

**Methodological Challenges
in Family Surveys and Survey Research**

Inaugural dissertation submitted in partial fulfillment of the requirements
for the degree Doctor of Social Sciences at the University of Mannheim

Submitted by
Almut Schumann

Dean of the School of Social Sciences:

Dr. Julian Dierkes

Supervisors:

Prof. Dr. Tobias Gummer

Prof. Dr. Karsten Hank

Evaluators:

Prof. Dr. Tobias Gummer

Prof. Dr. Karsten Hank

Prof. Dr. Thomas Gautschi

Date of the defense: March 17, 2025

Acknowledgements

I have always been captivated by the study of relationships, both within partnerships and families. The dynamics and processes that shape these relationships are a fundamental part of our lives, often influencing us in profound ways. As a quantitative researcher, I find it exciting to identify patterns and structures based on large datasets that provide empirical evidence into the realities of these relationships. Yet, as a survey researcher, I am acutely aware that these insights are only possible through respondents' willingness to participate in surveys and share glimpses into their intimate lives. For this reason, I extend my deepest gratitude to those anonymous individuals who have shared and continue to share their stories through surveys.

I would like to express my sincerest thanks to my supervisors, Tobias Gummer and Karsten Hank – from day one, it was a perfect match. I am deeply grateful for your steadfast support, thoughtful feedback, and exceptional guidance throughout the course of this research. Your mentorship has been truly invaluable, providing me with the knowledge and encouragement that significantly contributed to the success of this work and expanded my understanding of research. I would also like to extend a special thanks to Thomas Gautschi for serving as a third expert on my review committee. The time and effort you all have dedicated to this dissertation is deeply appreciated.

I am also profoundly grateful to my co-authors. To Detlev Lück, thank you for always having an open ear and for sharing your wealth of knowledge and experience. To Claudia Schmiedeberg, thank you for your candid advice and support that helped me to refine my ideas. The collaboration with both of you has been a tremendous asset to my work.

I also wish to express my heartfelt thanks to all of my (former) colleagues who have enriched these past few years. To my colleagues from the FReDA project at the Federal Institute for

Population Research (BiB), the GESIS Leibniz Institute for the Social Sciences, and the pairfam consortium, as well as to my colleagues from the GGP team at the Netherlands Interdisciplinary Demographic Institute (NIDI): Your dedication and collaboration have been truly inspiring. I am grateful for the insights and feedback I gained through our work within these data infrastructures, which have profoundly shaped this dissertation. Above all, I appreciate the support and encouragement from each of you, which has helped me along this journey.

Finally, my deepest gratitude goes to the focal point of this thesis—the intimate relationships I am fortunate to have in my own life. To my parents, siblings, friends, and my partner: Your support has meant everything to me. I am grateful not only to analyze relationships in my research, but to also experience them so deeply in my own life.

Table of Contents

1	General Introduction.....	1
1.1	Theoretical background.....	6
1.1.1	Nonresponse error	8
1.1.2	Measurement error	11
1.2	Practical relevance	13
1.3	Overview of chapters	18
1.3.1	Is consent to further panel participation selective? The case of a self-administered family panel survey announcing institutional change.....	18
1.3.2	Understanding the nonresponse process of secondary respondents: Evidence from a self-administered multi-actor survey.....	20
1.3.3	Better ask online when it concerns intimate relationships? Survey mode differences in the assessment of relationship quality	21
	References	24
2	Is Consent to Further Panel Participation Selective? The Case of a Self-Administered Family Panel Survey Announcing Institutional Change	35
	Abstract.....	35
2.1	Introduction.....	37
2.2	Theoretical background and literature	39
2.3	Data and methods.....	43
2.3.1	Data	43
2.3.2	Measurements	45
2.3.3	Methodological approach.....	48
2.4	Results.....	50
2.5	Discussion	56
	References	62
	Appendix	68
3	Understanding the Nonresponse Process of Secondary Respondents: Evidence From a Self-Administered Multi-Actor Survey	81
	Abstract.....	81
3.1	Introduction.....	83
3.2	Nonresponse process for secondary respondents.....	87

3.2.1	Steps leading to nonresponse	87
3.2.2	Predictors of nonresponse	90
3.2.2.1	Consent step	90
3.2.2.2	Participation step.....	93
3.3	Data and methods.....	98
3.3.1	Data	98
3.3.2	Measurements	101
3.3.3	Methodological approach.....	103
3.4	Results.....	105
3.5	Discussion	114
	References	122
	Appendix	131
4	Better Ask Online When It Concerns Intimate Relationships? Survey Mode	
	Differences in the Assessment of Relationship Quality	141
	Abstract.....	141
4.1	Introduction.....	143
4.2	Background and expectations	147
4.2.1	Face-to-face versus web interviewing.....	147
4.2.2	Social desirability bias and relationship quality.....	148
4.3	Data and methods.....	153
4.3.1	Experimental design and case selection.....	153
4.3.2	Methodological approach.....	156
4.3.3	Measurements	158
4.4	Results.....	162
4.4.1	Univariate analyses	162
4.4.2	Multivariate analyses	164
4.5	Discussion	167
	References	172
	Appendix	180
5	General Discussion and Conclusion.....	185
	References	195

List of Figures

1.1	Total survey error framework (adapted from Groves, Fowler, et al., 2004, fig. 2.5).....	6
2.1	LPM regression results of the probability of panel consent.....	52
2.2	LPM regression results of panel consent on relationship satisfaction (subsample: respondents with a partner)	54
2.3	LPM regression results of panel consent on conflict frequency (subsample: respondents with a partner)	56
3.1	Steps in the nonresponse process for secondary respondents	88
3.2	Nonresponse process for secondary respondents adapted to the FReDA study.....	106
3.3	Average marginal effects on the likelihood of providing consent by primary respondents and of participation by secondary respondents	108
A3.1	Invitation letter for secondary respondents	133
A3.2	Question text for obtaining consent and contact information to interview a secondary respondent.....	135
4.1	Effect of web mode on indicators of relationship quality with 95% confidence intervals under block-wise adjustment of demographic and family-related control variables.....	165

List of Tables

1.1	Overview: Design and outreach of three family surveys in Germany (status 2024)....	14
1.2	Overview of chapters.....	17
2.1	Summary statistics of variables under study	50
A2.1	LPM regression results of panel consent for models including only middle birth cohorts (model 1), DemoDiff sample as separate category (model 2); including survey mode (model 3).....	68
A2.2	Question wording of variables in the questionnaire	70
A2.3	Average marginal effects of panel consent	73
A2.4	Summary statistics of variables under study for respondents with a partner	74
A2.5	LPM regression results of panel consent for the full model, models separated by panel experience and an interaction model with panel experience	75
A 2.6	LPM regression results of panel consent on relationship satisfaction (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience	77
A2.7	LPM regression results of panel consent on conflict frequency (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience	79
3.1	Overview of tested hypotheses and whether they can be supported	110
3.2	Predicted probabilities for scenarios at the individual and dyadic level for providing consent and participating while holding other variables at their means.....	111
3.3	Predicted probabilities of participation, while holding other variables at their optimal values, for scenarios at the contact level	113
A3.1	Summary statistics for primary respondents regarding items under study.....	131
A3.2	Summary statistics for secondary respondents regarding items under study	132
A3.3	Descriptive distributions between self-reports by secondary respondents and proxy information by primary respondents	136
A3.4	Average marginal effects on the likelihood of providing consent by primary respondents, separated between CAWI and PAPI interviews.....	137
A3.5	Average marginal effects on the likelihood of providing consent by primary respondents	138
A3.6	Average marginal effects on the likelihood of participating by secondary respondents	139

A3.7	Average marginal effects on the likelihood of participating by secondary respondents, including only characteristics at the individual and dyadic level.....	140
4.1	Overview of the experimental design and case selection	155
4.2	Family-demographic distributions for respondents in the entire sample and for respondents with a partner.....	156
4.3	Means of indicators or percentages of confirmative answers by mode of data collection with 95% confidence intervals	162
A4.1	Distributions and sources of reference information from official statistics used for post-stratification weighting	180
A4.2	Original wording of questions and answers for items measuring relationship quality in the GGP Pilot Study 2018	181
A4.3	Post-stratification weighted means of indicators or percentages of confirmative answers by mode of data collection.....	182
A4.4	Effect of web mode on separate indicators of relationship quality under block-wise adjustment of demographic and family-related confounders	183

1 General Introduction

Studying family lives and intimate relationships is a crucial task in the social sciences for a variety of reasons. Family is among the most important spheres of life for the majority of individuals and has clear long-term effects on their well-being (e.g., Brown, 2010; Myrskylä & Margolis, 2014; Nelson-Coffey, 2018). The ways in which partnerships, parenthood, and other family ties are established and maintained are both important and challenging for shaping one's own life. A profound knowledge on family dynamics helps to understand demographic change as they influence outcomes such as fertility rates, marriage and divorce rates, and household composition (e.g., Esping-Andersen & Billari, 2015; Goldscheider et al., 2015; Seltzer, 2019). At the same time, these outcomes also indicate societal trends and have far-reaching effects on the economy, social stability, policy development, and cultural norms within a society. Thus, family research is integral to advancing scientific research and informing public administration and policy makers.

However, a solid empirical basis is needed for this task. Surveys, the classical tool for data collection in this realm, are still indispensable to gathering evidence-based information in demographic research:

“The sample survey has been a fundamental building block of demographic research. Many key advances in both empirical evidence and theoretical reasoning are founded on information from surveys.” (Axinn et al., 2011, p. 1127)

Accordingly, it is crucial to maintain high quality and comparability in survey data to guarantee valid findings on relevant family demographic outcomes. Survey data collection processes must

therefore ensure that data is accurate, reliable, and fit for its intended use (Couper, 2013; Groves, Fowler, et al., 2004).

Given the developments in survey participation and data collection in recent decades, conducting high-quality, large-scale surveys has become increasingly challenging (Couper, 2017; Krosnick et al., 2015; see Tourangeau, 2004 for an overview). One major development in Western Europe and the United States for almost all types of surveys is declining response rates (e.g., Daikeler et al., 2020; Heer & De Leeuw, 2002; Luiten et al., 2020; Williams & Brick, 2018). Possible contributing factors include increased data protection concerns, respondent fatigue, and reduced civic engagement and scientific trust (Bates et al., 2008; Couper et al., 2008; Kreuter, 2013). At the same time, there is an expansion of alternative modes of data collection and mixed-mode designs replacing interviewer-administered face-to-face and telephone interviews (Couper, 2011; De Leeuw, 2005; Olson et al., 2020; Wolf et al., 2021). Frequently, web surveys are used to increase participation, reduce survey costs, and adapt to technological developments and new modes of communication (Bethlehem & Biffignandi, 2012; Couper, 2013; Tourangeau et al., 2013), particularly mobile devices such as smartphones (Couper et al., 2017; Gummer et al., 2023). The COVID-19 pandemic and its restrictions further accelerated this shift in 2020 by reducing face-to-face contact worldwide. As a result, surveys were forced to adapt their designs, moving away from in-person, on-site interviews (e.g., Burton et al., 2020; Gummer et al., 2020; Sastry et al., 2020). In conclusion, it can be stated that:

“The one constant in survey research seems to be change, and we need to find ways to adapt existing methods and develop new methods, in response to both external changes and methodological research on ways to improve surveys.” (Couper, 2011, p. 905)

Consequently, survey researchers are addressing how developments in survey design and data collection procedures impact data quality, for example concerning response rates, nonresponse patterns, and response accuracy (e.g., Bianchi et al., 2017; Daikeler et al., 2020; Sakshaug et al., 2019). However, participation and response behavior can vary across target populations and questionnaire content (Groves, Presser, & Dipko, 2004; Krumpal, 2013; Tourangeau et al., 2010) making it a challenge to generalize the impact of changes in survey methods on data quality across surveys. This difficulty arises as most surveys focus on specific topics, being developed and designed to address particular research objectives and purposes (AAPOR, 2024). This focus can introduce unique challenges in representation and measurement, increasing vulnerability to nonresponse and measurement errors throughout the survey process.

