Research Statement for Brady T. West

My research covers several methodological and substantive areas, including survey methodology, applied statistics, and public health. As of April 25, 2021, my publications have been cited 13,278 times (Source: Google Scholar). My citation counts have increased each year during my academic career and 129 of my publications have been cited at least 10 times (Source: Google Scholar). My overall Google Scholar *h*-index is currently 50.

The Total Survey Error (TSE) framework is a guiding paradigm for my work. I co-organized an international conference on TSE in 2015 and co-edited a book titled *Total Survey Error in Practice* (Biemer et al., 2017) that disseminated research from that conference. With this framework motivating my research, I have made significant contributions to the literature in the following five areas:

- Interviewer Effects
- Survey Paradata
- Responsive Survey Design and Related Techniques for Increasing Data Collection Efficiency
- Measurement of (and Adjustment for) Selection Bias in Surveys
- Analysis of Clustered and Longitudinal Survey Data

In addition to these methodological and statistical contributions, I have used my technical and methodological expertise to make several significant substantive contributions with colleagues studying substance use and abuse, mental health, and medical care. I have also recently focused on providing students that I supervise with opportunities to serve as first authors on peer-reviewed publications that disseminate important contributions from their dissertations and other related methodological research. With this research statement, I aim to summarize my scientific contributions in each of these areas to date. All first-authored contributions from my students are denoted with an asterisk (*).

Interviewer Effects. My methodological research in this area has focused on the effects that interviewers can have on the quality of survey data, and how interviewers can introduce errors at all phases of the survey process if they are not carefully trained. To date, this literature has largely focused on measuring and reducing the variance in systematic measurement errors introduced by interviewers in surveys. I have advanced understanding of the effects that interviewers can have on multiple sources of error simultaneously (e.g., nonresponse error and measurement error); see West et al. (2020, *Journal of Survey Statistics and Methodology*), West et al. (2018, *JSSAM*), West et al. (2018, *Journal of the Royal Statistical Society, Series A*), West et al. (2013, *Journal of Official Statistics*), and West and Olson (2010, *Public Opinion Quarterly*). I have also developed statistical methods for comparing independent groups of interviewers in terms of their interviewer effects (West, 2020, *Interviewer Effects from a Total Survey Error Perspective*; West and Elliott, 2014, *Survey Methodology*), and studied the importance of accounting for interviewer effects in large health surveys (Elliott and West, 2015, *American Journal of Epidemiology*).

My initial work in this area led to a four-year NSF-funded project of which I was the PI, looking at the interviewer effects on nonresponse error and measurement error introduced by conversational and standardized interviewing techniques. I managed an original national data collection in Germany for this project that finished in late 2014. My team of research faculty and graduate students analyzed the survey data and linked the survey responses with administrative data provided by the German government. Three papers describing results from this research were accepted for publication by *JRSS-A* (West et al., 2018), *Field Methods* (Mittereder* et al., 2017), and *JSSAM* (West et al., 2018). This line of

research provided a comparison of the interviewer effects introduced by these two interviewing techniques for a variety of survey measures and also decomposed the total interviewer variance introduced by these techniques into sampling error variance, nonresponse error variance, and measurement error variance among interviewers. This work has provided the field with novel insights into the interviewer effects introduced by conversational interviewing and suggested that these effects do not offset the measurement benefits of this interviewing technique.

While this NSF-funded research was ongoing, I also first-authored an ambitious synthesis of the vast literature attempting to explain interviewer effects on a variety of survey outcomes (with Dr. Annelies Blom) using the TSE paradigm. This research synthesis, which provided the field with multiple fruitful directions for future studies of interviewer effects, was published in *JSSAM* in 2017 (West and Blom, 2017) and has already been cited 177 times (Source: Google Scholar). I also co-organized an NSF-funded international conference on interviewer effects in 2019 for both researchers and practitioners and co-edited a book presenting novel research from this conference (Olson et al., 2020, *Interviewer Effects from a Total Survey Error Perspective*).

