amount of interaction' and assume that both characteristics are negatively related to each other (increase of utterances · decrease of speed). One can also assume that longer interactions create more opportunities for interviewers to influence the results of the interaction process. This can contribute to more interviewer variance in substantive variables if the interviewers interact with the respondents in a different way. The interviewer assessments of different aspects of the respondent behavior such as the respondent's level of cooperation during the interview, asking for clarification or his/her understanding of the survey questions are a third type of interaction characteristics. This information focusses more on respondent behavior than on the interaction process and is appropriate to detect more difficult-to-interview respondents. One can assume that interactions with this type of respondents is more complex and sensitive for the impact of interviewers. In general, we hypothesize that in complex interactions differential influences of interviewers occur more frequently. This creates situations that are prone to the occurrence of systematic differences between interviewers which are reflected in the intra class correlation coefficients for substantive variables. Hence, interaction characteristics are used to identify types of interactions that contribute to the occurrence of high levels of interviewer variance. The assessment of the complexity is based on interaction characteristics. We also assume that more directly related interaction characteristics are more strongly related to interviewer variance.

Fri 20/1	<b>Hanyu Sun (Westat)</b>	
15.00-15.30	<i>Rapport between survey interviewers and respondents and its effect on disclosure: the role of interaction</i>	

Interviewer-respondent rapport is generally considered to be beneficial for the quality of the data collected in survey interviews; however, the impact of rapport on data quality has rarely been directly investigated. We conducted a laboratory experiment in which eight professional interviewers interviewed 128 respondents, to see how their rapport with respondents affected the quality of data - primarily disclosure of sensitive information - collected in these interviews. It is possible that increased rapport between interviewers and respondents might motivate respondents to be more conscientious, increasing disclosure; alternatively, increased rapport might inhibit disclosure because to the extent that it leads to a less favorable impression of the respondent, this is particularly aversive. We found that (1) increased rapport increases disclosure for questions that are highly sensitive compared to questions about topics of moderate or low sensitivity; and that (2) increased rapport increases item nonresponse when the questions concern moderately or highly sensitive topics as compared to topics that are not very sensitive. Presumably, increased rapport motivates respondents to be more conscientious, disclosing highly sensitive information, but also raises the cost of disclosing this kind of information, making disclosure more costly than no response. To further explore the interactional origins of disclosure, we coded interviewer and respondent speech and paralinguistic behaviors in one third of the cases. We present the evidences from our corpus connecting interviewer's and respondent's speech to disclosive and non-disclosive answers, when rapport is high vs. when rapport is low. Finally, we discuss a research agenda suggested by our results.

Fri 20/1	<b>Michael Schober (New School for Social Research)</b>	
15.30-16.00	<i>Hesitations in socially desirable responses in a smartphone survey</i>	

When survey respondents answer sensitive questions, they can sometimes present themselves in a more favorable light than is accurate. The current study examines how respondents produce "socially desirable" answers, testing three hypotheses about response latencies as paradata that may produce reliable indicators of socially desirable responding. These hypotheses are lent plausibility by evidence from prior studies of speech paradata in surveys about nonsensitive facts and behaviors, which have demonstrated that, for example, telephone respondents to an automated speech system delay longer when circumstances they are answering about do not correspond with question concepts straightforwardly (Ehlen et al., 2007), and that telephone respondents are more likely to produce disfluent answers in answers that turn out to be unreliable (Schober et al., 2012).

This study compares response latencies of answers to sensitive and non-sensitive questions in a corpus of 319 audiorecorded mobile telephone interviews from Schober et al. (2015), in which participants answered 32 questions (some sensitive and some not) on their iPhones. This allowed testing whether laboratory findings in controlled settings extend to settings where respondents might be mobile, multi-tasking, or distracted. Half the respondents (160) were interviewed by professional interviewers and the other half (159) by an automated spoken dialog interviewing system (speech-IVR). The main comparisons are (a) within-subject comparison of response latencies by the same person to sensitive and non-sensitive questions; (b) between-subject comparisons of response latencies for selecting more and less socially desirable answers to the same sensitive questions; and (c) within-question comparisons of response latencies when sensitive questions are asked by an automated system or human interviewer. Consistent with evidence in other survey modes, respondents produced more socially desirable responses with a human than an automated interviewer. The findings give clear evidence from matched (fair) comparisons across multiple questions that mobile survey respondents hesitate more when answering nonsensitive (vs. sensitive) questions; in answers to sensitive questions that are stigmatized (less socially desirable responses), for at least some survey questions; and when interviewed by an automated system (vs. a human interviewer). More generally, the findings demonstrate that speech paradata are indeed significantly associated with sensitivity of both questions and answers in mobile surveys.

These findings help distinguish the mechanisms underlying socially desirable responding—for example, distinguishing between the time pressure that results from talking at all vs. the pressure that results from having a potentially judgmental human interlocutor. They also raise important further questions about the response mechanisms at play: whether speed reflects less thoughtfulness or conscientiousness when respondents are offended or embarrassed by a topic of questioning, whether respondents want to minimize time spent or effort on embarrassing questions, or whether respondents find particular answers more salient or easily available for sensitive questions.

16.00-16.30	<b>Peter Miller (US Census) &amp; Tom Koole (University of Groningen)</b> Closing day 2
-------------	--