Fieldwork Decisions and Their Implications for Data Quality

Kumulative Arbeit

Inaugural-Dissertation

zur Erlangung der Doktorwürde

der

Philosophischen Fakultät

der

Rheinischen Friedrich-Wilhelms-Universität

zu Bonn

vorgelegt von

Manuela Schmidt, geb. Bustamante Hoyos

aus

Cali, Kolumbien

Bonn 2026

Gedruckt mit der Genehmigung der Philosophischen Fakultät

der Rheinischen Friedrich-Wilhelms-Universität Bonn

Zusammensetzung der Prüfungskommission:

Prof. Dr. David Kaldewey

(Vorsitzender)

Prof. Dr. Jörg Blasius

(Betreuer und Gutachter)

Prof. Dr. Paul Marx

(Gutachter)

Jun.-Prof. Dr. Hanno Kruse

(weiteres prüfungsberechtigtes Mitglied)

Tag der mündlichen Prüfung: 02.06.2026

Summary

This cumulative dissertation examines how the implementation and adaptation of survey design during fieldwork affect overall data quality. It is based on the idea that data quality is not just determined by the design of the questionnaire or the analytical techniques used, but is also influenced by how the survey is carried out in the field. Fieldwork decisions influence who participates in a survey, the conditions under which information is provided, and whether respondents remain part of a study over time. These processes operate simultaneously and interact with one another, particularly in longitudinal survey contexts. Drawing on the total survey error framework, data quality is conceptualized as a multidimensional outcome resulting from a series of design and implementation decisions involving necessary trade-offs between competing objectives.

The empirical analyses draw on three papers based on the Cologne Dwelling Panel, each focusing on a distinct yet interrelated aspect of fieldwork decision-making. The first paper examines adaptations to interview conditions introduced in response to the constraints of the COVID-19 pandemic, focusing on whether modifying the implementation of face-to-face interviews affects measurement quality while maintaining longitudinal comparability. The second paper investigates recruitment strategies aimed at addressing nonresponse among out-movers, using an experimental design to assess the influence of incentives and mode combinations on participation, sample composition, and survey costs. The third paper focuses on the monitoring of panel attrition in a longitudinal design with non-standard primary sampling units. It examines how attrition unfolds as a cumulative and design-dependent process and how it can be monitored when standard attrition frameworks are inadequate.

Overall, the papers demonstrate that fieldwork and design decisions influence various aspects of data quality, and that the impact of these decisions cannot be assessed in isolation. Adaptations made to address one challenge often introduce new trade-offs in other dimensions of survey quality. The overall discussion therefore argues for a more reflexive and transparent approach to the implementation of fieldwork, in which design assumptions are continuously questioned, and adaptations are evaluated not only in terms of immediate gains, but also with regard to their long-term implications for data quality. By highlighting fieldwork as a central site of data production, this dissertation contributes to a more nuanced understanding of how survey data quality is shaped in practice.

Contents

1. Introduction

Since empirical social research relies primarily on survey data, research into survey methods is essential for the social sciences. In order to collect high-quality data, it is necessary to implement valid, reliable, and cost-effective instruments that enable a quantified description of social phenomena and statistical inference about a population (Groves et al. 2009). Consequently, data collection practices need to constantly adapt to evolving cultural, technological, political, and social conditions (Groves 2011). As a result, survey methodology is concerned with ever-changing and multidimensional challenges to data quality.

Among these challenges, the continuous decline in survey response has become one of the central issues discussed in survey methodological literature. Unit nonresponse continues to rise over time and across different survey contexts (Brick and Williams 2013; de Leeuw, Hox and Luiten 2018), which can lead to substantial misrepresentations of individual opinions and societal issues, especially when nonresponse correlates with key variables (Cavari and Freedman 2023; Stein et al. 2025). The decline in response rates has two main consequences for survey practice. First, it increases the risk of nonresponse bias. Second, it forces researchers to implement more complex designs aiming to improve response rates, which in turn can increase survey costs, because these strategies often require additional fieldwork effort and greater financial resources (Peytchev 2013).

Strategies and techniques to reduce nonresponse target different potential causes of nonresponse and are implemented at various stages of fieldwork, such as deciding how, and when to contact respondents, contact frequency, interviewer training, and the use of incentives (Groves and Couper 1998). Importantly, both the implications of nonresponse for data quality and the potential for adaptation in fieldwork differ depending on survey design. In longitudinal surveys, the problem of unit nonresponse is intensified, as it affects both initial nonresponse at the recruitment stage as well as dropout in later waves (de Leeuw and Lugtig 2014). At the same time, however, longitudinal consistency constrains adaptation, since fieldwork decisions have cumulative consequences over time. Furthermore, panel studies strive to maintain consistent fieldwork procedures across waves to avoid comparability issues.

A specific challenge for survey research emerged during the COVID-19 pandemic, when external regulations and contact restrictions substantially constrained fieldwork practices, often making established methodological standards of survey implementation no longer feasible. Aggravated by the need for fast data production and rapid publication of results, survey practitioners were often required to simplify fieldwork strategies in the early phases of the pandemic in exchange for the potential lowering of data quality. In this context, again, the challenges were intensified for longitudinal surveys, which were mostly forced to switch from face-to-face to self-administered modes mainly due to contact restrictions, thereby risking the comparability of measurements across waves (Gummer et al. 2025a).

Despite the extensive methodological literature on survey design and data quality, fieldwork decisions are often treated as practical necessities rather than examined as substantive components of survey methodology. This dissertation argues that decisions made during the implementation of survey fieldwork have central implications for multiple dimensions of data quality. Drawing on the established total survey error framework (Groves and Couper 1998; Lyberg 2012; Weisberg 2005), this dissertation conceptualizes fieldwork as the primary site of data production, where design strategies are implemented and shape different dimensions of data quality. The three studies that make up the dissertation focus on different aspects of the fieldwork process in a longitudinal urban panel study with area-based sampling conducted in two neighborhoods in Cologne, Germany. This shared empirical context allows for the examination of survey data quality across different dimensions, while simultaneously shaping the range of feasible fieldwork decisions and design constraints faced by survey practitioners. Further, the shared empirical and conceptual context allows for a deeper understanding of how fieldwork decisions interact with substantive mechanisms and ultimately influence data quality. Taken together, the studies provide a multifaceted perspective on the role of fieldwork decisions within the data collection process.

In the first study, Schmidt and Estermann (2026) examine whether the use of face masks in face-to-face interviews, a fieldwork adaptation introduced under the design constraints of a late phase of the COVID-19 pandemic, is associated with lower data quality. By estimating the effects of face mask usage on multiple indicators while controlling for interviewer-related variation, the study evaluates the extent to which such adaptations affect data quality and longitudinal comparability and provides insights into the face-to-face survey process.

In the second study, Schmidt, Barth and Blasius (2025) focus on fieldwork strategies for addressing nonresponse among former residents of the urban panel's sample area. Using an experimental design varying survey mode and incentive strategies, the study examines differences in participation outcomes and sample composition across fieldwork conditions. The results show that the use of incentives positively affect participation regardless of survey mode and increase the share of participants from harder-to-reach populations. Moreover, by comparing the costs associated with each experimental condition, the study allows for a practical assessment of the cost efficiency of different fieldwork strategies and highlights trade-offs between survey expenses and data quality.

Finally, in the third study, Schmidt and Barth (2026) address the monitoring of panel attrition, again focusing on the urban panel study characterized by area-based sampling and non-standard primary sampling units (in this case, dwellings). The analysis considers predictors of attrition at both the individual and dwelling level, examining how both residents' characteristics and housing-related features are associated with nonresponse. The study contributes to a better understanding of how attrition processes unfold in panel designs that deviate from conventional individual- or household-based sampling approaches.

The shared discussion brings together the findings of the three studies and reflects on their implications for understanding the role of fieldwork decisions in survey data quality.

The three publications on which this dissertation is based are listed below:

- Schmidt, Manuela, and Karoline R. Estermann. 2026. "The Impact of Face Mask Usage on Data Quality in Face-to-face Interviews." *Field Methods* 38(1):19–32. doi:10.1177/1525822X251348080.
- Schmidt, Manuela, Alice Barth, and Jörg Blasius. 2025. "Effects of mode and incentives on response rate, sample composition, and costs – experience from a self-administered mixed-mode survey of movers." *Survey Methods: Insights from the Field*. doi:10.13094/SMIF-2025-00002.
- Schmidt, Manuela, and Alice Barth. 2026. "Unit nonresponse in a dwelling panel study – the role of residents' and housing characteristics." *methods, data, analyses*. doi: 10.25521/mda.595.

2. Data Quality in Social Research

At its core, sociology is an empirical science concerned with observable social phenomena (Durkheim 1982[1895]) and the interpretive understanding of social action and its subjective meaning (Weber 1922). Sociology therefore relies on a close interdependence between theoretical frameworks that structure the description of social reality and empirical research that renders social action observable for researchers. This interdependence implies that sociological theories, as all scientific theories, must be able to withstand confrontation with empirical reality (Popper 2002[1934]).

Empirical understanding, however, is not a direct process, but a *mental exercise* of reflexivity and construction, in which specific research situations shape what can be said, expressed, and ultimately recorded as data (Bourdieu 2017[1993]). As social phenomena are not immediately accessible, but the product of human activity and cultural context, their observation requires an analytical detachment from our own mental representations and a methodological break with preconceptions in order to treat them as external *social things*, as objects of empirical inquiry, and therefore as data (Durkheim 1982[1895]).

Within the broader methodological landscape, survey research is a prominent approach to the empirical observation of social phenomena that aims to produce quantitative descriptions of a study population that allow for standardized description and systematic comparison across social groups, populations, and over time (Fowler 2014[1984]). It is precisely this process of quantification that makes surveys particularly subject to the constraints, interpretations and subjective construction of social reality. Through the specification of constructs and the operationalization of measurements, scientists structure how attitudes and social phenomena are recorded. As a result, data are produced through the lens of a particular conceptual framework (Groves et al. 2009). And arguably more importantly, standardized survey designs and fieldwork decisions can systematically shape patterns of participation, therefore influencing the empirical basis of any survey data (Groves and Couper 1998).

The implementation of high-quality surveys plays a crucial role in the production of reliable data that empirical research relies on. For this, the field of survey methodology aims to identify and improve the design, collection, processing, and analysis of survey data by applying multidisciplinary principles and developing frameworks to improve data quality (Groves et al. 2009).

2.1 Data Quality as a Multidimensional Concept

If we adopt this epistemological approach to survey implementation as a method of empirical observation, the central issue becomes clear: instrument development and survey implementation are not trivial technical steps but integral to how social reality is observed.

The field of survey methodology is therefore concerned with improving the methods through which the assumed social reality is measured. Errors and inaccurate observations may translate into biased estimates and flawed empirical conclusions, with direct implications for sociological inference.