One branch of focused surveys are family surveys, which aim to understand family processes and dynamics. These collect detailed information about individual biographies, including union and childbirth histories, relationships with romantic partners and family members, as well as personal feelings and behaviors experienced in intimate relationships. However, collecting this type of information from respondents involves specific features and design elements that make conducting family surveys often unique and challenging, distinguishing them from other types of surveys. I will highlight three of the most important design aspects that characterize family surveys in this dissertation.

Firstly, family surveys aim to trace individual life course trajectories to track family life events, such as leaving the parental home, forming and dissolving partnerships, and becoming a parent. Therefore, longitudinal designs are a key element employed in family surveys to collect information at multiple timepoints in respondents' lives, allowing researchers to observe changes over time and understand family processes.

Secondly, family surveys must gather perspectives from different actors within a family or romantic relationship to track interactions and negotiations. Different viewpoints on family events and relationship dynamics enable researchers to understand processes such as parenting styles, conflict management, and decision-making between partners. Consequently, multi-actor designs are increasingly used to examine interactions and dynamics within intimate relationships through, for example, dyadic analyses in collecting data from multiple actors.

Thirdly, family surveys must collect information on sensitive and subjective topics related to relationships to gain deeper insights into respondents' feelings, experiences, and behaviors. Gathering personal experiences and feelings on intimate subjects—such as conflict behavior within partnerships, union dissolution, sexual behaviors, and (in)fertility—is essential to revealing the realities of relationships and to understanding the extent to which family events affect relationship quality and individual well-being. As responding to questions about intimate life may evoke both positive and negative emotions, family surveys must be designed to ensure that respondents feel comfortable and safe during the interview process in order to promote a high level of self-disclosure.

Consequently, family surveys involve complex design elements and sensitive content, making the rich data collected not only critical for family researchers, but also for survey research in general. Family surveys therefore serve as a valuable example of exploring how the content of a study influences respondents' motivation to engage with the survey and respond accurately, particularly due to topic-related factors. This has significant implications for the design and analysis of thematic surveys beyond family studies, especially those that track individuals over time such as household or income panel surveys, and those exploring personal relationships (e.g., social network studies) or addressing sensitive subjects (e.g., health or victimization surveys).

Accordingly, this investigation of family surveys aims to enhance overall survey research by suggesting improvements to the processes for collecting data from the same respondents over an extended period of time, gathering data from multiple interconnected respondents, and collecting private information on sensitive life domains. For instance, obtaining consent for future participation is particularly crucial for collecting panel data. The recruitment and participation process involving multiple respondents poses complex challenges in multi-actor data collection. In addition, measuring accurate responses is particularly difficult when collecting sensitive information prone to social desirability bias in different survey modes. Thus, there is a need to investigate how such processes are affected by changes in survey design or procedures.

The overall research question addressed by this dissertation is: How do changes in survey design and data collection procedures affect the data quality of family surveys? I explore the relationship between participation and response behavior and survey topic, specifically concerning family and partnership life. This work focuses on the challenges posed by three design characteristics often inherent in family surveys: longitudinal design, a multi-actor approach, and the collection of sensitive information. My goal is to investigate whether methodological challenges result in systematic nonresponse and measurement errors that may compromise the comparability of the sensitive data collected from multiple actors in longitudinal family surveys, which will be investigated in three phases.

In the following sections of this first chapter, I provide a theoretical and empirical background to highlight the significance of the research presented in this dissertation. First, I outline the theoretical process explaining how participation and response behavior can impact data quality in a family survey. Next, I demonstrate the practical relevance of the research questions addressed in chapters 2, 3, and 4. More specifically, I present the methodological challenges faced

by established German family surveys in recent years. Finally, I will provide extended summaries of each of these three chapters.

1.1 Theoretical background

This research is based on the total survey error framework developed by Groves, Fowler, et al. (2004) that identifies potential sources of survey error. The framework demonstrates the connection between survey design, data collection, and estimated parameters (see Figure 1.1).

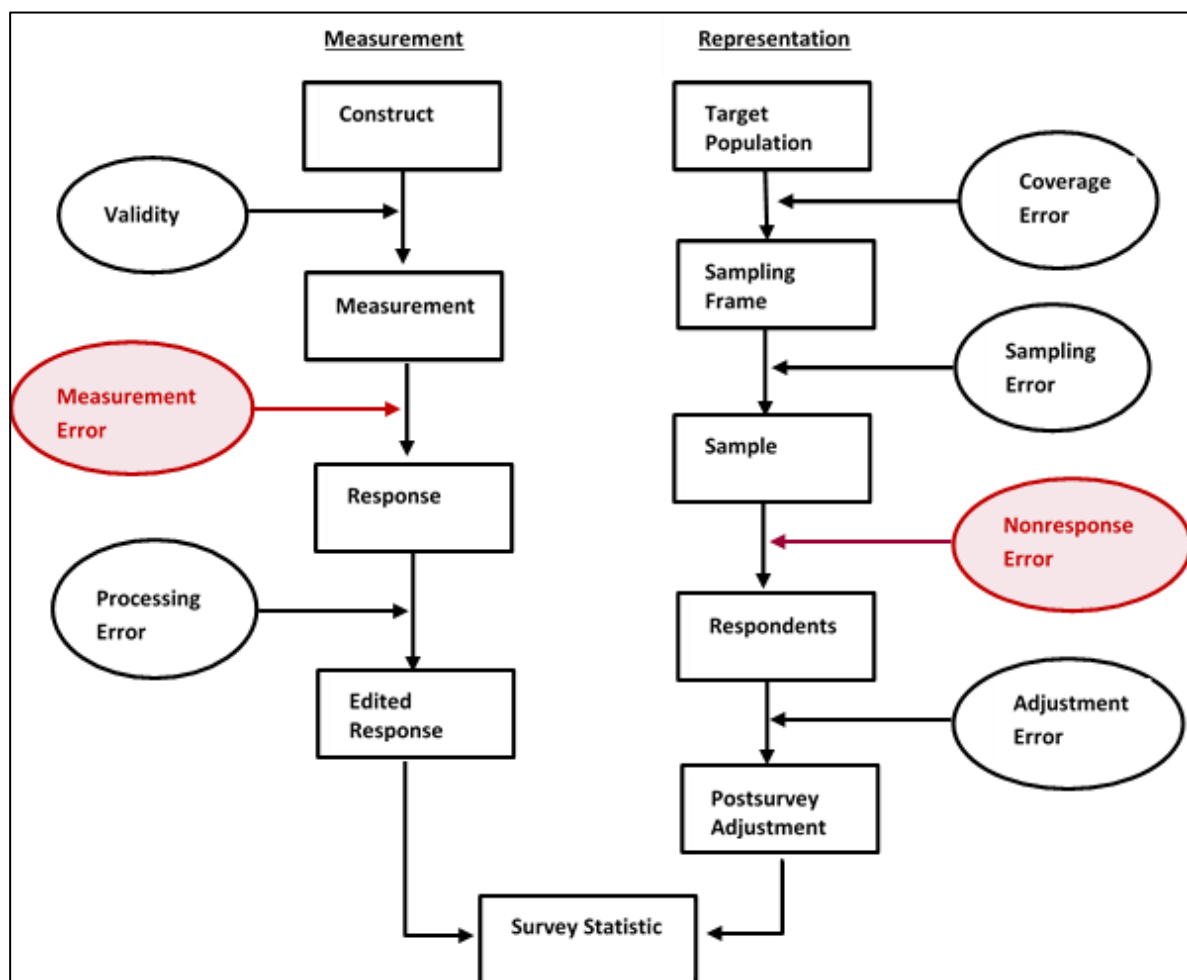


Figure 1.1: Total survey error framework (adapted from Groves, Fowler, et al., 2004, fig. 2.5)

The diagram depicts the lifecycle of a survey on two dimensions: representation and measurement. The “representation” dimension depicts the entire process, beginning from defining the target population to a list of all members of that target population from which the sample is drawn, and finally to the respondents who successfully participate in the survey (Groves, Fowler, et al., 2004). The “measurement” dimension describes the process from a construct of interest to a measurement implemented in the questionnaire, and finally to the resulting response provided by the respondent and recorded in the dataset.

Survey errors are likely to occur along both dimensions, potentially resulting in incorrect estimates in survey statistics. Nonresponse and measurement error are two prominent error types that are frequently addressed and investigated by survey methodologists (e.g., Biemer et al., 2013; Billiet & Matsuo, 2012; Groves, 2006; Gummer, 2019; Sakshaug et al., 2010; West et al., 2017; Yan, 2021). This dissertation concentrates on these two sources of error, as both are rooted in the interaction between the survey and the sample member (see Couper, 1997); in other words, between the survey invitation and the target sample (“representation” dimension) and between the questionnaire and/or interviewer and the respondent (“measurement” dimension).

In both survey processes of participation and responding, success depends mainly on the individual’s evaluation of the survey request. As (consenting to) participation and responding accurately to a question can both be seen as a compliance with a survey request, I use the term “survey request” for both of these actions in the following. According to rational choice theory, individuals face several options for how to act, and choose the option associated with the greatest ratio of benefit and cost for themselves (see Boudon, 2003 for a discussion; Coleman, 1986; see Kroneberg & Kalter, 2012 for a review). The evaluation and interpretation of a survey request takes place in the context of the individual’s personal situation, available resources, interests, and subjective expectations (Esser, 1986a, 1986b). These factors can vary depending

on the survey topic, as certain subjects may provoke more motivation or reluctance in certain individuals. In other words, psychological, social, emotional, and cognitive factors can influence how an individual reacts to the survey topic and, accordingly, handles the survey request. For instance, if the study topic is linked to negative experiences or low interest, such as unfavorable relationship experiences in a family survey or low interest in politics in an election survey, the individual may be less motivated or even reluctant to participate (Groves et al., 1992; Groves, Presser, & Dipko, 2004; Kalmijn, 2023; Stroud & Kenski, 2007). According to the leverage-saliency theory of survey participation adapted by Groves et al. (2000), topic interest is a crucial factor as it impacts the burden or motivation to participate in a survey. Similar cases can be made for response accuracy; for instance, questions about a specific life domain can be perceived as intrusive or sensitive depending on the respondent's personal situation (e.g., recently divorced, unemployed, struggling with mental illness) (Krumpal, 2013; Müller & Schmiedeberg, 2021; Tourangeau & Smith, 1996; Tourangeau & Yan, 2007). As DeMaio (1984) points out, personality and personal experiences as well as item characteristics affect response behavior.

The connection between an individual's situation and the survey topic may thus affect both participation likelihood and response accuracy. This link can lead to systematic nonresponse and measurement errors resulting in biased survey data. I therefore examine to what extent surveys targeting family dynamics and intimate relationships are prone to nonresponse and measurement error, and how this may lead to nonresponse and response bias.