There are many directions in which I hope to continue extending the literature on interviewer effects. First, the literature almost exclusively focuses on descriptive statistics. Together with a PhD candidate who I am currently supervising (Micha Fischer), I co-authored a paper in *JSSAM* (Fischer* et al., 2019) that outlines a statistical framework for studying interviewer effects on regression coefficients. I hope to build on this framework in future research. Second, I hope to research the mechanisms underlying unexpected interviewer effects on self-administered responses in surveys that involve a mix of face-to-face data collection and self-administration. My initial work in this area (West and Peytcheva, 2014, *Survey Practice*) suggests that interviewer behaviors during self-administration can play a critical role in shifting response. I am currently exploring a research proposal with Emilia Peytcheva in this area. Third, the literature on interviewer effects continues to be devoid of any practical methods for estimating interviewer variance when interviewers are not assigned random subsamples of the full sample. Mike Elliott and I have written a paper describing a new methodology for this estimation problem that recently received a revise and resubmit decision from *Survey Methodology*. Finally, together with Fred Conrad and Michael Schober, we have conducted an NSF-funded study of the interviewer effects introduced in live video interviewing, which will likely become more widely used in the near future. This research note recently received a revise and resubmit decision at *JSSAM*. I anticipate that this initial work will motivate many future studies of the interviewer effects introduced in live video interviewing.

Survey Paradata. I have contributed several methodological studies in this area that evaluate the role of survey paradata in improving the quality of survey data collection more generally (e.g., Mittereder* and West, *In Press*, *JSSAM*; Hu* et al., 2020, *Survey Research Methods*; McClain* et al., 2019, *Social Science Computer Review*; Hu* et al., 2017, *Journal of Nutrition*; Krueger and West, 2014, *POQ*; West and Peytcheva, 2014; West and Sinibaldi, 2013, *Improving Surveys with Paradata* [edited volume]; Wagner et al., 2012, *JOS*; West, 2011, *Survey Practice*). These contributions have focused on the use of paradata to make survey designs and survey estimation more efficient and also provided the field with practical techniques for monitoring interviewer performance (Guyer et al., *R&R*, *Survey Methods: Insights from the Field*; Ong* et al., 2018, *Public Health Nutrition*; West and Groves, 2013, *POQ*). My general focus on the quality of auxiliary information used for survey production has also resulted in a novel study of the ability of commercial data to improve survey operations and estimation (West et al., 2015, *JSSAM*).

In addition to these more general contributions on survey paradata, I have published several studies evaluating the utility of interviewer observations for survey estimation. Interviewer observations have

the potential to provide useful auxiliary information for improving survey estimates in the presence of survey errors like nonresponse, but their quality needs to be carefully monitored if they will be used for this purpose. The research that I performed for my dissertation resulted in three original contributions in this area, addressing gaps in the literature with regard to the quality of "pre-survey" interviewer observations, the implications of reduced quality in these observations (and auxiliary variables more generally) for survey estimation, and approaches for nonresponse adjustment of survey estimates using error-prone auxiliary information. These three studies were published in high-impact peer-reviewed journals (West, 2013, *JRSS-A*; West and Kreuter, 2013, *POQ*; West and Little, 2013, *Journal of the Royal Statistical Society, Series C*). With support from a recent NIH R03 grant, I have since published eight additional studies extending this line of work (West and Trappmann, 2019, *SMIF*; West and Li, 2019, *Sociological Methods and Research*; West and Kreuter, 2018, *Methodology*; West and Kreuter, 2015, *Field Methods*; West et al., 2014, *JSSAM*; West, 2013, *Improving Surveys with Paradata*; and two chapters in the 2018 *Palgrave Handbook of Survey Research*), providing the field with practical guidance on effective strategies for recording accurate observations and working with error-prone observations. In the future, I plan to experimentally evaluate the ability of specialized interviewer training that emphasizes effective observational strategies to improve the quality and utility of these observations.

Finally, I am also interested in the role that "post-survey" interviewer observations can play in improving survey estimates and providing an indication of the quality of the survey response process. These types of observations are routinely collected in different surveys at no small cost. Unfortunately, the literature has not benefitted from any systematic investigations of their potential to indicate response accuracy and the observations are almost never analyzed (regardless of the survey). I recently led an initial study of the potential of latent class analyses of these observations to classify respondents in two large national surveys based on their response quality, and this chapter was published in the aforementioned edited volume on interviewer effects (West et al., 2020). I hope to build on this work in the future and explore the possibility of using post-survey observations to construct data quality indicators for inclusion in public-use survey data sets, enabling sensitivity analyses.