2.1.1 The Survey Cycle

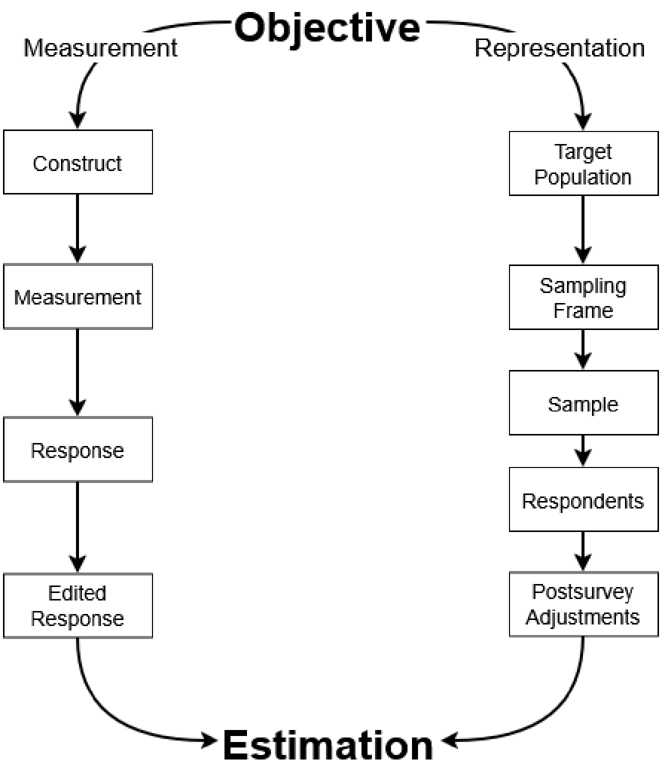
Within the general consensus of survey methodology, surveys are a systematic method for gathering information primarily by asking people questions. The information is usually gathered from a subset of a target population, and answers can be recorded by interviewers or self-administered by respondents. Ultimately, the goal of a survey is to allow for quantitative summaries and analytic statistics about the target population (Groves et al. 2009).

To achieve this goal, surveys are conceptualized as a sequence of interrelated steps, each associated with specific methodological prerequisites. These include the definition of a target population and a sampling frame, the operationalization and description of measurable properties, the selection of units according to sample design, data collection in accordance with measurement processes, and the computation of estimates based on collected measures (Biemer and Lyberg 2003:1–6). This survey cycle can be seen both from a design and a process perspective, because survey quality depends not only on how a survey is planned (design), but also on how these plans are implemented (process). While the design perspective emphasizes the conceptual and structural decisions underlying a survey, the process perspective highlights the iterative stages through which these decisions are implemented.

The core idea of survey design is the hierarchical sequence of steps that transforms an abstract, theory-based research question into survey results.

The survey circle (Figure 1) encompasses two parallel and interdependent dimensions of survey design: measurement, which focuses on the definition of conceptual constructs and their operationalization and collection in form of survey questions, and representation, which refers to the process of specifying a target population and selecting and including relevant observational units (Groves et al. 2009).

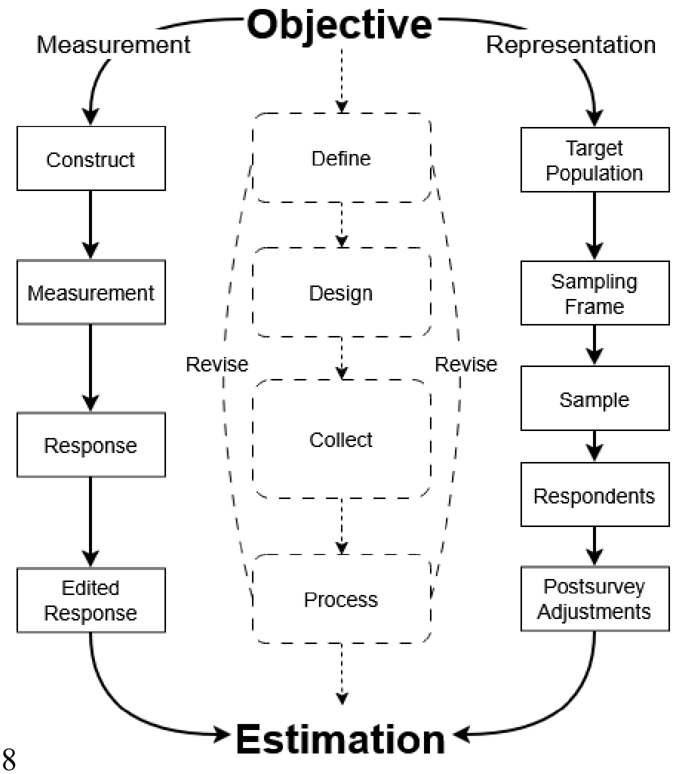
Figure 1. The Survey Cycle from a Design Perspective



(adapted from Groves et al., 2009, p. 42)

The process perspective is composed of the sequential and iterative stages necessary for planning and conducting a survey. The stages include defining research objectives, developing a measurement instrument, choosing a mode of administration, sampling the population, planning and implementing data collection and processing, and statistical analysis. At each stage, researchers might be confronted with new information about the feasibility of the design. As a consequence, previous stages might need adjustment, and the overall survey strategy may be revised (Biemer and Lyberg 2003). Figure 2 combines the perspectives of design and process by illustrating how the conceptual design components of measurement and representation are linked to specific stages of the operational survey process.

Figure 2. The Survey Cycle from a Design and Process Perspective



8

(adapted from Groves et al., 2009, p. 42 and Biemer & Lyberg, 2003, p. 27)

Conceptualizing the survey cycle as a sequence of interdependent design decisions and process stages highlights the relevance of data quality considerations. This is central to observing, measuring and controlling survey quality (Biemer and Lyberg 2003). Ideally, careful consideration of each step in the survey cycle should lead to the production of data that is *relevant* and *credible* for substantive research, *accurate* and *complete* with regard to the measured construct, *comparable* to the target population, *timely* as in available when needed, and *accessible* and *interpretable* for analyses. The overall implementation should also be *coherent* and *cost-efficient* (Biemer 2010; Lyberg 2012).

Historically, the relevance of survey quality has increased alongside the growing complexity of survey designs. Early survey research from the mid-1930s to the early-1960s focused primarily on developing first tools of data collection and sampling approaches and was mainly concerned with formalizing probability sampling assumptions (Groves 2011). The necessity of such efforts was illustrated by the incorrect prediction of the U.S. presidential election outcome in the infamous 1936 Literary Digest poll. The prediction was attributed to the biased sample selection using telephone directories and car registration lists, thereby grossly oversampling individuals with high income (Robinson 1937).

The period from the 1960s through the 1990s marked a new era of expansion in survey research. Technological developments, such as the telephone, along with the expansion of governmental social programs and publicly funded social research, led to an increase in the implementation of surveys by reducing survey costs and raising institutional demand for survey data (Groves 2011). This expansion occurred earlier in the United States, where approximately 75% of households had a telephone by the 1960s. In Germany, by contrast, only about 14% of households had a telephone at that time, limiting the feasibility of large-scale telephone surveys until the late 1970s and early 1980s (Blasius and Reuband 1995). Over time, the routinization of survey participation and the changing nature of interviewer–respondent interactions contributed to declining response rates, as survey interviews became less personal and less exceptional. During this era, the need for accurate data also brought developments around the psychological processes behind survey response (Groves 2011). Biemer and Lyberg (2003) describe this as the *quality revolution*, in which increasing attention to data quality laid the groundwork for the more comprehensive quality frameworks. This growth in complexity was necessary to cope with advancements in survey research at that time, where new technologies, changing environments, and the availability of alternative data sources had altered the conditions under which surveys were designed, implemented, and evaluated (Couper 2013). More recent contributions extend survey quality debates to a changing data environment addressing the role of big data (Amaya, Biemer and Kinyon 2020), the implications of large language models for online survey research (Westwood 2025), and broader efforts to modernize data collection under evolving technological conditions (Kreuter 2025).

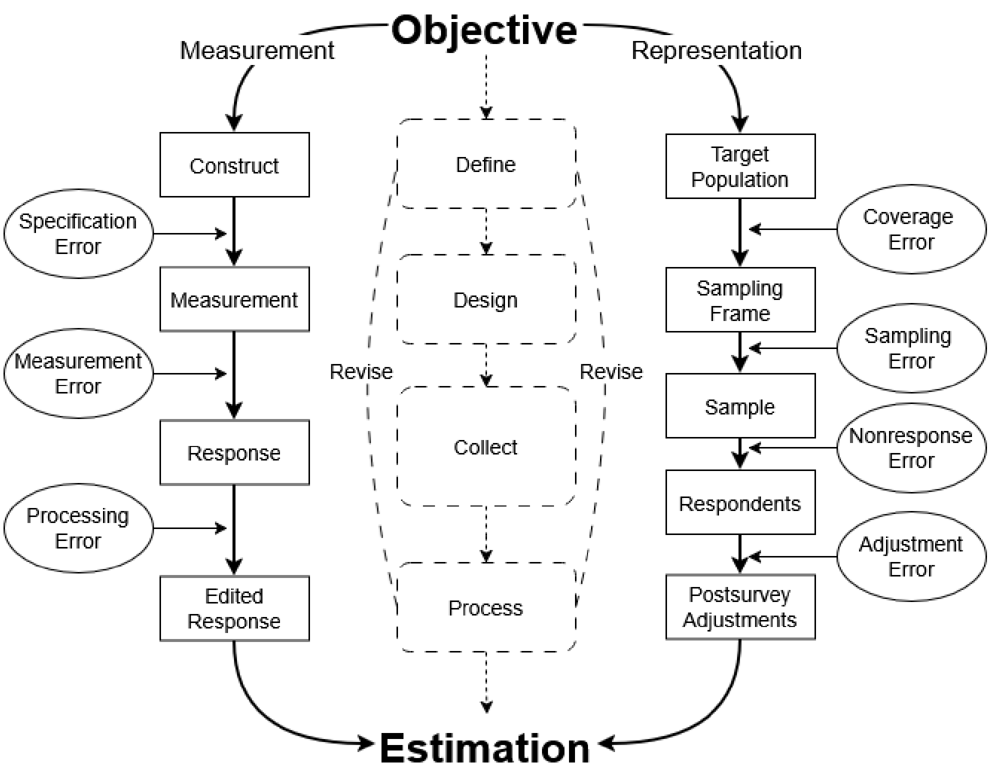
2.1.2 Total Survey Error

This growing attention to survey quality contributed to the development of the total survey error (TSE) paradigm, which in turn can be seen as conceptual foundation of the field of survey methodology (Groves and Lyberg 2010). Rather than treating data quality as a property of the final dataset alone, the TSE paradigm conceptualizes quality in relation to the entire survey process, focusing on sources of error across all stages of survey design and implementation.