1.1.1 Nonresponse error

Nonresponse error occurs when respondents differ from the entire sampling frame (i.e., respondents vs. nonrespondents) (Groves, Fowler, et al., 2004). To investigate nonresponse error,

it is important to understand the decision to participate in a survey. The motivation to engage in a survey largely depends on individuals' social and psychological predispositions, which correlate with various factors, including sociodemographic characteristics and those related to the survey topic. For some respondents, survey participation requires more effort, placing them at a higher respondent burden when answering cognitively demanding questions or understanding the survey request (Groves et al., 1992; Sharp & Franke, 1983; Yan & Williams, 2022). At the same time, an interesting and salient survey topic can increase the motivation to participate, as questions are perceived as pleasant and the study's purpose as relevant (Groves, Presser, & Dipko, 2004; Keusch, 2013; Zillmann et al., 2014). When nonresponse is not random but appears among specific groups, nonresponse bias is introduced, suggesting that those who do not respond differ systematically from those who do (Groves, Fowler, et al., 2004). Low response rates, (i.e., the proportion of contacted and eligible sample units that respond) (AAPOR, 2023) thus do not necessarily indicate nonresponse bias (Groves & Peytcheva, 2008; Hendra & Hill, 2019; Meterko et al., 2015).

Like most surveys, family surveys face challenges in accurately representing the sociodemographic distribution of a target population. Certain sociodemographic groups generally experience a higher respondent burden due to cognitive difficulties or language barriers, a lack of trust in science and research, or heightened concerns about data protection. Accordingly, sociodemographic groups that are generally associated with a lower likelihood of participation such as migrants, men, or those with low education levels are also likely to be underrepresented in family surveys (Bujard et al., 2023; Fokkema et al., 2016). This may not only affect the sociodemographic representation in the sample, but also induce biases in family-related variables that correlate with gender or educational level, such as fertility decisions or marital dissolution (Jalovaara et al., 2022; Martin, 2006). However, a meta-analysis by Peytcheva and Groves

(2009) revealed that nonresponse for specific demographic groups does not automatically lead to nonresponse bias for substantive indicators being studied in the survey.

In family surveys, the topic of family and partnerships also plays a crucial role for survey engagement as the judgment of the survey request largely depends on an individual's background, experiences, and preferences (Groves et al., 1992), which can vary across topics. When receiving the survey request, sample members are commonly aware of the general contents of the survey because the survey topic and scientific purpose must be mentioned in the invitation letter to comply with legal and ethical requirements. In this situation, contents of a family survey might appeal more to individuals with a partner, with children, and those with positive, stable experiences in their families and partnerships. This suggests a selective bias towards an overrepresentation of “happier” families and partnerships in such surveys. Studies using a Dutch multi-actor survey on family relations have found evidence that closer family ties between parents and children or experiences of divorce within the family have an influence on participation (Kalmijn, 2023; Kalmijn & Liefbroer, 2011). Similarly, findings from a German family panel showed that experiences of separation during a longitudinal study lowered the chances for participation in the next survey waves (Müller & Castiglioni, 2015). These findings support the assumption that the prevailing societal ideal of a stable family unit with strong emotional bonds could discourage individuals with negative relational experiences from participating (further) in a study as they may feel ashamed or uncomfortable revisiting these experiences.

As family-related characteristics are often not available in official statistics and register data, or may even be unobservable as in the case of subjective evaluations and dispositions, it is challenging to investigate biases due to topic-related nonresponse. Studies analyzing topic-related nonresponse therefore rely on panel data and utilize information provided in previous

waves as a reference for participation in consecutive waves. Alternatively, they may use multi-actor designs and information provided by the primary respondents as a reference for the participation of partners or children as secondary respondents.

1.1.2 Measurement error

Measurement errors occur when the “true” value deviates from the value provided by the respondent (Biemer et al., 2013; Groves, Fowler, et al., 2004). There are several reasons why a respondent may not provide an accurate answer, such as difficulties in understanding questions that require high cognitive effort or issues with recalling the information requested (Tourangeau et al., 2000). To shorten the processes of understanding, interpretation, and retrieval, respondents may satisfice; in other words, choose answers that require little cognitive effort (Krosnick, 1991). Response accuracy can also be affected by socially desirable responses when questions are perceived as sensitive (Krumpal, 2013; Tourangeau & Yan, 2007). If such deviations in response behavior occur in a systematic way for a specific type of question, response bias is introduced. Systematic distortions in reporting might lead to an over- or underestimation of the construct to be measured (Groves, Fowler, et al., 2004).

Family surveys cover questions that are particularly prone to measurement errors resulting from satisficing tendencies and socially desirable response behavior. Typically, questions in family surveys cover a broad range of topics and require detailed information about respondents’ personal lives, family members, and partnership histories. This typically results in a complex questionnaire structure, often leading to lengthy interviews and repetitive questions (Emery et al., 2023). This may trigger satisficing, as such questions may be cognitively exhausting. For example, respondents may begin to streamline their answers in grid questions, repeatedly choose the first or last answer category, respond “I don’t know”, or refuse to answer the question. This

may be especially true for respondents that have experienced numerous events in their family biographies, leading to a tendency to conceal information in order to minimize cognitive effort. As a result, numbers of (step)-children or ex-partners may be underreported, leading to distortions in demographic key performance indicators (Kreyenfeld et al., 2013; Ruckdeschel et al., 2016).

At the same time, questions regarding unconventional, unstable, or unhappy family and partnership formations may also be sensitive for respondents, as these delve into their personal and intimate sphere and may even trigger negative associations. Such questions might focus on separation and divorce experiences, partnership conflicts and partnership quality, infertility and fertility treatments, sexual behaviors, and attitudes towards traditional gender roles. As sensitive questions are perceived to be intrusive and personal (Tourangeau & Yan, 2007), respondents often do not want to disclose such information, fearing consequences for their reputation and social recognition. Accordingly, most sensitive questions are affected by social desirability bias, as they are often subject to societal expectations and norms (Krumpal, 2013). As a result, sensitive questions can lead to misreporting or item nonresponse, as respondents may choose not to (accurately) answer questions deemed too intrusive.

The presence or absence of an interviewer has a significant impact for these topics: Evidence exists that rates of self-disclosure are higher and socially desirable answers less frequent for questions concerning sexual behavior (Burkill et al., 2016; Kelly et al., 2013), health and well-being (Brackman et al., 2020; Christensen et al., 2013), and gender ideologies (Liu, 2017; Liu & Wang, 2016) in an anonymous interview setting as compared to personal interviews. However, social desirability effects may not only exist for sensitive questions, but also for normatively loaded behaviors. For instance, the normative idea of a “proper” family is widespread in Germany, favoring a heterosexual married couple living together with two children that share strong emotional attachments between family members (Lück & Ruckdeschel, 2018; Sobotka

& Beaujouan, 2014). Questions about family and partnership dynamics may therefore also provoke misreporting on behaviors, emotions, and life events among respondents that are not in line with a societal or individual concept or expectation.

1.2 Practical relevance

As with many large-scale surveys in the social sciences that have had to adapt to methodological shifts and advancements in recent years, family surveys have also faced distinct changes in survey design and fieldwork procedures. In this investigation, I focus on three well-established family surveys conducted in Germany: The German Family Panel pairfam, the German Family Demography Panel Study FReDA, and the Generations and Gender Survey (GGS) (see Table 1.1 for an overview).

Table 1.1: Overview: Design and outreach of three family surveys in Germany (status 2024)

	The German family panel pairfam	The German Family Demography Panel Study FReDA	The Generations and Gender Survey (GGG) Round I and II
Thematic focus	Partnership, parenting, child development, social networks	Family formation, fertility, partnership dynamics, employment	Fertility, partnership, intergenerational relations, gender roles
Starting year	2008	2021	Round I: 2004 Round II: 2021
Target population	Residents of Germany from birth cohort 1971-1973, 1981-1983, 1991-1993, 2001-2003	Residents of Germany aged 18-49	Round I: Residents of Germany aged 18-79 Round II: Residents of Germany aged 18-49
Initial sample size (n)	12,402	37,777	Round I: 10,017 Round II: 22,048
Survey mode	Wave 1-12: CAPI ¹ Wave 13: CATI ⁴ Wave 14: CAWI ²	CAWI ² /PAPI ³	Round I: CAPI ¹ Round II: CAWI ² / PAPI ³
Longitudinal design	Yes, annual interviews	Yes, bi-annual interviews	Yes, interviews every three years
Multi-actor design	Yes, including partners, parents, and children	Yes, including partners	No
Cross-national design	No	No	Round I: Yes, comparable across 19 countries Round II: Yes, comparable across 16 countries
Users ⁵	~2,300	~320	~5,800
Publications ⁶	~500	~20	~670

Note: n=observations; ¹Computer-assisted personal interview; ²Computer-assited web interview; ³Paper and pencil interview; ⁴Computer assisted telephone interview; ⁵Approximate number of registered users – available only for the entire GGS, without limitation to users of the German GGS; ⁶Approximate number of publications retrieved from each study's bibliography (FReDA, 2024; Generations and Gender Programme, 2024; Pairfam, 2024) – available only for the entire GGS, without limitation to users of the German GGS.

These three surveys incorporate design elements and questionnaire contents necessary to study processes and dynamics within intimate relationships. They all employ a longitudinal design, some utilizing a multi-actor approach, and all address topics that are sensitive, personal, and likely to be influenced by social desirability. Consequently, these surveys face specific

challenges when changes become necessary in survey design and data collection procedures that may result from process discontinuities, technological advancements, or the necessity for more efficient data collection. Given the shared thematic focus on families and intimate relationships across these three surveys, they allow me to determine whether methodological challenges introduce nonresponse or measurement error related to important aspects of family and partnership life that could compromise the accurate representation and measurement of target groups and content in a survey. Additionally, they provide an excellent opportunity to investigate methodological challenges in the context of different survey design characteristics, offering practical implications for collecting panel, multi-actor, and sensitive data. To achieve this, this dissertation poses three specific research questions, each targeting a unique methodological challenge related to design characteristics of family surveys necessary to study family dynamics and processes: longitudinal design, multi-actor approach and the measurement of sensitive and personal information (see Chapter 2, 3 and 4, respectively):

Chapter 2 examines the challenge of when the institution responsible for data collection changes in a running panel survey, making a renewal of panel consent mandatory. Economic reasons as well as technological advances may necessitate a change in the institution that owns survey and contact data. Following the conclusion of pairfam funding in 2022, the pairfam sample was integrated into the FReDA project to ensure the continued availability of pairfam panel data. As respondents were required to consent to the transfer of their address and survey data to the new data owner, the question arises as to whether this consent request was related to selection effects that could impact sample composition and compromise comparability between panel waves. Accordingly, this study addresses nonresponse error by assessing whether respondents in an ongoing panel survey are willing to participate in future waves when an institutional change is announced.

Chapter 3 addresses the complexity of collecting multi-actor data, exploring the challenges involved in the recruitment and participation process of respondents' partners, so-called secondary respondents, in a newly recruited self-administered multi-actor survey. Accelerated by the COVID-19 pandemic, the family demographic infrastructure FReDA was pushed to modify its survey design in 2021 from an intended face-to-face format to a fully self-administered format. This shift also had practical implications for the multi-actor design: The recruitment of secondary respondent had to be organized without the assistance of an interviewer, making the cooperation of both primary and secondary respondents in the participation process even more important. As dyadic data is intended to explain processes within relationships, it is crucial to investigate whether nonresponse of secondary respondents correlates with relationship characteristics and induces selectivity in the dyadic sample. Thus, this study focuses on nonresponse error and investigates the participation likelihood of secondary respondents under self-administered recruitment conditions.