Responsive Survey Design (RSD) and Increasing Data Collection Efficiency. My primary contributions in this area have focused on practical techniques for improving the practice of RSD, advancing the statistical science underlying RSD, and novel techniques for improving both the statistical and cost efficiency of survey data collection. Together with James Wagner and Mike Elliott, my former doctoral student Stephanie Coffey (who recently defended her dissertation), and my current doctoral student Xinyu Zhang, we are working to improve RSD techniques with Bayesian analysis, learning from prior data collection outcomes to improve future data collection decisions. We were awarded an R01 grant from NIH to conduct this research and we have already had three manuscripts accepted for publication (West et al., 2020, *JSSAM*; Wagner et al., 2020, *JOS*; Coffey* et al., 2020, *MDA*). A fourth manuscript evaluating the ability of Bayesian approaches to improve predictions of daily response propensity is currently under second review at *JSSAM*. We also have a fifth manuscript that recently received a revise and resubmit decision at *Survey Methodology*, presenting Bayesian approaches for predicting interviewer effects in non-interpenetrated sample designs.

I have also worked on experimental evaluations of the ability of case prioritization to reduce potential nonresponse bias (West et al., *R&R, JSSAM*; Wagner et al., 2012). These studies provide support for the practical strategy of flagging high-priority cases in a sample management system to focus interviewer effort and increase response rates among these high-priority cases. Experimentation that leverages survey paradata to make effective and efficient design decisions in real-time that reduce TSE is essential for improving survey quality. I recently co-edited a book titled *Experimental Methods in Survey Research*

(Lavrakas et al., 2019) that presents state-of-the-art approaches for embedding randomized experiments in surveys and using the results of the experiments to reduce TSE. My most recent publication in *Field Methods* provides a striking example of the importance of experimentation in surveys for improving population inferences about health disparities (West and McCabe, 2021).

I strive to advance the science of RSD and provide guidance on the implementation of RSD to other fields outside of survey methodology with the ultimate goal of improving the efficiency of data collections both large and small in applied fields. With these objectives in mind, I collaborated with James Wagner to develop an NIH R25 proposal outlining a research education program on RSD. This proposal was funded by NIH and in 2017 we offered our first three courses on RSD as part of the ISR Summer Institute in Survey Research Techniques. This training program has now blossomed into a year-round series of webinars and seminars on RSD and has also given rise to a new three-MOOC Coursera specialization on *Total Data Quality* that I am currently developing with Dr. Wagner. We have developed an interactive website for the RSD program (rsdprogram.si.isr.umich.edu) for researchers using RSD.

Finally, I have recently initiated a line of research into modular survey design (West et al., 2015, *Survey Research Methods*), where a survey is split into multiple modules that can be answered at different points in time and respondents do not need to answer a lengthy survey in one sitting. This approach takes full advantage of new technologies used for web survey data collection. There are many promising directions for this research, which has the potential to reduce perceived survey burden for respondents. I recently leveraged this initial work on modular design to obtain my first NIH R01 grant as a PI. This new project (known as the American Family Health Study, or AFHS; afhs.isr.umich.edu) aims to demonstrate that researchers can use novel sequential mixed-mode designs employing web and mail modes of data collection to collect high-quality family health and fertility data from a national probability sample in the U.S. at a fraction of the cost associated with in-person data collection. As part of this new project, we are experimentally evaluating the effectiveness of modular design when using web and mail methods to improve cost efficiency. In addition, as part of a recently-funded NSF contract, I'm working with a team of faculty and graduate students from MPSM and JPSM to explore new modular design approaches using smartphone apps for the Survey of Doctorate Recipients (SDR), via participatory design workshops with potential SDR respondents. Finally, I am also currently working on an NSF proposal with Ipek Bilgen at NORC to design an experiment that would evaluate the effectiveness of a modular web-based approach to collecting data on a variety of topics from a national probability-based panel.