Each step of the survey process introduces the possibility of "error", that is a mismatch between the intended and the realized outcome. TSE refers to the combination of all errors that may arise in the design and implementation of a survey. By providing a framework to identify, conceptualize, and assess distinct sources of error, TSE allows survey researchers to evaluate how different aspects of survey design and implementation contribute to overall survey quality (Biemer 2010). Errors cannot be fully avoided as they are often the consequence of necessary design and implementation decisions made under constraints such as cost, time, and ethics (Weisberg 2005). Cost constraints revolve around the financial resources available for the implementation of a survey. Generally, high-quality research is simply more expensive, including organizational and methodological survey costs beyond direct financial expenses such as interviewer training, the implementation of quality control procedures, and additional efforts to investigate nonresponse (Biemer 2010). Time constraints refer to limitations imposed by fixed fieldwork periods, reporting deadlines, and synchronization requirements with external events or institutional schedules. Finally, ethical, legal, and institutional constraints limit how survey research can be conducted (Weisberg 2005).

As errors can arise in each step of the survey cycle and may accumulate or interact across successive steps, survey researchers should aim to minimize the gap between two successive stages (Groves et al. 2009). Figure 3 situates the error sources within the survey cycle.

Figure 3. The Survey Cycle and Survey Errors



(adapted from Groves et al., 2009, pp. 42 and 48, and Biemer & Lyberg, 2003, p. 27)

The following paragraph provides an introductory and non-exhaustive overview of the core error concepts distinguished within the TSE paradigm, drawing on standard concepts in survey methodology and reflecting the distinction between measurement and representation in survey design (Biemer 2010; Biemer and Lyberg 2003; Groves et al. 2009).

Error in the measurement process

Across the measurement process, three types of error may arise. *Specification error* concerns the degree of mismatch between the measure and the underlying theoretical construct. It is a conceptual issue that comprises errors in concepts, objectives, and data elements, and it can be expressed as the correlation of measurement and the true value (or validity). *Measurement error* focuses on the mismatch between the true value and the value provided by a sample unit or the specific survey response. It arises due to errors in the information system, data collection setting, mode of data collection, respondent, interviewer, or instrument. It is generally viewed as one of the most damaging sources of error for surveys and, if systematic, it might lead to response bias, which is the systematic distortion of the response process. *Processing error* is introduced through data entry, coding, and editing, such as removal of outliers or imputation of data.

Error in the representation process

Issues in representation arise during the process of selecting a sample and recruiting sample units from a total population. *Coverage error* arises when the target population does not exactly match the sampling frame and can happen through omissions, erroneous inclusions, and duplications of covered units. Ideally, the entirety of the units in the target population are eligible in the sampling frame, but this is rarely the case. Undercoverage can happen if units are missing from the frame, for example, because they are unreachable for the researcher or via the selected mode (e.g., landline telephone surveys). Overcoverage, in contrast, may occur if units join the sample frame that should be excluded (i.e., the inclusion of business phone numbers in a telephone household survey). The function of the two makes up coverage bias. *Sampling error* consists of two components: sampling bias, which is when members of the frame are systematically excluded from the sample through a reduced or no chance of selection, and sampling variance, which reflects random differences between possible samples drawn from the same frame, limiting the precision and replicability of survey estimates. *Nonresponse error* occurs when sampled units fail to participate in the survey (*unit nonresponse*) or when respondents do not provide answers to specific questions (*item nonresponse*). If nonresponse is systematic, differences between respondents and nonrespondents can lead to nonresponse bias. *Adjustment error* arises due to inaccuracies introduced by statistical procedures intended to compensate for coverage, sampling, or nonresponse errors. Such errors may arise from incorrect weighting, imputation, or tabulation procedures, including errors in the construction or application of survey weights.

As shown, errors and biases are deeply interconnected. Within the TSE framework, survey quality is shaped by trade-offs between bias and variance. Bias refers to systematic deviations of survey estimates from the true population value, whereas variance reflects random fluctuations of estimates around that value due to sampling and measurement processes. Design decisions that reduce systematic error may increase random error, and vice versa. For example, an incomplete sampling frame that systematically excludes certain population groups might lead to bias in measurement because some characteristics are consistently under- or overrepresented. In contrast, a reduction in sample size may leave estimates unbiased on average but increase their variance, as results become more sensitive to random sampling fluctuations. Accordingly, TSE should be expressed as the sum of all variable errors and all biases, labeled *mean square error* (Groves 2004). However, TSE is generally used more as an intellectual framework than a statistical model, and the mean square error is not often computed (Groves and Lyberg 2010).

2.1.3 Data Quality in Longitudinal Surveys

While the TSE paradigm provides a general framework for conceptualizing survey quality across the survey process, it does not consider the unique challenges arising in longitudinal data collection. Here, some aspects of error, interactions between them, and trade-offs to survey design differ from cross-sectional surveys (Lynn and Lugtig 2017). Longitudinal surveys collect data from the same sample units repeatedly over time. As such, they provide analytical advantages over cross-sectional surveys, like the analysis of gross and unit-level change, measures of stability and time-related events, the computation of aggregate measures, and ultimately facilitate analyses of temporal ordering that are relevant for causal interpretation. To lever the advantages of longitudinal survey data and enable comparability over time, survey instruments and processes require stability across waves (Lynn 2009). In other words: “If you want to measure change, don’t change the measure” (AAPOR 2022:5).

The repetition of measurements over time, however, introduces two distinct challenges for survey quality through path dependency. Due to the repetition of questions, the experience from previous surveys may lead participants to answer in a way that systematically differs from respondents answering for the first time, a phenomenon known as *panel conditioning* (Lynn 2009; Warren and Halpern-Manners 2012). Moreover, issues of nonresponse become multidimensional and more complex: unit nonresponse becomes a pattern made up of wave nonresponse and total dropout, known as *panel attrition* (Lugtig 2014; Lynn 2009), which may interact with item nonresponse within and between waves (Lynn and Lugtig 2017). Therefore, the time dimension presents a complex issue to survey design and planning, which generates more initial and ongoing costs. The measurements that make up the first wave instrument cannot be designed independently of the instruments for later waves, as they need to consistently deliver the same estimates. The sample design also needs to be prepared to follow sample members throughout the planned time, for instance in the case of relocations (Lynn 2009).

A longitudinal design might also introduce additional error sources within the TSE error components. *Specification error* can emerge if the conceptual understanding of the research objective evolves over time. Adapting the measurement instruments then involves a trade-off between longitudinal stability and mitigating the growing mismatch between the measurement and the (evolving) underlying construct. *Measurement error* is especially consequential when observing change, as differences in measurement properties across waves, recall effects, mode or context changes, and panel conditioning interact and influence the estimates of change. *Processing error* can result from repeated editing and dependent data collection practices that, while intended to improve consistency, may introduce new errors or inflate between-wave correlations. *Coverage error* may increase over time as population dynamics introduce new discrepancies between the evolving target population and the available sampling frame. *Sampling error* can become systematic if the recruitment of additional samples needed to avoid undercoverage introduces unequal or inconsistent selection probabilities across waves. *Nonresponse error*, as explained above, becomes multidimensional and cumulative over time due to panel attrition, which might lead to more complex structures of nonresponse bias. *Adjustment error* may increase due to dynamic population processes and complex nonresponse, which complicate the estimation of selection probabilities and lead to imputation procedures that may change as information accumulates across waves (Lynn 2009; Lynn and Lugtig 2017).

In conclusion, longitudinal surveys are not the simple repetition of a survey instrument, but a complex system of design decisions and error sources that accumulate over time. Constant adaptation is necessary to accurately reflect the original, but evolving objectives, while continuously risking the loss of comparability, coherence, and interpretability if the trade-offs are not carefully managed. At the same time, constraints become tighter as survey cost increases through complexity and operational flexibility decreases due to path dependencies.

2.3 Strategies to Reduce Survey Error

Understanding the TSE paradigm shows that survey errors can be reduced but not fully avoided as survey quality is shaped by interdependent design decisions and implementations that are the result of trade-offs between errors and constraints. In practice, the TSE paradigm can be applied in both the design stage, where implementation is conceptually planned, and the implementation stage, where design decisions unfold in practice (Biemer 2010).

To reduce survey errors at the design stage, survey practitioners develop a tight plan that considers the goals and risks of each step within the survey cycle. Every decision is a unique preventive strategy that aims to reflect the needs of both the measurement and representation dimension while considering constraints (Groves et al. 2009). The sequential structure of the survey cycle gives a natural order to survey design: the first steps of defining the target population and the data requirements largely determine subsequent choices. For example, while an employee satisfaction survey in a white-collar company can be efficiently implemented as a self-administered web survey and is not urgently time sensitive, a pre-election survey aiming to reach a nationwide electorate requires modes that also cover offline populations, such as telephone or mail surveys, and needs fast implementation in order to deliver timely and reliable estimates. These early design choices, in turn, constrain the available sampling frames and shape instrument development, including the extent to which questionnaires must be adapted to the presence or absence of interviewer support (Blair, Czaja and Blair 2013). A prominent example of a design stage approach to survey error reduction is the Tailored Design Method, which conceptualizes survey quality as the result of coordinated design decisions across sampling, measurement, contact strategies, and incentives (Dillman, Smyth and Christian 2014). Approaches like *adaptive survey design* go a step further, aiming to use available information about a target population and its subgroups to assign more appropriate data collection protocols, reducing nonresponse error (Schouten, Calinescu and Luiten 2013; Schouten, Peytchev and Wagner 2017).

These design choices shape subsequent implementation strategies, which comprise specific actions carrying out the survey plan, such as contact protocols, incentive structures, and follow-up intensity. They influence the types of errors that can realistically be addressed during fieldwork. Strategies at the implementation stage often differ from the initial plan due to disruptive factors such as unforeseen issues, unanticipated costs, or scheduling problems. Dealing with these uncertainties needs a flexible and proactive approach to survey implementation in fieldwork (Biemer 2010). This has motivated the development of implementation stage approaches that explicitly treat fieldwork as a dynamic process, such as *responsive design* and some implementations of *adaptive design* (Tourangeau et al. 2017), which were primarily developed to deal with increasing issues of nonresponse and nonresponse bias. The goal is to identify design features that might affect survey costs and error and monitor them using paradata during data collection. If necessary, interventions are implemented in the field to reduce survey errors and manage costs (Brick and Tourangeau 2017; Groves and Heeringa 2006; Tourangeau et al. 2017). Drawing on these ideas, Luiten and Schouten (2013) implemented an experimental *Tailored Fieldwork Design* in a running panel. For this, they used a sequential mixed mode design of a mail and/or web survey with a telephone follow-up for nonrespondents, tailored to the expected contact and co-operation propensities of the sample units. The experiment demonstrated that adaptive allocation of fieldwork resources can improve the representativeness of survey response under fixed budget constraints, even when overall response rates change only marginally.