Chapter 4 investigates the challenges of measuring sensitive and personal questions, experimenting with interview situations to improve data accuracy and enhance respondent comfort when addressing private or stigmatized topics. Like many other large-scale social science surveys, the Generations and Gender Programme (GGP) transitioned from personal face-to-face interviews to web interviews and mixed-mode survey designs in its second round of data collection (GGS II) beginning in 2020. This decision aimed to combat declining response rates and increasing costs associated with face-to-face interviews in many European countries. The shift from interviewer-administered surveys to self-administered modes of data collection is crucial for measuring sensitive and personal questions. The more anonymous nature of web surveys may reduce socially desirable responding. Therefore, Chapter 4 addresses measurement error and investigates whether responses to sensitive and subjective questions are affected by the transition from face-to-face to web interviews.

Table 1.2: Overview of chapters

	Survey used:		
	pairfam	FReDA	GGs
Design characteristic addressed:			
Longitudinal design	Chapter 2		
Multi-actor design		Chapter 3	
Sensitive questions			Chapter 4
Error addressed:			
Nonresponse error	Chapter 2	Chapter 3	
Measurement error			Chapter 4
Challenge addressed:			
Consent to future participation due to institutional change	Chapter 2		
Participation of secondary respondents in a self-administered design		Chapter 3	
Measurement differences between survey modes			Chapter 4

Each research topic addressed in the following three chapters provides valuable insights for survey methodological and empirical family research. First, as survey methodology must continually adapt to evolving data collection methods, evaluating consent, participation and response behavior under new conditions is essential for addressing emerging methodological challenges, preserving data quality, and ensuring the validity of findings. In this respect, this thesis highlights that the context of a survey plays a fundamental role in evaluating data quality, as potential biases are shaped by both the survey topic and the individuals' personal circumstances. Moreover, this thesis provides practical implications for collecting panel, multi-actor, and sensitive data, emphasizing how survey processes can be influenced by changes in design and procedures.

Second, survey data from projects such as pairfam, FReDA, and the GGS have served as the empirical basis for many publications and research projects advancing family and population research over the last twenty years (see Table 1.1 for scientific outreach). Considering that the nature of these surveys allows for the analysis of changes over time, relationships between

individuals, and - in the case of the GGS - for international differences, there is a risk of confounding effects related to changes in survey design and processes with effects, for example over time or across countries. An accurate representation and measurement of families and partnerships is key to understanding changes in family dynamics and processes within and across populations over time. This thesis emphasizes the importance for family researchers to consider methodological changes in survey data and to evaluate how these changes may impact data quality and the comparability of datasets used to study family life.

1.3 Overview of chapters

1.3.1 Is consent to further panel participation selective? The case of a self-administered family panel survey announcing institutional change¹

In this chapter, I investigate whether a request for consent to continue participating in a panel study leads to higher selectivity within the remaining panel sample. Retaining respondents in the panel is crucial, in particular those who are generally more difficult to reach such as specific sociodemographic groups or individuals who are less likely to participate in family surveys due to experiencing stigmatizing events in their family lives. At the same time, these groups are of great interest to family researchers due to their unique life histories and deviations from societal norms. By requesting consent, panel members are given an opportunity to withdraw from the panel, which bears the risk that certain groups of individuals, for example those who have recently had negative family and/or partnerships experiences, are more likely than others to decline consent at the end of the interview and leave the panel. Therefore, the primary objective

¹ This study was conducted in cooperation with Dr. Claudia Schmiedeberg, and has been accepted for publication in *Journal of Survey Statistics and Methodology* published by Oxford University Press, DOI: 10.1093/jssam/smaf010

of this study is to examine selection effects when an institutional change is announced and consent for data forwarding and future participation is requested.

The analyses are based on data from wave 14 of the pairfam panel, the last wave executed under the initial project team. At the end of the interview, respondents were informed about the institutional change and asked to provide consent to transfer their contact data to the new host data infrastructure FReDA in order to be invited to participate in future waves. Based on this information, we examined which respondents were less likely to provide consent based on socio-demographic characteristics, topic-related characteristics, and panel experience.

Results show that migrants, respondents with lower educational levels, and older cohorts were less likely to consent to future participation under a new host institution. In terms of topic-related characteristics, we only found evidence that respondents who are in a partnership were more likely to consent. Neither parenthood status, recent separation, nor relationship quality (e.g., relationship satisfaction, conflict frequency) affected the decision to provide consent. Participants who had been part of the panel for a longer period of time had higher consent rates compared to those who joined later as part of the refreshment sample in wave 11.

This study demonstrates that a significant change within organizational structures during a long-running family panel can accelerate the increase in selectivity in the remaining sample. Socio-demographic groups who may already face greater difficulties in responding are more likely to discontinue their involvement. However, most aspects related to family life and intimate relationships typically studied in family research remain unaffected. Although topic-related findings are encouraging for pairfam data users who analyze changes in partnership or family life over time, results also indicate that dropout among non-consenting participants is not random, which increases the selectivity of specific sociodemographic groups in the remaining pairfam sample.

1.3.2 Understanding the nonresponse process of secondary respondents: Evidence from a self-administered multi-actor survey²

Chapter 3 investigates the recruitment and participation process of respondents' partners in the family demography panel study FReDA, a self-administered multi-actor panel survey. In its recruitment wave, FReDA approached primary respondents' partners to participate as secondary respondents in order to explore partnership dynamics and processes. Accordingly, the content of the partner survey is centered around their relationship with the primary respondent, addressing emotions and experiences within the relationship. However, the collection of multi-actor data is challenging, especially without the support of an interviewer, as it involves several steps and the cooperation of both respondents: Before the secondary respondent can decide to participate, the primary respondent first needs to identify the target person, then consent to contacting and interviewing the partner, as well as provide valid contact information. In addition to personal predispositions, relationship characteristics can also affect the cooperation of both respondents. Thus, this study investigates whether dropout of secondary respondents is associated with selection effects related to the characteristics of each respondent and/or characteristics of the relationship.

In order to trace the steps that lead to the highest dropout rate of secondary respondents, I matched survey data from primary and secondary respondents of wave 1 of the FReDA panel with contact data from secondary sample members. Having identified the main sources of respondent drop-out, I examined whether individual characteristics of each respondent and dyadic characteristics of their relationship predict nonresponse at these stages. Additionally, I examined how aspects related to contactability affect nonresponse, as establishing contact in a self-administered multi-actor survey becomes more challenging.

² This study is planned for submission to a journal.

My findings indicate that obtaining consent from the primary respondent and participation of the secondary respondent are critical steps in successfully completing a secondary respondent interview that significantly contribute to the nonresponse rate. Further, results suggest that relationship characteristics associated with higher commitment and closeness, such as marriage, co-residence, and a high level of relationship satisfaction, decrease the probability of nonresponse throughout the process. Individual aspects of each respondent such as sociodemographic and personal characteristics associated with a higher individual burden or lower motivation to comply with a survey request lowered the probability of participation, but not necessarily of consent. Moreover, establishing direct contact with the secondary respondents and sending invitations in earlier were both found to increase the chances of participation.

This study emphasizes the challenges of collecting dyadic data in a self-administered multi-actor design. In addition to individual attributes of each respondent such as sociodemographic and personal characteristics, nonresponse among secondary respondents correlates strongly with relationship-related characteristics, inducing selectivity in the dyadic sample. This leads to an overrepresentation of more committed and satisfied couples in multi-actor data. Considering that these factors can influence dynamics and interactions between partners, the ability to draw general conclusions from dyadic analyses based on FReDA multi-actor data is limited.

1.3.3 Better ask online when it concerns intimate relationships? Survey mode differences in the assessment of relationship quality³

Chapter 4 investigates the differences in measurement across survey modes in terms of the assessment of relationship quality in partnerships. This analysis uses data from a German GGP

³ This study was conducted in cooperation with Dr. Detlev Lück and has been published in *Demographic Research* (see Schumann & Lück, 2023), DOI: 10.4054/DemRes.2023.48.22

experimental pilot study conducted in 2018. Relationship quality is an important construct in research on families and intimate lives and is frequently analyzed as an outcome variable as well as a predictor of demographic events such as union dissolution or family formation (e.g., Karney & Bradbury, 1995; Rijken & Liefbroer, 2009). However, asking individuals about their current relationship and feelings towards their partner can feel intrusive, especially for respondents who have had negative experiences. The presence of an interviewer in a face-to-face interview can further strain the interview situation due to the perceived social pressure, leading to more socially desirable and less open answers compared to a self-administered interview with a higher level of anonymity. The aim of this study is to investigate differences in response behavior for questions on relationship quality in personal interviews conducted in face-to-face mode and self-administered web interviews, to test whether such items are prone to mode measurement effects.

For this purpose, an experimental pilot study was carried out by the GGP and conducted in Germany. Sample members were randomly assigned to an experimental group who received an invitation to participate in a web interview or a control group that was interviewed face-to-face in their homes, as was the case in the previous round of data collection (GGS I). After accounting for sociodemographic selection effects, we examined measurement differences between the two survey modes and found pronounced mode differences in measurements for a variety of items assessing intimate relationships. Our findings show that respondents who participated in web mode assessed their overall relationship quality with the partner significantly lower than respondents in the control group: They reported lower levels of satisfaction with different domains of the relationship such as household chores and childcare, a higher conflict frequency, and more negative conflict styles. This suggests that a higher level of anonymity due to the absence of an interviewer leads to less socially desirable responding and a higher willingness to self-disclose.

This study thus provides evidence that the switch from face-to-face interviews to web interviews is both an opportunity and a challenge for family surveys. On the one hand, less socially desirable responses to sensitive questions such as relationship quality might suggest that respondents answer more honestly online as compared to a personal interview. On the other hand, this implies that these indicators, which are frequently used by data users to research partnerships, are prone to measurement bias. The use of different survey modes, particularly a combination of interviewer- and self-administered interviews, may compromise the comparability of measurements from survey data. This issue is especially concerning when comparing differences across countries and/or panel waves.