Measurement of (and Adjustment for) Selection Bias in Surveys. My extensive line of research in this area demonstrates my consistent commitment to the improvement of survey estimates via effective adjustment for selection bias, whatever the source of this bias (nonresponse, non-probability sampling, etc.; Antoun* et al., 2019, *JSSAM*; Taylor and West, 2019, *Educational Researcher*; Hu* et al., 2019, *JSSAM*; West and McCabe, 2017, *AJE*; Hu* et al., 2017; Raykov and West, 2016, *Structural Equation Modeling*; McCabe and West, 2016, *Social Psychiatry and Psychiatric Epidemiology*; Krueger and West, 2014; West and Little, 2013; Groves et al., 2012, *POQ*). I recently completed work as the PI of an NIH R21 grant, where the aim of the project was to develop and evaluate statistical measures of selection bias in surveys subject to non-ignorable selection. This project produced four peer-reviewed manuscripts that have been accepted for publication in high-impact statistical journals (West et al., *In Press*, *Annals of Applied Statistics*; Boonstra et al., *In Press*, *JOS*; Little et al., 2020, *JSSAM*; Andridge et al., 2019, *JRSS-C*) and describe new statistical measures that researchers can use to quantify the potential bias in a variety of survey estimates (means, proportions, regression coefficients) that arises due to non-ignorable selection. These novel measures offer advantages over prior measures in that they can detect the bias that arises from participating individuals being unique in terms of the survey measures of interest. The

measures produced by this project will enable researchers to quantify this potential selection bias, examine the sensitivity of any population inferences made to assumptions about the bias, and adjust the estimates for this potential bias. I am currently exploring the possibility of applying these new measures to non-ignorable selection bias problems in political polling.

Analysis of Clustered and Longitudinal Survey Data. I have first-authored a book that is now in its second edition (with a third edition currently in progress and scheduled to be published in 2022) on the use of alternative statistical software procedures for fitting linear mixed models (West, Welch, and Galecki, 2014). This book has been cited 2,366 times to date (Source: Google Scholar). Working on this book has inspired me to continually understand and inform the literature with regard to state-of-the-art developments in statistical software for multilevel modeling (West et al. 2015, *American Journal of Public Health*; Galecki and West, 2013, *The SAGE Handbook of Multilevel Modeling*; West and Galecki, 2011, *The American Statistician*; West, 2009, *Evaluation and the Health Professions*).

I have also co-authored a second book that presents applied approaches to design- and model-based analyses of complex sample survey data, which is now in its second edition (Heeringa et al., 2017). This book, together with its accompanying website, provides a comprehensive review of the relevant literature and describes best practices with regard to the analysis of survey data with a more applied audience in mind. This book has been cited 1,115 times to date (Source: Google Scholar). I have written numerous collaborative articles with other colleagues in public health and medicine that apply these analytic approaches to large complex sample survey data sets to address substantive research questions. These collaborations have led to numerous methodological contributions in this area (e.g., Smith et al., 2021, *Health Services Research*; Kolenikov et al., 2020, *The Survey Statistician*; Westgate and West, 2020, *JSSAM*; West, 2019, *SAGE Research Methods Foundations*; West et al., 2018, *JOS*; West et al., 2015, *AJPH*; Heeringa et al., 2015, *Annals of Epidemiology*; Raykov, West, and Traynor, 2015, *SEM*; Sakshaug and West, 2014, *AJPH*; West and McCabe, 2012, *Stata Journal*; West et al., 2008, *Stata Journal*). My research agenda in this area and my steadfast dedication to ensuring that applied researchers are analyzing survey data correctly has also produced an influential line of work on the importance of analytic error as a part of the larger TSE framework. I was the PI of an NSF-funded research program in this area that has provided the field with in-depth assessments of how frequently apparent analytic errors occur in scientific publications and the implications of such errors for the quality of survey estimates (West and Sakshaug, 2018, *SMIF*; West et al., 2017, *Total Survey Error in Practice*; West et al., 2016, *PLoS ONE*; Sakshaug and West, 2014, *AJPH*).

Substance Abuse, Mental Health, and Medical Care Research. I have been a co-Investigator on several large NIH-funded research projects with prominent scholars in the epidemiology of substance abuse (e.g., McCabe), mental health (e.g., Gonzalez), and medical care (e.g., Kullgren). These substantive investigations have taken full advantage of my aforementioned methodological contributions to multilevel modeling and the analysis of complex sample survey data. My contributions in these areas have addressed many important knowledge gaps in public health. Most recently, Dr. McCabe and I have had an R01 proposal funded that will examine nonresponse bias in longitudinal surveys, and we are currently writing papers that build on our prior work in this area (e.g., West and McCabe, 2017, *AJE*). I also have a pilot project on measurement of motives for opioid use using mobile applications included in an NIH P50 proposal (led by Danny Almirall and Inbal Nahum-Shani) that recently received a very favorable score. I am eager to continue contributing to knowledge in each of these areas.