Overall, these strategies highlight the reality of survey practice: theoretically sound research and perfectly planned designs are always confronted by trade-offs, constraints, uncertainties and real-life implications. Understanding the actual process of data production is therefore essential for assessing how survey quality is generated in practice.

2.4 Fieldwork as a Site of Data Production

Survey fieldwork then logically emerges as the site in which survey data is actively produced through the interaction of design decisions, organizational routines, and relevant constraints. It is not merely the practical execution of a design, but the place where costs, time, and ethics are negotiated in practice. Even carefully planned survey designs cannot fully anticipate the conditions under which fieldwork takes place, making the implementation of surveys somewhat uncertain. Surveys are susceptible to unforeseen issues and external “shocks”, such as institutional changes, staffing issues, or even the weather during data collection, that make adaptations necessary (Biemer 2010), and these challenges are intensified in longitudinal designs due to path dependencies.

By acknowledging this perspective, the empirical question becomes not how surveys should be implemented, but how concrete fieldwork conditions shape survey data quality in practice. Addressing this question requires empirical analyses that examine fieldwork decisions and conditions as substantive components of the survey process rather than treating them as technical and deterministic aspects of implementation. The following empirical chapters build on this perspective by investigating how fieldwork decisions are associated with survey data quality under real-world constraints.

3. Empirical Setting: The Cologne Dwelling Panel

The empirical analyses focus on different aspects of the survey implementation process of the Cologne Dwelling Panel (*Kölner Wohnungspanel*), a longitudinal study designed to examine housing conditions, residential mobility, and neighborhood dynamics in two neighborhoods of Cologne, Germany. The panel structure, combined with its spatial constraint and repeated fieldwork over several waves, makes the study particularly suitable for analyzing how fieldwork decisions interact with dimensions of survey quality over time.

The survey design and the institutional, organizational, and temporal conditions under which data collection takes place structure the range of feasible design adaptations and shape survey outcomes. This is important for longitudinal surveys, where early design choices and fieldwork implementation create path dependencies that affect later waves. Across its waves, the Cologne Dwelling Panel faced these types of practical challenges, particularly declining response rates, increasing costs of face-to-face fieldwork, adaptations in fieldwork strategies introduced by external shocks, and the design-induced exclusion of moved-out individuals, resulting in systematic missing information.

The Cologne Dwelling Panel was designed to measure neighborhood change in the neighborhoods of Deutz and Mülheim in Cologne, Germany (Friedrichs and Blasius 2015; Friedrichs and Blasius 2020). The selection of the neighborhoods was based on the assumption of a high susceptibility to gentrification processes, mainly through their location close to the city center, the characteristics of their buildings, and the social structure of working and lower middle class residents (Hamnett 1991).

By defining dwellings as the primary sampling units that are followed over the panel waves, the study enables the assessment of socio-structural change in the research areas by tracing individual moves in and out of the neighborhoods. In the interview, the current resident of a dwelling acts as its "speaker", making them the secondary sampling unit, providing information on both dwelling characteristics and individual information, such as household characteristics and attitudes. This unique design offers analytical advantages for observing urban change compared to individual-level panel data, where spatial comparability is lost when respondents move; aggregated data, which lack information on intra-individual change; and repeated cross-sectional surveys, which lack comparability over time (Friedrichs and Blasius 2015).

For the first wave, a random sample of N=2,372 individuals was drawn by the Cologne Office for Urban Development and Statistics, based on previously identified natural areas in the neighborhoods made up of specific streets. Participants were contacted via personalized postal invitations explaining the research goals and announcing a visit by the interviewer in the next days. During the first wave, interviewers noted the position of each respondent's name on the respective housing unit's bell-board in multi-family buildings to allow for tracking over time (Friedrichs and Blasius 2020). In the following panel waves, inhabitants of the same dwellings were interviewed, defined by the bell-board position noted in the first wave. This sampling and follow-up design reflects initial trade-offs: while drawing individuals from register data is a clean approach to aim for statistical inference, it may be problematic as a sampling strategy for a dwelling panel as it increases the selection probability of larger households. The use of bell-board tracking, in contrast, constitutes a pragmatic solution to longitudinal follow-up in multi-family buildings, but it also introduces uncertainty due to possible name changes and alterations of building façades and bell-boards.

From the initial N=1,009 dwellings entering the panel in wave 1 (2010), N=896 participated in wave 2 (2011), N=843 in wave 3 (2013), N=848 in wave 4 (2014) and N=460 wave 5 (2022). Between each of the first four waves, 11–16% of households moved out of the dwelling units, identified primarily by changes in the bell-board and then by comparison of respondents' information in the case of participation. In order to include out-movers and address the systematic lack of information on reasons for leaving the neighborhood, a complementary cross-sectional survey of former residents was conducted alongside the fifth wave of the panel (2023). Based on a random sample of individuals who moved away from the sampled neighborhoods between 2018 and 2022, the survey targeted movers irrespective of prior panel participation and yielded N=928 valid interviews.

The panel data (Blasius and Barth 2025) has been used in several substantial applications around gentrification and urban research. For example, cross-sectional analyses have explored differences in neighborhood perceptions between long-term residents and newcomers (Atakan 2023) or reconstructed residents' symbolic space of neighborhood perceptions to inspect social cohesion and expectations of change (Atakan and Barth 2025). Longitudinal analyses based on the data have mainly focused on the description of gentrification stages and changes in social structure (Blasius, Friedrichs and Rühl 2016a; Blasius, Friedrichs and Rühl 2016b), changing perceptions and expectations of the neighborhood (Blasius and Friedrichs 2016), and the analysis of rent dynamics (Barth and Blasius 2023).

From a methodological perspective, the Cologne Dwelling Panel provides an empirical setting in which fieldwork decisions and their consequences for data quality become visible. The combination of area-based sampling, non-standard primary sampling units, repeated face-to-face fieldwork, and externally imposed adaptations creates a context in which survey implementation is shaped by multiple, interacting constraints. The three empirical studies presented in the following chapters draw on different components of this shared research design and examine fieldwork-related influences on measurement processes, participation outcomes, and longitudinal nonresponse.

4. Data Quality Trade-offs in the Field

According to the conceptual discussion of data quality in the context of TSE, decisions about survey design and implementation involve trade-offs that are specific to the research topic, the target population, and institutional constraints.

This chapter focuses on three domains of fieldwork decision-making: interview conditions, recruitment strategies, and monitoring of panel attrition. It draws on empirical evidence from the Cologne Dwelling Panel, which offers a common background of design features, fieldwork requirements, and general constraints, guiding decision-making of trade-offs.

Zooming in on these domains situationally, enables the examination of how concrete fieldwork decisions shape different dimensions of data quality in practice.

The first subsection examines adaptations to interview conditions in answer to external shocks, implemented as a trade-off in favor of longitudinal continuity. It examines the implications of this adaptation for measurement quality, drawing on the evidence by Schmidt and Estermann (2026) and current methodological findings. The second subsection builds on the long tradition around the use of incentives and their impact on nonresponse and nonresponse bias to analyze how the use of different recruitment strategies influence response patterns and sample composition. Drawing on empirical evidence by Schmidt et al. (2025), the subsection highlights the trade-off between increasing response and managing potential challenges in representation and cost. The third subsection addresses panel attrition as a cumulative process shaped by early design decisions. Drawing primarily on Schmidt and Barth (2026), it discusses the general literature of attrition with regard to individual and household panels, offering an alternative perspective for monitoring attrition when dealing with non-standard sampling units.

The empirical analyses are united by a common focus on fieldwork decisions under constrained design decisions. Together, they illustrate how data quality is not shaped by single factors, but by interacting decisions made before, during, and after data collection, whose consequences are context-specific and cannot directly be transferred into other survey settings.

4.1 Interview Conditions and Measurement Quality

The quality of survey answers is not only a matter of good survey questions. To gain an understanding of the influences on respondents' answers to survey questions, the response process must be considered. Survey answers are the product of a cognitive process, with the steps of comprehension, information retrieval, judgment, and estimation before reporting an answer (Tourangeau, Rips and Rasinski 2000). This process is heavily influenced by contextual factors: the mode and setting of data collection, the wording of questions, the order of questions in the survey, and, in personal interviews, the presence of an interviewer (Groves et al. 2009; Tourangeau et al. 2000).

A particularity of data collected in face-to-face survey interviews is that it is a product of the interaction between interviewer and respondent. Although these interactions are standardized by design, they are ultimately a social situation, vulnerable to individual behavior (Maynard and Schaeffer 2002). Research has shown that interviewers might influence survey data quality on several TSE dimensions, intentionally or unintentionally: coverage error may be introduced through modification of address lists and screening-interviews; nonresponse error may arise through systematic patterns specific to interviewers' conversion success; processing error may occur in the case of inaccurate coding of open-ended answers; and measurement error through intervention or assistance in the response process and through personal characteristics extending from age to physical appearance (West and Blom 2017).

In longitudinal data collection, measurements aimed at quantifying change require that respondents receive the exact same question stimulus each time. Using a different mode or significantly changing the survey instrument has been shown to introduce unintentional changes in answers that are associated with the changing stimulus, but not actual change (Dillman 2009).

Considering this, any external shock that disrupts the mode and setting of data collection or the role of the interviewer can significantly change measurement. In this sense, the COVID-19 pandemic presented a major challenge to survey implementation, specifically to longitudinal surveys that were forced to adapt survey designs and data collection protocols to enable fieldwork during the pandemic due to restrictions and interventions implemented by governments and institutions (Gummer et al. 2025b).

As a response to the contact restrictions during the pandemic, many running longitudinal surveys previously using face-to-face interviews switched their mode of data collection. Gummer et al. (2025b) investigate the adaptations made by n=53 survey programs, demonstrating that most programs using face-to-face interviews switched their mode of data collection (12 out of 19), while other modes were barely impacted. Implementing these changes, especially mid-field, was the result of a careful consideration of trade-offs: In the Survey of Health, Ageing and Retirement in Europe, the face-to-face mode was replaced by telephone interviews, accepting an anticipated moderate mode-related measurement error in order to avoid potentially more severe interactional and comprehension problems associated with self-administration among older respondents (Scherpenzeel et al. 2020) and in the

Refugees in the German Educational System panel, about 40% of interviews were completed before lockdown, leading survey practitioners to accept an increased mode-related measurement uncertainty and adapt the survey instrument in order to preserve temporal comparability, panel participation, and fieldwork feasibility (Will, Becker and Weigand 2020). In contrast, the Understanding Society panel already had an infrastructure for the implementation of mixed-mode surveys. They suspended all face-to-face interviews and very quickly shifted into a full web-first design with telephone follow-ups, allowing for the continuity of the questionnaire. While acknowledging the risk of underrepresentation of low web-propensity households, the switch lead to financial savings due to less travel time and expenses (Burton, Lynn and Benzeval 2020).