References

- AAPOR. (2023). *AAPOR's standard definitions, Version 10*. <https://aapor.org/response-rates/>
- AAPOR. (2024). *Best practices for survey research*. <https://aapor.org/standards-and-ethics/best-practices/>
- Axinn, W. G., Link, C. F., & Groves, R. M. (2011). Responsive survey design, demographic data collection, and models of demographic behavior. *Demography*, 48(3), 1127–1149. <https://doi.org/10.1007/s13524-011-0044-1>
- Bates, N., Dahlhamer, J., & Singer, E. (2008). Privacy concerns, too busy, or just not interested: Using doorstep concerns to predict survey nonresponse. *Journal of Official Statistics*, 24(4), 591–612.
- Bethlehem, J., & Biffignandi, S. (2012). *Handbook of web surveys*. Wiley. <https://doi.org/10.1002/9781118121757>
- Bianchi, A., Biffignandi, S., & Lynn, P. (2017). Web-face-to-face mixed-mode design in a longitudinal survey: effects on participation rates, sample composition, and costs. *Journal of Official Statistics*, 33(2), 385–408. <https://doi.org/10.1515/jos-2017-0019>
- Biemer, P. P., Groves, R. M., Lyberg, L. E., Mathiowetz, N. A., & Sudman, S. (Eds.) (2013). *Measurement errors in surveys*. Wiley.
- Billiet, J., & Matsuo, H. (2012). Non-response and measurement error. In L. Gideon (Ed.), *Handbook of survey methodology for the social sciences* (pp. 149–178). Springer. https://doi.org/10.1007/978-1-4614-3876-2_10
- Boudon, R. (2003). Beyond rational choice theory. *Annual Review of Sociology*, 29(1), 1–21. <https://doi.org/10.1146/annurev.soc.29.010202.100213>
- Braekman, E., Charafeddine, R., Demarest, S., Drieskens, S., Berete, F., Gisle, L., van der Heyden, J., & van Hal, G. (2020). Comparing web-based versus face-to-face and paper-

- and-pencil questionnaire data collected through two Belgian health surveys. *International Journal of Public Health*, 65, 5–16. <https://doi.org/10.1007/s00038-019-01327-9>
- Brown, S. L. (2010). Marriage and child well-being: Research and policy perspectives. *Journal of Marriage and Family*, 72(5), 1059–1077. <https://doi.org/10.1111/j.1741-3737.2010.00750.x>
- Bujard, M., Gummer, T., Hank, K., Neyer, F. J., Pollak, R., Schneider, N. F., Spieß, C. K., Wolf, C., Bauer, I., Börlin, S., Bretsch, D., Brüggemann, K., Christmann, P., Edinger, R., Eigenbrodt, F., Frembs, L. C., Groß, K., Jost, C., Kunz, T., . . . Weih, U. (2023). *FReDA – Das familiendemografische Panel*. <https://doi.org/10.4232/1.14080>
- Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., Conrad, F., Wellings, K., Johnson, A. M., & Erens, B. (2016). Using the web to collect data on sensitive behaviours: a study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles. *PloS One*, 11(2), e0147983. <https://doi.org/10.1371/journal.pone.0147983>
- Burton, J., Lynn, P., & Benzeval, M. (2020). How understanding society: the UK household longitudinal study adapted to the COVID-19 pandemic. *Survey Research Methods*, 14(2), 235–239. <https://doi.org/10.18148/srm/2020.v14i2.7746>
- Christensen, A. I., Ekholm, O., Glümer, C., & Juel, K. (2013). Effect of survey mode on response patterns: comparison of face-to-face and self-administered modes in health surveys. *The European Journal of Public Health*, 24(2), 327–332. <https://doi.org/10.1093/eurpub/ckt067>
- Coleman, J. S. (1986). Social theory, social research, and a theory of action. *American Journal of Sociology*, 91(6), 1309–1335.
- Couper, M. P. (1997). Survey introductions and data quality. *Public Opinion Quarterly*, 61(2), 317–338.

- Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75(5), 889–908. <https://doi.org/10.1093/poq/nfr046>
- Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods*, 7(3), 145–156. <https://doi.org/10.18148/srm/2013.v7i3.5751>
- Couper, M. P. (2017). New developments in survey data collection. *Annual Review of Sociology*, 43(1), 121–145. <https://doi.org/10.1146/annurev-soc-060116-053613>
- Couper, M. P., Antoun, C., & Mavletova, A. (2017). Mobile web surveys. In P. P. Biemer, E. D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, N. C. Tucker, & B. T. West (Eds.), *Total survey error in practice* (pp. 133–154). Wiley.
- Couper, M. P., Singer, E., Conrad, F. G., & Groves, R. M. (2008). Risk of disclosure, perceptions of risk, and concerns about privacy and confidentiality as factors in survey participation. *Journal of Official Statistics*, 24(2), 255–275.
- Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: an updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8(3), 513–539. <https://doi.org/10.1093/jssam/smz008>
- De Heer, W., & De Leeuw, E. D. (2002). Trends in household survey nonresponse: A longitudinal and international comparison. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 41–54). Wiley.
- De Leeuw, E. D. (2005). To mix or not to mix data collection modes in surveys. *Journal of Official Statistics*, 21(2), 233–255.
- DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena: Volume 2* (pp. 257–282). Russell Sage Foundation.

- Emery, T., Cabaco, S., Fadel, L., Lugtig, P., Toepoel, V., Schumann, A., Lück, D., & Bujard, M. (2023). Breakoffs in an hour-long, online survey. *Survey Practice*, 16(1). <https://doi.org/10.29115/SP-2023-0008>
- Esping-Andersen, G., & Billari, F. C. (2015). Re-theorizing family demographics. *Population and Development Review*, 41(1), 1–31. <https://doi.org/10.1111/j.1728-4457.2015.00024.x>
- Esser, H. (1986a). Können Befragte lügen? Zum Konzept des“ wahren Wertes“ im Rahmen der handlungstheoretischen Erklärung von Situationseinflüssen bei der Befragung. *Zuma Arbeitsbericht*, 2, 2–35.
- Esser, H. (1986b). Über die Teilnahme an Befragungen. *Zuma Nachrichten*, 10(18), 38–47.
- Fokkema, T., Liefbroer, A. C., Emery, T., Hiekel, N., & Kveder, A. (2016). Generations and Gender Programme Wave 1 data collection: An overview and assessment of sampling and fieldwork methods, weighting procedures, and cross-sectional representativeness. *Demographic Research*, 34(18), 499–524. <https://doi.org/10.4054/DemRes.2016.34.18>
- FReDA. (2024). *Scientific Papers*. <https://www.freda-panel.de/FReDA/EN/Publications/Scientific-Papers/Scientific-Papers.html>
- Generations and Gender Programme. (2024). *Community/ Bibliography*. <https://www.ggp-i.org/bibliography/>
- Goldscheider, F., Bernhardt, E., & Lappegård, T. (2015). The gender revolution: A framework for understanding changing family and demographic behavior. *Population and Development Review*, 41(2), 207–239. <https://doi.org/10.1111/j.1728-4457.2015.00045.x>
- Groves, R. M., Cialdini, R., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475–495. <https://doi.org/10.1086/269338>
- Groves, R. M. (2006). Nonresponse rates and nonresponse bias in household surveys. *Public Opinion Quarterly*, 70(5), 646–675. <https://doi.org/10.1093/poq/nfl033>

- Groves, R. M., Fowler Jr., F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Wiley.
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189. <https://doi.org/10.1093/poq/nfn011>
- Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation Decisions. *Public Opinion Quarterly*, 68(1), 2–31. <https://doi.org/10.1093/poq/nfh002>
- Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, 64(3), 299–308. <http://www.jstor.org/stable/3078721>
- Gummer, T. (2019). Assessing trends and decomposing change in nonresponse bias: The case of bias in cohort distributions. *Sociological Methods & Research*, 48(1), 92–115. <https://doi.org/10.1177/0049124117701479>
- Gummer, T., Höhne, J. K., Rettig, T., Roßmann, J., & Kummerow, M. (2023). Is there a growing use of mobile devices in web surveys? Evidence from 128 web surveys in Germany. *Quality & Quantity*, 57(6), 5333–5353. <https://doi.org/10.1007/s11135-022-01601-8>
- Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T., Lück, D., & Neyer, F. J. (2020). The impact of Covid-19 on fieldwork efforts and planning in pairfam and FReDA-GGS. *Survey Research Methods*, 14(2), 223–227. <https://doi.org/10.18148/srm/2020.v14i2.7740>
- Hendra, R., & Hill, A. (2019). Rethinking response rates: New evidence of little relationship between survey response rates and nonresponse bias. *Evaluation Review*, 43(5), 307–330. <https://doi.org/10.1177/0193841X18807719>

- Jalovaara, M., Andersson, L., & Miettinen, A. (2022). Parity disparity: Educational differences in Nordic fertility across parities and number of reproductive partners. *Population Studies*, 76(1), 119–136. <https://doi.org/10.1080/00324728.2021.1887506>
- Kalmijn, M. (2023). Are national family surveys biased toward the happy family? A multiactor analysis of selective survey nonresponse. *Sociological Methods & Research*, 52(2), 561–586. <https://doi.org/10.1177/0049124120986208>
- Kalmijn, M., & Liefbroer, A. C. (2011). Nonresponse of secondary respondents in multi-actor surveys: Determinants, consequences, and possible remedies. *Journal of Family Issues*, 32(6), 735–766. <https://doi.org/10.1177/0192513X10390184>
- Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin*, 118(1), 3–34. <https://doi.org/10.1037/0033-2909.118.1.3>
- Kelly, C. A., Soler-Hampejsek, E., Mensch, B. S., & Hewett, P. C. (2013). Social desirability bias in sexual behavior reporting: Evidence from an interview mode experiment in rural Malawi. *International Perspectives on Sexual and Reproductive Health*, 39(1), 14–21. <https://doi.org/10.1363/3901413>
- Keusch, F. (2013). The role of topic interest and topic salience in online panel web surveys. *International Journal of Market Research*, 55(1), 59–80.
- Kreuter, F. (2013). Facing the nonresponse challenge. *The ANNALS of the American academy of political and social science*, 645(1), 23–35. <https://doi.org/10.1177/0002716212456815>
- Kreyenfeld, M., Hornung, A., & Kubisch, K. (2013). The German Generations and Gender Survey: Some critical reflections on the validity of fertility histories. *Comparative Population Studies*, 38(1). <https://doi.org/10.12765/CPoS-2013-02>

- Kroneberg, C., & Kalter, F. (2012). Rational choice theory and empirical research: Methodological and theoretical contributions in Europe. *Annual Review of Sociology*, 38(1), 73–92. <https://doi.org/10.1146/annurev-soc-071811-145441>
- Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology*, 5(3), 213–236. <https://doi.org/10.1002/acp.2350050305>
- Krosnick, J. A., Presser, S., Fealing, K. H., Ruggles, S., & Vannette, D. L. (2015). *The future of survey research: Challenges and opportunities*. Report, The National Science Foundation advisory committee for the social, behavioral and economic sciences subcommittee on advancing SBE survey research, Arlington.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025–2047. <https://doi.org/10.1007/s11135-011-9640-9>
- Liu, M. (2017). Data collection mode differences between national face-to-face and web surveys on gender inequality and discrimination questions. *Women's Studies International Forum*, 60, 11–16. <https://doi.org/10.1016/j.wsif.2016.11.007>
- Liu, M., & Wang, Y. (2016). Comparison of face-to-face and web surveys on the topic of homosexual rights. *Journal of Homosexuality*, 63(6), 838–854. <https://doi.org/10.1080/00918369.2015.1112587>
- Lück, D., & Ruckdeschel, K. (2018). Clear in its core, blurred in the outer contours: Culturally normative conceptions of the family in Germany. *European Societies*, 20(5), 715–742. <https://doi.org/10.1080/14616696.2018.1473624>
- Luiten, A., Hox, J., & De Leeuw, E. D. (2020). Survey nonresponse trends and fieldwork effort in the 21st Century: Results of an international study across countries and surveys. *Journal of Official Statistics*, 36(3), 469–487. <https://doi.org/10.2478/jos-2020-0025>

- Martin, S. P. (2006). Trends in marital dissolution by women's education in the United States. *Demographic Research*, 15, 537–560.
- Meterko, M., Restuccia, J. D., Stolzmann, K., Mohr, D., Brennan, C., Glasgow, J., & Kaboli, P. (2015). Response rates, nonresponse bias, and data quality: Results from a national survey of senior healthcare leaders. *Public Opinion Quarterly*, 79(1), 130–144. <https://doi.org/10.1093/poq/nfu052>
- Müller, B., & Castiglioni, L. (2015). Stable relationships, stable participation? The effects of partnership dissolution and changes in relationship stability on attrition in a relationship and family panel. *Survey Research Methods*, 9(3), 205–219. <https://doi.org/10.18148/srm/2015.v9i3.6207>
- Müller, B., & Schmiedeberg, C. (2021). Do respondents get used to answering sensitive questions? Refusal of items on sexuality and fertility in a panel survey. *Public Opinion Quarterly*, 84(3), 654–674. <https://doi.org/10.1093/poq/nfaa041>
- Myrskylä, M., & Margolis, R. (2014). Happiness: Before and after the kids. *Demography*, 51(5), 1843–1866. <https://doi.org/10.1007/s13524-014-0321-x>
- Nelson-Coffey, S. K. (2018). Married... with children: The science of well-being in marriage and family life. In E. Diener, S. Oishi, & L. Tay (Eds.), *Handbook of wellbeing* (chapter 26). DEF Publishers.
- Olson, K., Smyth, J. D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N. A., McCarthy, J. S., O'Brien, E., Opsomer, J. D., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z. T., Turakhia, C., & Wagner, J. (2020). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR Task Force Report. *Journal of Survey Statistics and Methodology*, 9(3), 381–411. <https://doi.org/10.1093/jssam/smz062>
- Pairfam. (2024). *Bibliography*. <https://www.pairfam.de/en/publications/bibliography/>

- Peytcheva, E., & Groves, R. M. (2009). Using variation in response rates of demographic subgroups as evidence of nonresponse bias in survey estimates. *Journal of Official Statistics*, 25(2), 193–201.
- Rijken, A. J., & Liefbroer, A. C. (2009). The Influence of partner relationship quality on fertility. *European Journal of Population/ Revue Européenne De Démographie*, 25(1), 27–44. <https://doi.org/10.1007/s10680-008-9156-8>
- Ruckdeschel, K., Sauer, L., & Naderi, R. (2016). Reliability of retrospective event histories within the German Generations and Gender Survey: The role of interviewer and survey design factors. *Demographic Research*, 34, 321–358.
- Sakshaug, J. W., Cernat, A., & Raghunathan, T. E. (2019). Do sequential mixed-mode surveys decrease nonresponse bias, measurement error bias, and total bias? An experimental study. *Journal of Survey Statistics and Methodology*, 7(4), 545–571. <https://doi.org/10.1093/jssam/smy024>
- Sakshaug, J. W., Yan, T., & Tourangeau, R. (2010). Nonresponse error, measurement error, and mode of data collection: Tradeoffs in a multi-mode survey of sensitive and non-sensitive items. *Public Opinion Quarterly*, 74(5), 907–933. <https://doi.org/10.1093/poq/nfq057>
- Sastry, N., McGonagle, K., & Fomby, P. (2020). Effects of the COVID-19 crisis on survey fieldwork: Experience and lessons from two major supplements to the US Panel Study of Income Dynamics. *Survey Research Methods*, 14(2), 241–245. <https://doi.org/10.18148/srm/2020.v14i2.7752>
- Schumann, A., & Lück, D. (2023). Better to ask online when it concerns intimate relationships? Survey mode differences in the assessment of relationship quality. *Demographic Research*, 48, 609–640. <https://www.jstor.org/stable/48728217>

- Seltzer, J. A. (2019). Family change and changing family demography. *Demography*, 56(2), 405–426. <https://doi.org/10.1007/s13524-019-00766-6>
- Sharp, L. M., & Franke, J. (1983). Respondent burden: A test of some common assumptions. *Public Opinion Quarterly*, 47(1), 36–53. <https://doi.org/10.1086/268765>
- Sobotka, T., & Beaujouan, É. (2014). Two is best? The persistence of a two-child family ideal in Europe. *Population and Development Review*, 40(3), 391–419. <https://doi.org/10.1111/j.1728-4457.2014.00691.x>
- Stroud, N. J., & Kenski, K. (2007). From agenda setting to refusal setting: Survey nonresponse as a function of media coverage across the 2004 election cycle. *Public Opinion Quarterly*, 71(4), 539–559. <https://doi.org/10.1093/poq/nfm042>
- Tourangeau, R. (2004). Survey research and societal change. *Annual Review of Psychology*, 55(1), 775–801. <https://doi.org/10.1146/annurev.psych.55.090902.142040>
- Tourangeau, R., Conrad, F. G., & Couper, M. P. (2013). *The science of web surveys*. Oxford University Press.
- Tourangeau, R., Groves, R. M., & Redline, C. D. (2010). Sensitive topics and reluctant respondents: Demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly*, 74(3), 413–432. <https://doi.org/10.1093/poq/nfq004>
- Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response*. Cambridge University Press.
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60(2), 275–304. <https://doi.org/10.1086/297751>
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859–883.

- West, B. T., Conrad, F. G., Kreuter, F., & Mittereder, F. (2017). Nonresponse and measurement error variance among interviewers in standardized and conversational interviewing. *Journal of Survey Statistics and Methodology*, 6(3), 335–359. <https://doi.org/10.1093/jssam/smx029>
- Williams, D., & Brick, J. M. (2018). Trends in US face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology*, 6(2), 186–211. <https://doi.org/10.1093/jssam/smx019>
- Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (2021). Conducting general social surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*, 85(2), 623–648. <https://doi.org/10.1093/poq/nfab039>
- Yan, T. (2021). Consequences of asking sensitive questions in surveys. *Annual Review of Statistics and Its Application*, 8(1), 109–127.
- Yan, T., & Williams, D. (2022). Response burden – Review and conceptual framework. *Journal of Official Statistics*, 38(4), 939–961. <https://doi.org/10.2478/jos-2022-0041>
- Zillmann, D., Schmitz, A., Skopek, J., & Blossfeld, H.-P. (2014). Survey topic and unit nonresponse. *Quality & Quantity*, 48(4), 2069–2088. <https://doi.org/10.1007/s11135-013-9880-y>

2 Is Consent to Further Panel Participation Selective? The Case of a Self- Administered Family Panel Survey Announcing Institutional Change⁴

Abstract

Changes in study design or procedures bear the risk of affecting survey participation and sample composition. Changing the institution in charge of data collection during a running panel may constitute a risk factor for increasing selectivity, as respondents' explicit consent to transferring data and contact information to the new institution is required for further participation. Based on data from wave 14 of the German Family Panel pairfam⁵, we investigate which respondent characteristics are associated with providing consent to data transfer and whether the timing of such an institutional change plays a role. We focus on respondents' sociodemographic characteristics, factors related to the survey topic such as respondents' relationship status and relationship quality, and panel experience. Further, we investigate if selection effects are more pronounced among respondents with a short or long panel experience. While we find that sociodemographic characteristics such as migration background, level of education, and age affect the propensity to provide panel consent, topic-related characteristics do not consistently impact the decision to consent. Moreover, our findings reveal that a short panel experience leads not only to less panel consent but also to stronger selection effects. As respondent consent is mandatory in the case of institutional change, panel surveys should be aware that this step might be

⁴ This study was conducted in cooperation with Dr. Claudia Schmiedeberg, and has been accepted for publication in *Journal of Survey Statistics and Methodology* published by Oxford University Press, DOI: 10.1093/jssam/smaf010

⁵ This study used data from the German Family Panel pairfam, coordinated by Prof. Dr. Josef Brüderl, Prof. Dr. Sonja Drobnič, Prof. Dr. Karsten Hank, Prof. Dr. Johannes Huinink, Prof. Dr. Bernhard Nauck, Prof. Dr. Franz J. Neyer, and Prof. Dr. Sabine Walper. The German family panel pairfam was funded as priority program and long-term project by the German Research Foundation (DFG) from 2004 to 2022.

a vulnerable point for respondents who are generally less likely to participate, particularly when panel experience is short.

2.1 Introduction

“If you want to measure change, don’t change the measure” (AAPOR, 2022). It is basic knowledge that panel studies, which aim to measure change over time, should avoid changes in design, instruments, and processes. Nevertheless, this is not always feasible, be it to keep up with technological progress and societal change or due to institutional developments (Lynn & Lugtig, 2017). One such major change is transferring the panel to a new host institution, here referred to as institutional change. According to current European law (EU GDPR, 2016), respondents must be informed about the change and provide their consent for their personal data and contact information to be transferred to the new host institution. This procedure may trigger dropout, be it because respondents use the consent question as an easy exit option or because the legal statements concerning data transfer and data protection cause irritation among respondents and lead to a reflection of their previous participation and interview experience in deciding whether they wish to participate further. As a consequence, sample selectivity may increase if this consent is correlated with specific (observed or unobserved) respondent characteristics (Lynn, 2018).

But while most research on selection effects in surveys examines biases due to unit nonresponse or attrition (Dillman et al., 2002; Groves et al., 2004; Groves & Peytcheva, 2008; Lugtig, 2014), research on selective panel consent is scarce, in particular when it comes to data storage and transfer for further participation in a running panel study. A study by G. Müller (2011) documented selectivity in the objection to data transfer with the change of the fieldwork agency among respondents of the German labor market panel PASS. A more recent example, which refers to additional consent due to ministerial requirements rather than institutional change, also confirms selectivity in panel consent (Homuth et al., 2017).

The present study investigates consent to data storage and transfer required for further panel participation after an institutional change in a long-running panel study. Based on wave 14 data of the German Family Panel pairfam, in which a change in the project team and the organizational structure was announced, this analysis explores which respondent characteristics are related to providing consent to transferring address and survey data to the new project team required for contact in future panel waves. We examine three sets of characteristics that we expect to have an impact on respondents' motivation and commitment to provide panel consent: socio-demographic characteristics, which have been found to affect overall nonresponse and consent in surveys, panel experience, and characteristics related to the survey topic, in this case family and intimate relationships. Therefore, we aim to assess if respondents who experience high-quality relationships and subscribe to a socially desirable concept of family life, such as having children or a partner, are more likely to consent. It might be argued that those who feel stigmatized by the survey questions and find the questionnaire burdensome and unpleasant due to their personal circumstances might be less open to comply with the request (Lynn, 2008, 2018; Schräpler, 2004; Yan & Williams, 2022). In addition, we examine which role the duration of panel experience plays for providing panel consent, and to which degree it moderates the effects of sociodemographic and topic-related characteristics.

Our study contributes to current knowledge on the degree to which a panel study might be affected by selective nonresponse after several waves as non-consenting respondents cannot be re-contacted and as a consequence necessarily drop out in the next wave. In this regard, selective non-consent can be regarded the first step to selective nonresponse. Our findings are of significant importance for data collection in running panel surveys as expected cost savings, new funding sources, or switches in data collection and survey design, such as mode changes, easily lead to an institutional change, for example to a new fieldwork agency or data owner. Our study reveals not only whether such an announcement causes selectivity for socio-

demographic and topic-related aspects that are particularly important for data users, but also provides information on whether the strength of possible biases varies according to duration of panel participation. Thereby, we go beyond selectivity in sociodemographic characteristics as reported by Homuth et al. (2017) and examine if the requirement to obtain consent for future contact and data transfer increases selectivity and, thus, compromises data quality for substantive analysis in longitudinal research.

2.2 Theoretical background and literature

A respondent's decision to provide panel consent when an institutional change is announced may be affected by several factors that come into play. Respondents must process and understand the information, decide whether they want to remain part of the panel study under the new institutional setting, and agree to transfer their data and contact information to the new host institution.