Evidence from panels switching from face-to-face to other modes outside of the pandemic context further suggests that a change in mode can lead to differences in response distribution, but that these are largely explained by the parallel change in sample composition (Penker, Eder and Hadler 2025; Schröder et al. 2025).

In the case of the Cologne Dwelling Panel, however, other factors had to be considered while deciding on the implementation of the fifth panel wave during the pandemic. In this case, a change in survey mode was considered particularly costly with regard to measurement and longitudinal comparability. The panel is based on repeated face-to-face interviews conducted inside sampled dwellings, with the dwelling serving as the primary sampling unit and the current resident acting as its speaker. More importantly, the tracking of units using the bell-board method relies on interviewers being on-site. This design implies a strong dependence on physical presence. A switch to self-administered or remote modes would therefore not only have altered the mode of data collection but also disrupted the spatial logic through which the panel follows neighborhood change over time.

Instead of changing the mode of data collection, the fieldwork strategy in the fifth wave of the panel focused on adapting interview conditions within the existing face-to-face design. Conducted in the summer of 2022, after the peak of the pandemic but under continued reduced governmental regulations, interviewers were instructed to offer respondents the option of wearing a face mask during the interview. This adaptation aimed at preserving longitudinal continuity and mode consistency while addressing ethical concerns, respondent preferences, and interviewer safety.

It should be acknowledged that this adaptation was feasible only because the survey wave took place at a relatively late stage of the pandemic. Other panels, confronted with fieldwork disruptions during periods of strict lockdown, did not have such options, as legal restrictions, ethical considerations regarding respondent and interviewer health, and the absence of prior evidence on the effects of face mask use in interviews substantially constrained feasible fieldwork strategies.

The paper by Schmidt and Estermann (2026) provides some of the very little empirical evidence available on this adaptation, suggesting only minor consequences for measurement quality. The preference for interviewer mask usage was not found to be systematically associated with respondents' sociodemographic characteristics, indicating that the adaptation did not introduce selective measurement or participation effects. Moreover, mask usage was not associated with increased item nonresponse or longer interview duration, suggesting that core indicators of response quality and interview efficiency remained stable. While answers to open-ended questions were shorter when interviewers wore face masks, their substantive content, measured by the number of topics mentioned, was unaffected. Taken together, these findings indicate that adapting interview conditions within the face-to-face mode introduced potential measurement error in a limited and manageable way, without undermining the overall quality or longitudinal comparability of the data.

These results underline that the consequences of fieldwork adaptations for measurement quality are highly contingent on survey design and context, and that preserving face-to-face interviewing while modifying interview conditions can be a viable strategy for managing trade-offs under constrained fieldwork options.

4.2 Recruitment Strategies, Incentives, and Participation

Unit nonresponse has remained a relevant issue in survey research since its conception (Hansen and Hurwitz 1946; Smith 1995) and has become even more relevant due to the continued decline of response rates (Brick and Williams 2013; de Leeuw et al. 2018). This decline leads to a possible increase of nonresponse bias while simultaneously requiring more complex and costly survey designs (Peytchev 2013).

Accordingly, understanding why individuals choose to participate in a survey and how fieldwork procedures may impact this decision is important. First, sample units need to be contacted, which is influenced by many factors, such as sociodemographic characteristics, accessibility at home, or the intensity of contact efforts (Groves and Couper 1998; Nicoletti and Peracchi 2005). Once the sample unit is contacted, there are aspects within and outside of the researchers' control that influence survey cooperation, described by Groves and Couper (1998) as a framework for survey cooperation. Outside of control are, again, aspects like the social environment and economic conditions of the sample unit, as well as their sociodemographic characteristics and psychological dispositions (Groves and Couper 1998). Within a researchers control is the survey design (topic, mode, sample) as well as contact strategies like the use of interviewers with more experience, the communication through invitation letters, or a tailored mix of any of those (Brick and Williams 2013; Dillman et al. 2014; Groves and Couper 1998).

Building on this conceptual framework, leverage-saliency theory helps to understand how these factors interact in the decision to participate in a survey. According to this perspective, survey participation is not driven by a single motive, but by the relative salience of different aspects of the survey invitation, such as topic, burden, trust, and perceived benefits. These aspects act as leverage points whose influence depends on respondents' pre-existing social positions, experiences, and psychological dispositions. Survey design features can increase participation by making certain leverage points more salient, but their effectiveness necessarily varies across individuals and contexts (Groves, Singer and Corning 2000).

Survey design may use one or more of these levers. For example, the effectiveness of incentives, especially if they are unconditional, prepaid and monetary, has long been established in survey research (Abdelazeem et al. 2023; Church 1993; Pforr et al. 2015; Singer and Ye 2013), being particularly effective for respondent groups that are otherwise difficult to motivate due to their social position, lack of interest in research or altruistic motives (Singer, Groves and Corning 1999). Research also shows that, as respondents prefer modes of data collection depending on their characteristics (Wolf et al. 2021), giving respondents the option to choose their own mode of data collection can lead to an increase in the response rate (Biemer et al. 2018). Recently, it has been shown that mixing an optimized incentive strategy and an optimized survey mode strategy works better than either strategy alone (Stadtmüller et al. 2023).

While these strategies are effective in increasing overall participation, they are not necessarily neutral with respect to data quality. From a leverage-saliency perspective, increasing the salience of certain incentives or contact modes may disproportionately affect sample composition. Whether this is beneficial or not is open to discussion: on the one hand, boosting response in groups that might otherwise not participate in the survey can help to reduce nonresponse bias. On the other hand, efforts to influence sample composition through targeted design interventions may introduce additional sources of bias and measurement error, thereby altering the error structure of the survey (Peytchev 2013). For example, monetary incentives may be particularly effective among younger respondents, increasing their representation in the sample composition, while simultaneously changing response behavior. At the same time, either of these strategies may also increase survey costs.

Nonresponse bias is a particular issue if attributes associated with unit nonresponse are in the focus of research. For example, in studies of residential mobility and neighborhood change, individuals who move out of the sample area constitute a population that is both harder to reach and of central analytical interest, such as disadvantaged groups that might be more impacted by gentrification processes (Newman and Wyly 2006). In the Cologne Dwelling Panel, this challenge is further shaped by the area-based sampling design. While the panel allows for a detailed observation of neighborhood change among residents who live in the sample dwellings, it does not, by design, follow individuals who leave the sampled dwellings. As a consequence, the panel allows dwelling-level analyses but not investigation of out-moving residents. Addressing this gap therefore requires active recruitment strategies that balance the goals of maximizing participation, maintaining representativeness, and managing survey costs.

To address these challenges, Schmidt et al. (2025) experimented on how the used fieldwork strategies affected participation in the out-mover population. A cross-sectional survey of former residents was implemented in 2023, targeting individuals who had moved out of the sampled natural areas between 2018 and 2022. The fieldwork design varied recruitment strategies by combining web-first and concurrent (paper/web) modes, and the use of prepaid monetary incentives. These design choices were motivated by the need to evaluate participation gains in relation to potential changes in sample composition and their cost implications, rather than by the goal of maximizing response rates alone.

The findings of Schmidt et al. (2025) demonstrate that the use of incentives substantially increased participation largely independent of the chosen survey mode. Incentives were especially effective in increasing participation among population groups that are typically harder to reach, including younger respondents and individuals with lower socioeconomic status. At the same time, the analysis shows that mode variation alone had a comparatively smaller effect on participation once incentives were taken into account.

From a data quality perspective, these results highlight that recruitment strategies shape both response rates and sample composition. By increasing participation among groups that are otherwise underrepresented, the use of incentives has the potential to reduce nonresponse bias in the out-mover sample. However, these gains come at a higher cost per completed interview.

Taken together, the findings echo the overall state of research about the effectiveness of monetary incentives as fieldwork tools. However, their effectiveness when balanced against the significantly higher cost to survey implementation depends largely on the specific research objective, the characteristics of the target population, and the design constraints of the survey. This illustrates that strategies aimed at increasing participation must be evaluated jointly with their implications for sample composition and survey costs, rather than solely in terms of response rates.

4.3 Panel Maintenance and Longitudinal Data Quality

The issue of nonresponse, either at unit or item level, can have a big impact on many dimensions of survey data quality. For example, high rates of nonresponse can influence the survey estimates, if people participating systematically differ from the total sample, leading to nonresponse bias (Groves et al. 2009). Because of this, survey methodologists have continuously developed new, and sometimes more complex, strategies to increase response rates (Peytchev 2013).

When applied to panel surveys, the issue of nonresponse becomes multidimensional. First, nonresponse might apply to either the first contact or recruitment survey, or to the continuous missing participation in later waves. As such, unit nonresponse becomes a complex pattern made up of wave nonresponse and total dropout, known as *panel attrition*, in which different mechanisms take place at the different stages (Lugtig 2014; Lynn 2009). Further, these mechanisms can interact with item nonresponse within and between waves (Lynn and Lugtig 2017).

From a TSE perspective, panel attrition constitutes a particularly complex source of error because it is cumulative and path dependent. Nonresponse in a single wave may be temporary, but repeated instances of wave nonresponse or permanent dropout accumulate over time and can interact with earlier design and fieldwork decisions (Lynn and Lugtig 2017). As a result, the quality of longitudinal data cannot be assessed solely on the basis of response rates in individual waves, but must take into account patterns of participation and loss across the entire life course of the panel.

At the same time, attrition processes are not uniform across panel designs. The mechanisms driving attrition depend on what entities are followed over time, how they are tracked, and how changes in respondents' life circumstances intersect with the sampling design. Lepkowski and Couper (2002) conceptualize unit nonresponse in later waves of longitudinal household surveys as the result of three distinct processes: non-location, non-contact, and refusal to cooperate, each driven by different mechanisms and thus requiring separate consideration. Non-location refers to cases in which sample members cannot be found in later waves because their contact information has changed, most commonly due to residential mobility (Minderop and Weiß 2022), the time interval between waves and the survey organization's efforts in tracking sample members (Uhrig 2008). Non-contact occurs when sample members are eligible and locatable but cannot be reached during fieldwork. This process is shaped by patterns of presence at home, physical accessibility, household composition, and interviewer effort, and is therefore closely linked to fieldwork organization and contact strategies (Groves and Couper 1998; Nicoletti and Peracchi 2005; Uhrig 2008). Refusal is the active decision not to participate after contact has been established and represents the largest share of nonresponse in many panel surveys (Lipps 2009). While prior participation generally increases continued cooperation (Nicoletti and Peracchi 2005), refusal is selectively associated with individual characteristics, social integration (Lipps 2009; Lugtig 2014; Uhrig 2008), attitudes toward surveys, and prior survey experiences (Gummer and Daikeler 2020; Struminskaya 2014).