In order to understand a respondent's decision for future participation, we draw on rational choice theory. As for general participation in surveys, a respondent will provide panel consent if the expected utility of future participation is larger than its costs (Esser, 1986), which are both influenced by survey-specific as well as situational factors (Groves et al., 1992; Groves & Couper, 1998). Accordingly, utility is determined by factors such as future incentives, social desirability of survey participation, commitment to the study and to supporting scientific research, and experienced pleasantness of the interview, whereas costs depend on respondent burden, for example regarding survey content and demanding questions, as well as privacy- and data protection concerns (Groves et al., 1992; Lynn, 2008; Uhrig, 2008). For instance, respondents who have experienced the interview situation as unpleasant or the topics as threatening or boring will be likely to use the opportunity of a (renewed) panel consent request to drop out of

the panel, whereas respondents who are strongly committed to the study and feel a social obligation to participate may be more likely to provide panel consent. In the following, we describe our expectations concerning three sets of characteristics: sociodemographic, topic-related, and panel experience.

Sociodemographic characteristics

Participating in a survey, in particular in a self-administered interview mode without the support of an interviewer, may be less burdensome for those with higher education levels as they are better equipped to comprehend the survey questions. Moreover, they may feel more connected to scientific research and feel more of an obligation to support research by complying with the survey request (Groves & Couper, 1998; Lynn, 2008). This assumption is supported by a general underrepresentation of individuals with lower levels of education in surveys, as those with higher levels of education are more likely to participate (Tourangeau, 2017) and less likely to attrite (Radler & Ryff, 2010). The same may be the case for individuals with migration background, who may have difficulties in understanding the survey language. Studies on nonresponse show that ethnic minorities generally have higher nonresponse rates across waves than the native population (Radler & Ryff, 2010; Satherley et al., 2015). An additional factor that might increase subjectively perceived costs are privacy concerns. Transferring personal data to an unknown institution might generate mistrust and concern, especially among migrant groups who are less familiar with regulations and institutions in the host country. The explicit mention of the transfer of personal data to a third party under strict panel consent regulations increases attention (Singer et al., 1993), in particular for those who have existing concerns about data protection and fear an invasion of privacy (G. Müller, 2011; Schräpler, 2004). Research on consenting to administrative data and record linkage also supports that higher concerns of privacy and confidentiality lead to a decrease in consent rates for data linkage (Sakshaug et al., 2012). The same might apply to older respondents participating in web surveys, as they may be

less experienced with online environments and therefore more concerned about data transfer and data security. This is also confirmed by findings from a German household panel in which older respondents were more likely to object data transfer (G. Müller, 2011). The author indicates that older respondents may particularly prefer stability and continuity, so that an institutional change has a higher subjective importance and might cause a break in their commitment. We therefore expect that respondents with lower levels of education, those with a migration background, and older respondents are less likely to provide panel consent.

Topic-related characteristics

Surveys with a thematic focus may discourage individuals from specific groups if they perceive the questions to be of little interest or irrelevant, in particular if socially desirable expectations are implicitly present (Groves et al., 2000; Lynn, 2008). This association is supported by findings on topic-related nonresponse in surveys: For instance, Campbell et al. (2020) find that women from sexual minorities are more likely to drop out of the survey than are heterosexual women in a longitudinal study on women's health in Australia. Studies based on a family survey in the Netherlands report that respondents with happier and closer family relationships are more likely to participate (Kalmijn, 2023; Kalmijn & Liefbroer, 2011). A study based on German pairfam data reveals that respondents who experienced a separation since the prior wave are less likely to participate further (B. Müller & Castiglioni, 2015). These findings indicate that individuals with characteristics that do not correspond to a social ideal touched on in a study are less inclined to participate. Accordingly, respondents' personal situations may affect the perceived pleasantness or burden of an interview, as respondents who feel not sufficiently addressed by the questionnaire may feel less connected to the survey and may have a lower motivation to provide panel consent. In the context of the present study, this might be true for single and/or childless respondents as the focus of the pairfam panel mainly covers partnership and family topics. The questionnaire may be of less interest for these participants and make

them possibly feel that their lives do not fit in to the panel. The same applies to respondents in a conflictual or unsatisfying relationship as well as to those who recently separated from their partner as they may perceive answering questions about their current or past relationship such as pertaining to conflicts, infidelity, and intimate partner violence as threatening, burdensome, and unpleasant. Accordingly, these respondents may use the consent request at the end of the survey as an occasion to drop out of the panel. We therefore expect that respondents without children, those without a partner, those who have separated from their partner since the last interview, and those who experience lower levels of relationship quality are less likely to give panel consent.

Panel experience

Research on panel attrition shows that nonresponse due to attrition declines in the course of a panel, be it due to selectivity (i.e., the least motivated respondents have already dropped out in the first waves) or because respondents' commitment to the study increases over time, or simply because they get used to being surveyed (Laurie et al., 1999; Lugtig, 2014). The same may hold for panel consent, leading to higher consent rates among participants who joined the panel a longer time ago. On the other hand, an opposite effect may play a role: After several years of panel participation, panel fatigue may set in (Lemay, 2009; Lugtig, 2014). Respondents may continue their participation out of habit or commitment, but when it comes to explicit panel consent, they may use the opportunity to end their participation if the burden outweighs commitment or habit. However, as commitment can be expected to be higher among respondents who started their panel participation in the initial wave and participated annually over several years than those who joined the panel later as part of a refreshment sample, we expect that the latter are less likely to provide panel consent.

In addition to a direct effect of panel experience on panel consent, we expect a moderating effect leading to smaller impacts of sociodemographic and topic-related factors among respondents with longer panel experience. This may be the case in particular due to past selective attrition as respondents who feel less connected to the panel (e.g., those who experienced negative relationship events) might have already dropped out (Kalmijn, 2023; B. Müller & Castiglioni, 2015). Moreover, the costs of participation (e.g., burden of reporting a separation) may play a smaller role for long-term participants who may feel more committed to the study (Lugtig, 2014; Laurie et al., 1999; Lemay, 2009). Trust in the survey project and fieldwork team may also be higher and, thus, data protection concerns may be less important after many years of panel participation so that long-term participants will be more likely to provide consent to data storage and transfer. Accordingly, we assume that effects of sociodemographic and topic-related characteristics are larger for respondents with a shorter panel experience than for those with a longer panel experience.

2.3 Data and methods

2.3.1 Data

This analysis is based on data from the German Family Panel pairfam, Release 14.1 (Brüderl, Drobnič, et al., 2023), a panel study with a thematic focus on intimate relationships, fertility, and parent-child relationships (Huinink et al., 2011). The panel started in 2008 with a nationwide randomly-drawn sample of 12,402 individuals from the three birth cohorts 1971-73, 1981-83, and 1991-93. Further samples were added over the course of the panel: The DemoDiff sample includes respondents from the two older cohorts living in Eastern Germany who were first surveyed in 2009, and a refreshment sample drawn in wave 11 includes respondents of the two

younger cohorts as well as an additional youth cohort (born 2001-03) who were first surveyed in 2018.

The panel consent request was included in wave 14, which was the last wave before the switch to the new project team. Data collection for the 14th panel wave was self-administered and used a sequential mixed-mode design. Respondents first received an invitation to participate in a computer-assisted web interview (CAWI), via postal letter including a €5 cash incentive in October 2021, followed by two reminders after two and four weeks if they had not participated by that time. A mail questionnaire (PAPI) was enclosed with the second reminder. Respondents who had not participated by January 18, 2022, were reminded by telephone and asked to participate in PAPI or CAWI (Brüderl, Schmiedeberg, et al., 2023). The average duration of a web interviews was about 20 minutes. Wave 14 was the first wave applying a sequential mixed-mode design, whereas in prior waves data were collected in face-to-face interviews (CAPI), and in some cases in waves 12 and 13 in telephone interviews due to the COVID-19 pandemic (Gummer et al., 2020).⁶ Given the changes in survey design implemented in wave 14, the response rate in wave 14 was only 71%, which is considerably lower than in previous waves. For comparison, the panel stability (i.e., wave-to-wave retention rates) amounted to 87-90% over the previous panel waves (Brüderl, Schmiedeberg, et al., 2023).⁷ However, the respondents included in the present analysis had already participated in wave 14 and thus accepted the new design. A total of 4,444 individuals participated in self-administered mode in wave 14.

⁶ A randomly selected subsample of 957 respondents was interviewed via personal interviews in wave 14 to experimentally analyze effects of the mode change. These respondents are excluded from the present analysis.

⁷ Due to specific disposition codes for sampling and data collection in the German context (see Stadtmüller et al., 2019) we refer to the response rates released in pairfam documentation rather than those defined in the AAPOR standards.

After listwise deletion of missing values for the variables under study, the analytical sample includes 4,321 respondents, of which 3,018 had a partner in order to analyze relationship aspects. Stata 18 was used for the estimations.

Wave 14 was the last panel wave conducted by the original project team, a cooperation of the German universities of Munich, Cologne, Jena, and Bremen. Due to changes in the funding source, the pairfam project became part of the data infrastructure project FReDA – The German family demography panel study, which is run by two German public research institutions (Federal Institute for Population Research (BiB) and GESIS – Leibniz-Institute for the Social Science) in cooperation with two of the previous universities (University of Cologne and University of Jena) (Schneider et al., 2021). Respondents were informed about the institutional change at the end of the survey in wave 14 and asked for their consent that, first, the current project team transferred their address data to the new project team and, second, survey data from previous waves were linked to the data from future waves.

2.3.2 Measurements

The dependent variable is a dichotomous variable indicating respondents' consent to the address transfer to the new project team and linkage of all panel waves. As consent to both transfer and linkage was asked within the same request, consenting only to one of the two elements was not an option. The wording of the panel consent request was as follows:

“Next year, the study will no longer be conducted by Kantar on behalf of the universities, but by GESIS – Leibniz Institute for Social Sciences and the Federal Institute for Population Research (BiB). For this reason, we would like to ask you for your consent:

- to link your previous responses to the survey with the new data via an identifier

- to disclose your address information to the two research institutes mentioned so that they may contact you concerning participation in the survey next year.

All information you provide will be treated as strictly confidential and in full compliance with all data protection policies. Data protection policies of GESIS and the BiB can be found here: www.gesis.org / www.bib.bund.de. Your participation is voluntary but nevertheless very important to the success of the study. It is essential that you participate in the survey next year in order for changes over time to be observed.

Do you consent? [Yes / No]”

Accordingly, the outcome variable is coded 0 “no panel consent” and 1 “panel consent”. Following our theoretical argumentation, the main explaining factors consist of sociodemographic characteristics, topic-related characteristics, and panel experience.

As sociodemographic characteristics we include respondent educational level, differentiated between tertiary education, no tertiary education, and currently in education. Migration background is captured using the generated variable in the pairfam dataset that differentiates between respondents without migration background, first generation, and second-generation migrants according to their and their biological parents’ nationality and country of birth (Brüderl, Edinger, et al., 2023). Respondent age is considered using the four birth cohorts 1971-73, 1981-83, 1991-93, and 2001-03.