While this framework provides a useful analytical distinction between different sources of panel nonresponse, it is important to note that it is rooted in the logic of household- and person-based panel designs. In these designs, individuals are the primary sampling, with attrition largely understood as the result of changes in respondents' availability, willingness, or contactability. Consequently, attrition is commonly treated as a respondent-driven process, and strategies for panel maintenance typically focus on tracking individuals, reducing refusals, and correcting resulting imbalances through weighting or adjustment.

In the Cologne Dwelling Panel, these assumptions do not directly apply, due to a sampling design that defines dwellings rather than individuals as the primary sampling units. Dwellings are followed over time to capture neighborhood change, while the current residents act as "speakers" for the dwelling, or secondary sampling units, providing information on both the dwelling and their own characteristics. This design choice introduces a dual structure of eligibility and nonresponse: attrition may occur because a dwelling becomes ineligible or inaccessible, because no eligible resident can be contacted at a given wave, or because the current resident refuses participation. At the same time, individual residents who move out of a sampled dwelling automatically exit the panel by design, even though the primary sampling unit remains part of the study. As a result, panel loss cannot be understood solely as respondent dropout but must be interpreted in relation to the interaction between residential mobility, dwelling stability, and fieldwork implementation. These challenges are further shaped by early design decisions, such as the use of natural areas instead of standardized statistical units, which reduce the availability of external contextual information and constrain the analysis of how structural characteristics influence attrition.

Against this background, the central methodological question is not whether attrition occurs in the Cologne Dwelling Panel, but how it unfolds under a design in which eligibility, nonresponse, and sample loss operate at multiple levels and are shaped by early design decisions. Rather than treating attrition as a single outcome driven by individual respondents, the focus shifts to identifying how different mechanisms of panel loss are associated with characteristics of both dwellings and residents, and how these mechanisms accumulate over time. The analysis by Schmidt and Barth (2026) addresses this challenge by examining predictors of nonresponse at the individual and dwelling level, thereby offering an alternative perspective on panel maintenance that is tailored to non-standard sampling units and constrained design settings.

The findings confirm that attrition mechanisms behave differently between both levels of sampling units. Individual nonresponse follows known patterns for attrition at respondents' level and is primarily associated with characteristics of the residents' life situations and attachment to their dwelling and neighborhood. Shifting the view to the dwellings shows that nonresponse is also associated with their structural conditions and local context. However, these mechanisms do not operate independently: residential mobility, meaning changes in dwelling occupancy, plays a role in shaping which units remain observable over time. As a consequence, attrition in the Cologne Dwelling Panel cannot be interpreted as a uniform loss of respondents, but as a patterned process that selectively affects different types of dwellings and residents.

From a panel maintenance perspective, these findings underscore that attrition monitoring must be adapted to the underlying survey design and its constraints. In panels with non-standard primary sampling units, traditional approaches that focus exclusively on individual-level predictors or wave-specific response rates are insufficient to capture how longitudinal data quality evolves. Instead, maintaining data quality requires continuous monitoring of how design decisions, fieldwork implementation, and structural change jointly shape patterns of panel loss over time. More generally, this shows that panel attrition is not merely a problem to be minimized, but a process to be understood and contextualized. Recognizing attrition as cumulative and design-dependent is therefore essential for assessing the limits of inference and the interpretability of longitudinal findings in complex panel studies. This perspective also shows that not all challenges in longitudinal fieldwork need straightforward solutions through intervention, but in some cases, analytical accommodation and transparent documentation are more appropriate responses.

5. Discussion

Making up a central component of empirical evidence in sociological research, surveys are tools through which social reality is made observable, comparable, and analyzable. The quality of survey data is therefore directly related to the scope and validity of sociological evidence. As shown throughout the survey methodological literature, data quality is a multidimensional and highly complex concept that can be understood as the result of a sequence of design and implementation decisions involving unavoidable trade-offs between competing objectives, such as accuracy, comparability, cost efficiency, and feasibility. The TSE framework provides a useful perspective for conceptualizing survey quality as the cumulative result of multiple error sources. Following this idea, it can be said that there is no survey without error, just as there is no single way to conduct survey research. Any survey design necessarily prioritizes certain dimensions of quality over others, and these priorities are shaped by substantive research objectives, target populations, institutional constraints, and fieldwork conditions.

While surveys are designed with these considerations in mind, the empirical chapters of this dissertation demonstrate how data quality is ultimately negotiated in the implementation of fieldwork. Design plans are confronted with real-world conditions, unexpected disruptions, and practical limitations, which require adaptations that may alter the balance of errors initially anticipated. Fieldwork thus emerges not as a neutral execution phase, but as a central site at which data quality is actively produced through decisions about how to respond to constraints, risks, and opportunities.

This dissertation examines these processes by focusing on three domains of fieldwork decisions: interview conditions, recruitment strategies, and the monitoring of panel attrition. Drawing on empirical evidence from the Cologne Dwelling Panel, the studies address how adaptations introduced under constrained conditions affect different dimensions of data quality.

A first contribution of this dissertation is to provide empirical evidence showing how small modifications in the implementation of fieldwork can be used to preserve mode consistency and, in turn, maintain longitudinal continuity in the panel without compromising measurement quality (Schmidt and Estermann 2026). Many panels struggled to continue fieldwork as planned and introduced drastic design changes that effectively split panel histories into periods before and after the COVID-19 pandemic. In contrast, the first paper in this dissertation offers an alternative path. Beyond this context, the findings of this paper can be applied to survey settings in which face-to-face data collection must be conducted under heightened health, safety, or trust constraints, such as research involving at-risk populations, healthcare settings, or situations in which physical proximity is perceived as sensitive. In these contexts, adapting interview conditions while maintaining the face-to-face mode may offer a viable strategy for balancing measurement quality, ethical considerations, and fieldwork feasibility.

A second contribution lies in the empirical analysis of recruitment strategies aimed at addressing nonresponse among analytically relevant but hard-to-reach populations. By experimentally varying the use of incentives and survey modes in a survey of out-movers, the second paper demonstrates how participation gains depend not only on the choice of recruitment tools, but also on how these tools interact with the characteristics of the target population (Schmidt et al. 2025). Rather than treating higher response rates as a straightforward indicator of improved data quality, the findings highlight that incentives can simultaneously increase participation and reshape sample composition, with implications for nonresponse bias and survey costs. This contribution underscores the importance of evaluating recruitment strategies in relation to substantive research objectives, and it provides empirical support for a more differentiated assessment of incentives as selective fieldwork tools rather than universally optimal solutions. In addition, the paper is among the few studies reporting actual costs per completed interview, thereby enabling researchers to make more informed and transparent decisions about the trade-offs between participation gains and financial resources.

A third contribution of this dissertation is to provide empirical evidence on the analysis and monitoring of panel attrition in longitudinal surveys with non-standard primary sampling units. By examining nonresponse processes at both the dwelling and resident levels, the third paper demonstrates that attrition in such designs cannot be understood solely as respondent dropout, but results from the interaction between residential mobility, dwelling characteristics, and fieldwork implementation (Schmidt and Barth 2026). The paper offers an alternative perspective by emphasizing the systematic evaluation of sample loss under design-induced constraints, but it does not provide strategies to reduce attrition. This contribution highlights the importance of adapting attrition monitoring strategies to the underlying survey design and shows how early design decisions shape not only attrition patterns, but also the interpretability of longitudinal findings over time. It reinforces that not all fieldwork problems can be resolved through intervention, but may instead need analytical accommodation and careful interpretation.

Taken together, these papers constitute the main conceptual contribution of this dissertation. By shifting the perspective towards fieldwork as a site of data production rather than a mere execution step, survey data quality can be understood as the outcome of cumulative, context-specific decisions that interact across different stages of the survey process. Understanding fieldwork in this manner allows for the transparent and realistic assessment of trade-offs in survey implementation. This perspective is particularly relevant for longitudinal research, where data quality cannot be separated from design history and where adaptations always take place within constraints created by earlier methodological choices. Effective fieldwork management requires naming and documenting trade-offs explicitly, and evaluating interventions in relation to survey goals, rather than relying on single quality indicators.

This dissertation also offers practical contributions for survey organizations and survey practitioners designing and managing longitudinal surveys. By empirically examining fieldwork decisions under constrained conditions, the findings offer some guidance for the implementation of panels that are sufficiently flexible to accommodate necessary adaptations while remaining coherent in their core design logic. The results further underscore the importance of evaluating fieldwork interventions beyond overall response rates by also considering their implications for measurement quality, sample composition, and survey costs. Finally, the analyses highlight the need for transparent documentation of design constraints and adaptation choices, as this allows data users to better assess interpretive limits and supports more informed use of survey data in substantive research.

The findings of this dissertation should be interpreted in light of several limitations. First, all empirical analyses are based on a single longitudinal study with a specific design and research objective, namely an area-based panel with dwellings as primary sampling units. While this design is well suited to studying neighborhood change and residential mobility, it limits the generalizability of the findings to surveys that follow individuals or households across changing locations. Second, the empirical context of the Cologne Dwelling Panel introduces additional constraints that shape the observed trade-offs. The study is embedded in two urban neighborhoods in Germany and reflects institutional, legal, and cultural conditions specific to this setting. Adaptations in response to the COVID-19 pandemic were influenced by both the timing of the survey wave and the regulatory environment at that stage. As a result, the feasibility and consequences of similar fieldwork decisions may differ in other national contexts or under different external conditions. Third, the assessment of data quality relies on observable indicators and available auxiliary information. While the studies examine multiple dimensions of survey quality, including participation, sample composition, and indicators of measurement quality, some error processes, such as other forms of measurement bias or unobserved mechanisms of attrition, were not measured. The selectivity of observed indicators is inherent to empirical research on survey implementation and reinforces the importance of cautious interpretation of quality indicators. Finally, the dissertation focuses primarily on fieldwork decisions and adaptations, rather than earlier stages of survey design such as questionnaire development or sampling frame construction. Although these stages are closely linked to later implementation, the analytical emphasis of this work lies on how design choices unfold and are negotiated during fieldwork. Future research could extend this perspective by examining how design-stage decisions and fieldwork adaptations jointly shape survey data quality over longer time frames.

Nonetheless, the findings emphasize that survey data quality, particularly in longitudinal research, is inseparable from the fieldwork decisions leading to its production. Making these decisions analytically visible is essential for interpreting empirical findings.

6. Literatur

AAPOR. 2022. *Best Practices for Survey Research*. Retrieved December 30, 2025 (https://aapor.org/wp-content/uploads/2023/06/Survey-Best-Practices.pdf).

Abdelazeem, Basel, Aboalmagd Hamdallah, Marwa A. Rizk, Kirellos S. Abbas, Nahla A. El-Shahat, Nouraldeen Manasrah, Mostafa R. Mostafa, and Mostafa Eltobgy. 2023. "Does Usage of Monetary Incentive Impact the Involvement in Surveys? a Systematic Review and Meta-Analysis of 46 Randomized Controlled Trials." *PloS one* 18(1)e0279128. doi:10.1371/journal.pone.0279128.

Amaya, Ashley, Paul P. Biemer, and David Kinyon. 2020. "Total Error in a Big Data World: Adapting the TSE Framework to Big Data." *Journal of Survey Statistics and Methodology* 8(1):89–119. doi:10.1093/jssam/smz056.

Atakan, Rebekka. 2023. "Dimensionen der Raumwahrnehmung – eine relationale Analyse in Vierteln im Gentrifizierungsprozess." *Psychologie & Gesellschaftskritik* 47:213–48. doi:10.2440/007-0003.

Atakan, Rebekka, and Alice Barth. 2025. "Interrelations of Social, Physical and Symbolic Space – Assessing Residents' Spatial Perceptions of Gentrifying Neighbourhoods with Multiple Factor Analysis." *Housing, Theory and Society* 42(3):312–34. doi:10.1080/14036096.2024.2406826.

Barth, Alice, and Jörg Blasius. 2023. "Assessing Rental Price Dynamics in Two Gentrified Neighbourhoods in Cologne by Means of a Dwelling Panel." *Survey Research Methods* 17(3):395–410. doi:10.18148/SRM/2023.V17I3.7987.

Biemer, Paul P., and Lars Lyberg. 2003. *Introduction to Survey Quality*. Hoboken, NJ: John Wiley & Sons.

Biemer, Paul P. 2010. "Total Survey Error: Design, Implementation, and Evaluation." *Public Opinion Quarterly* 74(5):817–48. doi:10.1093/poq/nfq058.

Biemer, Paul P., Joe Murphy, Stephanie Zimmer, Chip Berry, Grace Deng, and Katie Lewis. 2018. "Using Bonus Monetary Incentives to Encourage Web Response in Mixed-Mode Household Surveys." *Journal of Survey Statistics and Methodology* 6(2):240–61. doi:10.1093/jssam/smx015.

Blair, Johnny, Ronald Czaja, and Edward Blair. 2013. *Designing Surveys: A Guide to Decisions and Procedures*. 3rd ed. Los Angeles, Calf.: SAGE Publications.

Blasius, Jörg, and Karl-Heinz Reuband. 1995. "Telefoninterviews in der empirischen Sozialforschung: Ausschöpfungsquoten und Antwortqualität." *ZA-Information / Zentralarchiv für Empirische Sozialforschung* 37:64–87.

Blasius, Jörg, and Jürgen Friedrichs. 2016. “Gentrification – Verlauf und Wahrnehmung durch die Bewohner.” *GWP – Gesellschaft. Wirtschaft. Politik* 65(3):347–58. doi:10.3224/gwp.v65i3.08.

Blasius, Jörg, Jürgen Friedrichs, and Heiko Rühl. 2016a. “Gentrifikation in zwei Wohngebieten von Köln.” *KZfSS Kölner Zeitschrift für Soziologie und Sozialpsychologie* 68(3):541–59. doi:10.1007/s11577-016-0372-7.

Blasius, Jörg, Jürgen Friedrichs, and Heiko Rühl. 2016b. “Pioneers and gentrifiers in the process of gentrification.” *International Journal of Housing Policy* 16(1):50–69. doi:10.1080/14616718.2015.1071029.

Blasius, Jörg, and Alice Barth. 2025. “Kölner Wohnungspanel (Version 2.0.0) [Data set].”.

Bourdieu, Pierre. 2017[1993]. “Verstehen.” Pp. 779–822, in *Das Elend der Welt: Zeugnisse und Diagnosen alltäglichen Leidens an der Gesellschaft,* edited by P. Bourdieu and et al. Köln: Herbert von Halem.

Brick, J. M., and Douglas Williams. 2013. “Explaining Rising Nonresponse Rates in Cross-Sectional Surveys.” *The ANNALS of the American Academy of Political and Social Science* 645(1):36–59. doi:10.1177/0002716212456834.

Brick, J. M., and Roger Tourangeau. 2017. “Responsive Survey Designs for Reducing Nonresponse Bias.” *Journal of Official Statistics* 33(3):735–52. doi:10.1515/jos-2017-0034.

Burton, Jonathan, Peter Lynn, and Michaela Benzeval. 2020. “How Understanding Society: the UK Household Longitudinal Study Adapted to the COVID-19 Pandemic.” *Survey Research Methods* 14(2):235–59. doi:10.18148/srm/2020.v14i2.7746.

Cavari, Amnon, and Guy Freedman. 2023. “Survey Nonresponse and Mass Polarization: The Consequences of Declining Contact and Cooperation Rates.” *American Political Science Review* 117(1):332–39. doi:10.1017/S0003055422000399.

Church, Allan H. 1993. “Estimating the Effect of Incentives on Mail Survey Response Rates: A Meta-Analysis.” *Public Opinion Quarterly* 57(1):62–79. doi:10.1086/269355.

Couper, Mick P. 2013. “Is the Sky Falling? New Technology, Changing Media, and the Future of Surveys.” *Survey Research Methods* 7(3):145–56. doi:10.18148/srm/2013.v7i3.5751.

de Leeuw, Edith, and Peter Lugtig. 2014. “Dropouts in Longitudinal Surveys.” Pp. 1–6, in *Wiley StatsRef: Statistics Reference Online,* edited by N. Balakrishnan, T. Colton, B. Everitt, W. Piegorsch, F. Ruggeri, and J. L. Teugels. Hoboken, N.J: John Wiley & Sons.

de Leeuw, Edith, Joop Hox, and Annemieke Luiten. 2018. “International Nonresponse Trends across Countries and Years: An analysis of 36 years of Labour Force Survey data.” *Survey Methods: Insights from the Field*. doi:10.13094/SMIF-2018-00008.

Dillman, Don A. 2009. “Some Consequences of Survey Mode Changes in Longitudinal Surveys.” Pp. 127–40 in *Wiley series in survey methodology, Methodology of longitudinal surveys,* edited by P. Lynn. Chichester, West Sussex: Wiley.

Dillman, Don A., Jolene D. Smyth, and Leah M. Christian. 2014. *Internet, Phone, Mail, and Mixed-Mode Surveys: The Tailored Design Method*. Hoboken, NJ: John Wiley & Sons.

Durkheim, Émile. 1982[1895]. *The rules of Sociological Method: And selected texts on sociology and its method*. New York, NY: The Free Press.

Fowler, Floyd J. 2014[1984]. *Survey Research Methods*. 5th ed. Los Angeles, California: SAGE.

Friedrichs, Jürgen, and Jörg Blasius. 2015. “The Dwelling Panel – A New Research Method for Studying Urban Change.” *Raumforschung und Raumordnung* 73(6):377–88. doi:10.1007/s13147-015-0369-0.

Friedrichs, Jürgen, and Jörg Blasius. 2020. “Neighborhood change – results from a dwelling panel.” *Housing Studies* 35(10):1723–41. doi:10.1080/02673037.2019.1699032.

Groves, Robert M., and Mick Couper. 1998. *Nonresponse in Household Interview Surveys*. New York, NY: John Wiley & Sons.

Groves, Robert M., Eleanor Singer, and Amy D. Corning. 2000. “Leverage-Saliency Theory of Survey Participation: Description and an Illustration.” *Public Opinion Quarterly* 64(3):299–308.

Groves, Robert M. 2004. *Survey Errors and Survey Costs*. Hoboken, N.J: Wiley.

Groves, Robert M., and Steven G. Heeringa. 2006. “Responsive Design for Household Surveys: Tools for Actively Controlling Survey Errors and Costs.” *Journal of the Royal Statistical Society Series A: Statistics in Society* 169(3):439–57. doi:10.1111/j.1467-985X.2006.00423.x.

Groves, Robert M., Floyd J. Fowler, Mick Couper, James M. Lepkowski, Eleanor Singer, and Roger Tourangeau. 2009. *Survey Methodology*. Hoboken, NJ: John Wiley & Sons.

Groves, Robert M., and Lars Lyberg. 2010. “Total Survey Error: Past, Present, and Future.” *Public Opinion Quarterly* 74(5):849–79. doi:10.1093/poq/nfq065.

Groves, Robert M. 2011. “Three Eras of Survey Research.” *Public Opinion Quarterly* 75(5):861–71. doi:10.1093/poq/nfr057.

Gummer, Tobias, and Jessica Daikeler. 2020. “A Note on How Prior Survey Experience With Self-Administered Panel Surveys Affects Attrition in Different Modes.” *Social Science Computer Review* 38(4):490–98. doi:10.1177/0894439318816986.

Gummer, Tobias, Karolina von Glasenapp, Thomas Skora, Elias Naumann, Saskia Bartholomäus, and Sophia Piesch. 2025a. “Survey Data Collection during the COVID-19 Pandemic in Germany: Recommendations for an Improved Data Collection Infrastructure.” *Health Policy* 165:105551. doi:10.1016/j.healthpol.2025.105551.

Gummer, Tobias, Karolina von Glasenapp, Thomas Skora, Saskia Bartholomäus, and Elias Naumann. 2025b. “The Impact of the COVID-19 Pandemic on the Design of Repeated Cross-Sectional and Panel Surveys in Germany.” *International Journal of Public Opinion Research* 37(4). doi:10.1093/ijpor/edaf062.

Hamnett, Chris. 1991. “The Blind Men and the Elephant: The Explanation of Gentrification.” *Transactions of the Institute of British Geographers* 16(2):173–89. doi:10.2307/622612.

Hansen, Morris H., and William N. Hurwitz. 1946. “The Problem of Non-Response in Sample Surveys.” *Journal of the American Statistical Association* 41(236):517–29.

Kreuter, Frauke. 2025. “Modernizing Data Collection.” *Journal of Official Statistics* 41(3):863–72. doi:10.1177/0282423X251318452.

Lepkowski, James M., and Mick P. Couper. 2002. “Nonresponse in the Second Wave of Longitudinal Household Surveys.” Pp. 259–72 in *Wiley series in survey methodology, Survey nonresponse,* edited by R. M. Groves, D. A. Dillman, J. L. Eltinge, and R. J. A. Little. New York, NY: John Wiley & Sons.

Lipps, Oliver. 2009. “Attrition of Households and Individuals in Panel Surveys.” *SOEPpapers on Multidisciplinary Panel Data Research* 164.

Lugtig, Peter. 2014. “Panel Attrition.” *Sociological Methods & Research* 43(4):699–723. doi:10.1177/0049124113520305.

Luiten, Annemieke, and Barry Schouten. 2013. “Tailored Fieldwork Design to Increase Representative Household Survey Response: An Experiment in the Survey of Consumer Satisfaction.” *Journal of the Royal Statistical Society Series A: Statistics in Society* 176(1):169–89. doi:10.1111/j.1467-985X.2012.01080.x.

Lyberg, Lars. 2012. “Survey Quality.” *Survey Methodology* 38(2):107–30.

Lynn, Peter. 2009. “Methods for Longitudinal Surveys.” Pp. 1–19 in *Wiley series in survey methodology, Methodology of longitudinal surveys,* edited by P. Lynn. Chichester, West Sussex: Wiley.

Lynn, Peter, and Peter Lugtig. 2017. “Total Survey Error for Longitudinal Surveys.” Pp. 279–98, In *Total Survey Error in Practice,* edited by P. P. Biemer, E. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. Lyberg, C. N. Tucker, and B. T. West. Hoboken, New Jersey: John Wiley & Sons Inc.

Maynard, Douglas W., and Nora C. Schaeffer. 2002. “Standardization and Its Discontents.” Pp. 3–46 in *Wiley series in survey methodology, Standardization and tacit knowledge: Interaction and Practice in the Survey Interview,* edited by D. W. Maynard. New York, NY: Wiley.

Minderop, Isabella, and Bernd Weiß. 2022. “Now, Later, or Never? Using Response-Time Patterns to Predict Panel Attrition.” *International Journal of Social Research Methodology* :1–14. doi:10.1080/13645579.2022.2091259.

Newman, Kathe, and Elvin K. Wyly. 2006. “The Right to Stay Put, Revisited: Gentrification and Resistance to Displacement in New York City.” *Urban Studies* 43(1):23–57. doi:10.1080/00420980500388710.

Nicoletti, Cheti, and Franco Peracchi. 2005. “Survey Response and Survey Characteristics: Microlevel Evidence from the European Community Household Panel.” *Journal of the Royal Statistical Society Series A: Statistics in Society* 168(4):763–81. doi:10.1111/j.1467-985X.2005.00369.x.

Penker, Matthias, Anja Eder, and Markus Hadler. 2025. “From Face-to-Face to Push-to-Web: Methodological Lessons from the Transition of the “Social Survey Austria”.”. doi:10.13094/SMIF-2025-00012.

Peytchev, Andy. 2013. “Consequences of Survey Nonresponse.” *The ANNALS of the American Academy of Political and Social Science* 645(1):88–111. doi:10.1177/0002716212461748.

Pforr, Klaus, Michael Blohm, Annelies G. Blom, Barbara Erdel, Barbara Felderer, Mathis Fräßdorf, Kristin Hajek, Susanne Helmschrott, Corinna Kleinert, Achim Koch, Ulrich Krieger, Martin Kroh, Silke Martin, Denise Saßenroth, Claudia Schmiedeberg, Eva-Maria Trüdinger, and Beatrice Rammstedt. 2015. “Are Incentive Effects on Response Rates and Nonresponse Bias in Large-scale, Face-to-face Surveys Generalizable to Germany? Evidence from Ten Experiments.” *Public Opinion Quarterly* 79(3):740–68. doi:10.1093/poq/nfv014.

Popper, Karl R. 2002[1934]. *The Logic of Scientific Discovery*. London: Routledge Classics.

Robinson, Claude E. 1937. “Recent Developments in the Straw-Poll Field.” *Public Opinion Quarterly* 1(3):45. doi:10.1086/265091.

Scherpenzeel, Annette, Kathrin Axt, Michael Bergmann, Salima Douhou, Andrea Oepen, Gregor Sand, Karin Schuller, Stephanie Stuck, Melanie Wagner, and Axel Börsch-Supan. 2020. “Collecting Survey Data Among the 50+ Population During the COVID-19 Outbreak: the Survey of Health, Ageing and Retirement in Europe (SHARE).” *Survey Research Methods* 14(2):217–21. doi:10.18148/srm/2020.v14i2.7738.

Schmidt, Manuela, Alice Barth, and Jörg Blasius. 2025. “Effects of mode and incentives on response rate, sample composition, and costs – experience from a self-administered mixed-mode survey of movers.” *Survey Methods: Insights from the Field*. doi:10.13094/SMIF-2025-00002.

Schmidt, Manuela, and Karoline R. Estermann. 2026. “The Impact of Face Mask Usage on Data Quality in Face-to-face Interviews.” *Field Methods* 38(1):19–32. doi:10.1177/1525822X251348080.

Schmidt, Manuela, and Alice Barth. 2026. “Unit nonresponse in a dwelling panel study – the role of residents’ and housing characteristics.” *methods, data, analyses*. doi:10.25521/mda.595.

Schouten, Barry, Melania Calinescu, and Annemieke Luiten. 2013. “Optimizing quality of response through adaptive survey designs.” *Survey Methodology* 39(1):29–58.

Schouten, Barry, Andy Peytchev, and James Wagner. 2017. *Adaptive Survey Design*. Boca Raton, FL: Chapman and Hall/CRC.

Schröder, Jette, Claudia Schmiedeberg, Josef Brüderl, and Bozoyan Christiane. 2025. “Switching a Face-to-Face Panel to Self-Administered Survey Modes: Experimental Evidence on Effects of Mode Assignment on Response and Selectivity.” *Survey Research Methods* 19(1):1–11. doi:10.18148/srm/2025.v19i1.8427.

Singer, Eleanor, Robert M. Groves, and Amy D. Corning. 1999. “Differential incentives: Beliefs about practices, perceptions of equity, and effects on survey participation.” *Public Opinion Quarterly* 63(2):251–60.

Singer, Eleanor, and Cong Ye. 2013. “The Use and Effects of Incentives in Surveys.” *The ANNALS of the American Academy of Political and Social Science* 645(1):112–41. doi:10.1177/0002716212458082.

Smith, Tom W. 1995. “Trends in non-response rates.” *International Journal of Public Opinion Research* 7(2):157–71.

Stadtmüller, Sven, Christoph Beuthner, Pablo Christmann, Tobias Gummer, Rebekka Kluge, Matthias Sand, and Henning Silber. 2023. “The interplay of incentives and mode-choice design in self-administered mixed-mode surveys." *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique* 159(1):49–74. doi:10.1177/07591063231184243.

Stein, Annika, Tobias Gummer, Elias Naumann, Björn Rohr, Henning Silber, Roman Auriga, Michael Bergmann, Arne Bethmann, Michael Blohm, Carina Cornesse, Pablo Christmann, Mustafa Coban, Jean P. Décieux, Britta Gauly, Caroline Hahn, Susanne Helmschrott, Oshrat Hochman, Johannes Lemcke, Dörte Naber, Steffen Pötzschke, Joss Roßmann, Jan-Lucas Schanze, Tobias Schmidt, Silke L. Schneider, Heike Spangenberg, Tobias Rettig, Mark Trappmann, Michael Weinhardt, and Bernd Weiß. 2025. "Education bias in probability-based surveys in Germany: evidence and possible solutions." *International Journal of Social Research Methodology* :1–18. doi:10.1080/13645579.2025.2508889.

Struminskaya, Bella. 2014. *Data Quality in Probability-Based Online Panels: Nonresponse, Attrition, and Panel Conditioning: PhD Diss.* Utrecht: Utrecht University.

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

Tourangeau, Roger, J. Michael Brick, Sharon Lohr, and Jane Li. 2017. "Adaptive and Responsive Survey Designs: A Review and Assessment." *Journal of the Royal Statistical Society Series A: Statistics in Society* 180(1):203–23. doi:10.1111/rssa.12186.

Uhrig, S. C. N. 2008. *The Nature and Causes of Attrition in the British Household Panel Study.* ISER Working Paper Series. 2008-05. Colchester: University of Essex, Institute for Social and Economic Research (ISER).

Warren, John R., and Andrew Halpern-Manners. 2012. "Panel Conditioning in Longitudinal Social Science Surveys." *Sociological Methods & Research* 41(4):491–534. doi:10.1177/0049124112460374.

Weber, Max. 1922. "Methodische Grundlagen der Soziologie." Pp. 503–23, in *Gesammelte Aufsätze zur Wissenschaftslehre,* edited by M. Weber. Tübingen: J. C. B. Mohr (Siebeck).

Weisberg, Herbert F. 2005. *The Total Survey Error Approach: A Guide to the New Science of Survey Research*. Chicago, IL: The University of Chicago Press.

West, Brady T., and Annelies G. Blom. 2017. "Explaining Interviewer Effects: A Research Synthesis." *Journal of Survey Statistics and Methodology* 5(2)175-211. doi:10.1093/jssam/smw024.

Westwood, Sean J. 2025. "The Potential Existential Threat of Large Language Models to Online Survey Research." *Proceedings of the National Academy of Sciences of the United States of America* 122(47)e2518075122. doi:10.1073/pnas.2518075122.

Will, Gisela, Regina Becker, and Dominik Weigand. 2020. "COVID-19 Lockdown During Field Work." *Survey Research Methods* 14(2):247–52. doi:10.18148/srm/2020.v14i2.7753.

Wolf, Christof, Pablo Christmann, Tobias Gummer, Christian Schnaudt, and Sascha Verhoeven. 2021. "Conducting General Social Surveys as Self-Administered Mixed-Mode Surveys." *Public Opinion Quarterly* 85(2):623–48. doi:10.1093/poq/nfab039.

7. Artikel

Parts of this thesis have been published in peer-reviewed journals:

7.1 Schmidt, Manuela, and Karoline R. Estermann. 2026. The Impact of Face Mask Usage on Data Quality in Face-to-face Interviews.

Schmidt, Manuela, and Karoline R. Estermann. 2026. "The Impact of Face Mask Usage on Data Quality in Face-to-face Interviews." *Field Methods* 38(1):19–32. First published online October 17, 2025. doi:10.1177/1525822X251348080.

7.2 Schmidt, Manuela, Alice Barth, and Jörg Blasius. 2025. Effects of mode and incentives on response rate, sample composition, and costs – experience from a self-administered mixed-mode survey of movers.

Schmidt, Manuela, Alice Barth, and Jörg Blasius. 2025. "Effects of mode and incentives on response rate, sample composition, and costs – experience from a self-administered mixed-mode survey of movers." *Survey Methods: Insights from the Field*. doi:10.13094/SMIF-2025-00002.

7.3 Schmidt, Manuela, and Alice Barth. 2026. Unit nonresponse in a dwelling panel study – the role of residents' and housing characteristics.

Schmidt, Manuela, and Alice Barth. 2026. "Unit nonresponse in a dwelling panel study – the role of residents' and housing characteristics." *methods, data, analyses*. doi: 10.25521/mda.595.