The topic-related, here family-related, characteristics cover relationship status, indicating whether the respondent currently has an intimate relationship, and parental status, indicating whether the respondent has children (without differentiating between biological, adoptive, and stepchildren). A further aspect that might impact the decision to consent is whether a respondent has recently broken up with a partner. We therefore include a dichotomous variable indicating a reported separation since the previous wave. To measure the subjective quality of a

respondent's relationship, we rely on two different aspects asked in pairfam: The first indicator is relationship satisfaction captured by the question "All in all, how satisfied are you with your relationship?" with a response scale ranging from 0 ("Very dissatisfied") to 10 ("Very satisfied"). The second indicator gives the frequency of conflicts in the relationship, measured by two items adapted from the Network of Relationships Inventory (NRI) (Furman & Buhrmester, 1985) indicating how often respondents disagree and quarrel and are annoyed and angry with their partner rated on a scale from 1 ("Never") to 5 ("Always"); from these two items, an additive index is calculated.

Regarding panel experience, we distinguish between respondents from the original samples, who were part of the panel since the initial wave in 2008 (DemoDiff: 2009), and those from the refreshment sample who entered the panel study in 2018 and were thus only in their fourth wave when the institutional change was announced. We run models for these three samples separately (see Table A2.1), but as results are similar for the DemoDiff and the pairfam original sample, we decided for a dichotomous variable differentiating between the long-running samples (pairfam and DemoDiff) and the refreshment sample in the final model. Furthermore, we include several control variables that may have confounding effects with our variable of interest (i.e., panel consent) and the treatment variables. As women are generally more likely to participate in surveys (Lynn, 2009), gender might also influence the decision to consent. At the same time, gender is associated with relationship satisfaction (Jackson et al., 2014). Further, we include a control variable indicating whether a respondent had participated in all previous waves or had skipped one or more waves in the past. As analyses have shown that temporary dropouts differ from the other respondents in several aspects (B. Müller & Castiglioni, 2017), they may also differ in terms of their propensity to consent to data transfer. Specifically, their commitment to the survey may be lower than among panel respondents who participated in all waves so that they may be less likely to give panel consent. Moreover, for the models that include only

respondents with a partner to assess the effect of relationship quality, we use a further binary control variable distinguishing between cohabitating and non-cohabiting relationships, as this can impact how a respondent assesses relationship quality (e.g., Rhoades et al., 2008). Note that we do not include mode as a control variable although prior research found that in mixed-mode surveys, offline respondents were less likely to provide panel consent in recruitment waves (Dollmann et al., 2023; Schmid et al., 2023) because self-selected mode is a mediator rather than a confounder. However, the association between our characteristics under study and panel consent are not affected by mode (see Table A2.1). The question wording of all variables can be found in Table A2.2 in the Appendix.

2.3.3 Methodological approach

We run linear probability models (LPM) to test the effect of our treatment variables on the probability of giving panel consent. We choose this approach rather than logistic regression because the interpretation of the results of logistic models is complicated by scaling issues, in particular when comparing estimates from differently specified models or from the same model fitted to different groups (see Breen et al., 2018 for a discussion). To test for the robustness of the results, we additionally perform a logistic regression analysis and calculate average marginal effects for the total sample; results are not affected (see Table A2.3). Moreover, as one weakness of LPM is that models can yield unrealistic predicted values (smaller than 0 or larger than 1), we calculate predictions: values lie within the $[0;1]$ boundaries in 97.5 % of cases and the maximum value outside this range is 1.03.

In a first step, we estimate a baseline model which assesses whether there is any indication of selective consent including the total analytical sample. In a second step, we estimate to which degree panel experience moderates' selectivity in panel consent. We run separate models for

respondents with a longer panel experience (pairfam original and DemoDiff sample) and for respondents with a shorter panel experience (pairfam refreshment sample) and test for significance of differences using a fully interacted model with panel experience.

In addition to our baseline model, we run a second set of models in which only respondents who reported having a partner at the time of the interview are included in order to assess the impact of relationship quality on panel consent. To avoid multicollinearity, we estimate separate models including only one indicator of relationship quality (i.e., relationship satisfaction or conflict frequency) at a time. Again, after assessing consent for the full sample we run separate models for long and short panel experience and test the significance of differences in a fully interacted model.

When investigating panel experience, we include only the two middle birth cohorts (1981-83, 1991-93) while excluding the oldest (1971-73) and youngest cohort (2001-03) as the refreshment sample differs from the pairfam original and DemoDiff samples not only in terms of length of panel participation, but also in its age structure: The oldest birth cohort (1971-73) is represented exclusively in the two older samples and the youngest cohort (2001-03) is represented exclusively in the refreshment sample. To avoid confounding effects of age and panel experience we exclude the oldest and youngest cohorts from the models to keep the subgroups as similar as possible. In the main baseline model and the model including partnered respondents only, in contrast, we keep all cohorts and account for birth cohort to retain a higher number of cases in the analytical sample. Additionally, we test the baseline model when only including the middle cohorts, which yielded similar results as when including all four birth cohorts (see Table A2.1).

2.4 Results

In the total analytical sample, 92.9% of respondents agreed to transfer their data and contact information to the new institutions while 7.1% did not. Even if the majority of respondents consented, it is evident that a part of respondents declared the end of their participation in the panel by not consenting to data transfer, recontact and thus future participation.

Table 2.1 provides a descriptive overview of the degree to which the sample composition differs between respondents who did and did not provide panel consent. This provides a first indication of whether the group of non-consenting respondents is selective with respect to the characteristics under study.

Table 2.1: Summary statistics of variables under study

	Total	Consent	No consent
<i>Full sample</i>			
Panel consent			
Yes	92.9	100.0	0.0
No	7.1	0.0	100.0
Gender			
Male	43.7	43.4	47.2
Female	56.3	56.6	52.8
Birth cohort			
1971-1973	20.3	19.7	28.7
1981-1983	30.1	29.6	36.2
1991-1993	26.2	26.8	17.9
2001-2003	23.4	23.9	17.3
Education			
Enrolled	15.9	16.4	8.8
Non-tertiary	44.4	43.4	57.7
Tertiary	39.7	40.2	33.6
Migration background			
None	82.0	82.8	71.7
First generation	7.1	6.5	15.3
Second generation	10.9	10.7	13.0
Relationship status			
No partner	30.2	30.0	32.3
Has partner	69.8	70.0	67.8

Table 2.1: Summary statistics of variables under study (*continued*)

Parental status			
No children	54.8	55.5	45.6
Children	45.2	44.5	54.4
Experienced separation since last wave			
No	95.2	95.1	96.7
Yes	4.8	4.9	3.3
Temporary dropout previous waves			
No	75.6	76.3	66.8
Yes	24.4	23.7	33.2
Panel experience			
Long	54.7	54.6	56.0
Short	45.3	45.4	44.0
Observations	4,321	4,014	307
In percent	100	92.9	7.1
<i>Respondents with partner</i>			
Cohabitation status			
Not cohabiting	18.9	19.5	10.6
Cohabiting	81.1	80.5	89.4
Relationship satisfaction (0-10)	7.9 (1.9)	7.9 (1.8)	7.7 (2.2)
Conflict frequency (1-5)	2.4 (0.7)	2.4 (0.7)	2.4 (0.7)
Observations	3,018	2,810	208
In percent	100	93.1	6.9

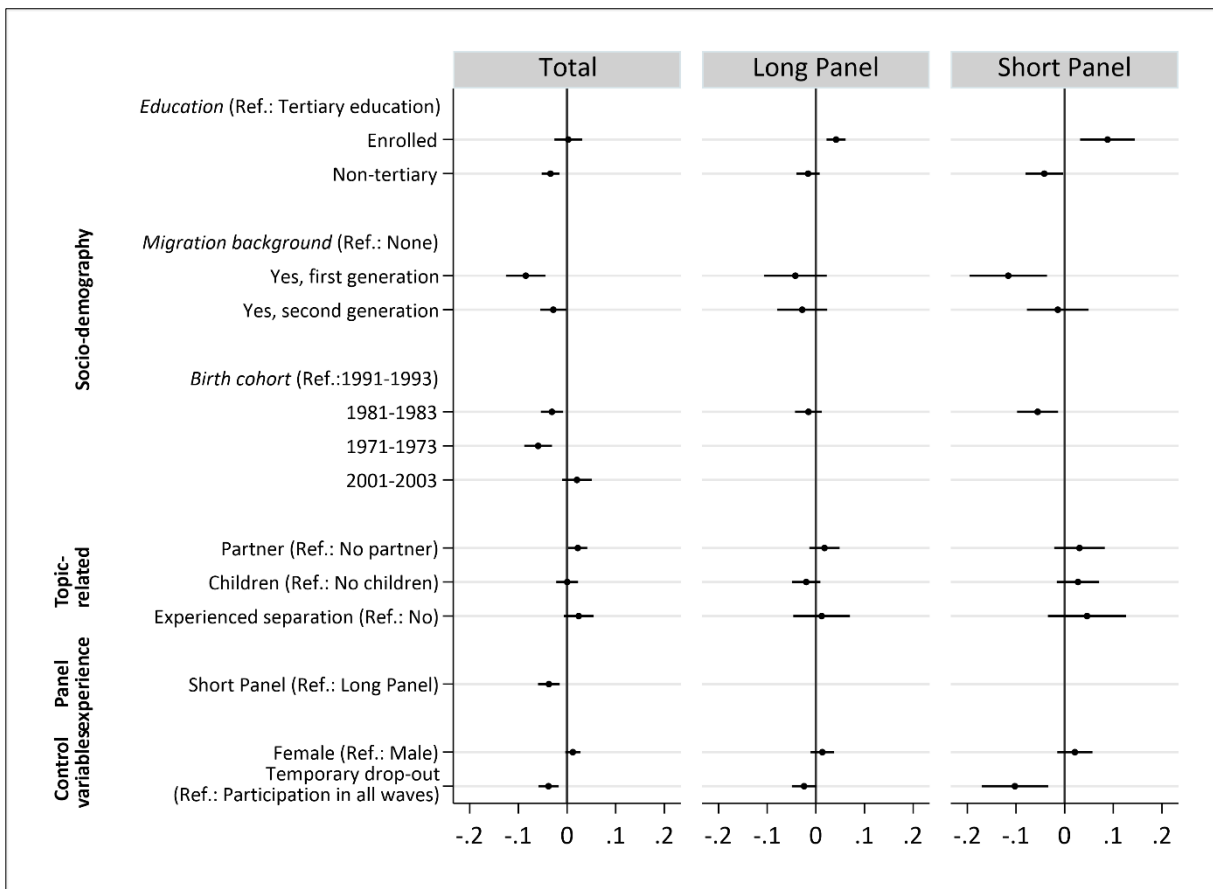
Source: pairfam Release 14.1.

Note: Standard deviation for means in parentheses. Distribution for categorical variables in percent (%) and means for metric variables for the total sample (total) are separated between those who provided panel consent (consent), and those who did not (no consent) for the full sample. Relationship characteristics are only reported for respondents with a partner.

The samples differ in particular with regard to sociodemographic characteristics, whereas the distributions of topic-related characteristics do not differ greatly between respondents who provide panel consent and those who do not. Note that relationship-related characteristics are based on the subsample of partnered respondents; a full description of this subsample can be found in Table A2.4.

Figure 2.1 shows the coefficients of the multivariate models for the full analytical sample including all respondents, irrespective of having a partner. The effects on the probability of providing panel consent are grouped by sociodemographic characteristics, topic-related characteristics, panel experience, and control variables. These are displayed, first, for the total

sample and then separately for respondents who had previously participated for 14 or only 4 waves in the panel.



Source: pairfam Release 14.1.
 Note: Coefficients from linear probability models (LPM) with 95% confidence intervals for the total sample (total), and for respondents who participated longer (long panel) and shorter (short panel) in the panel separately. N_{Total}=4,321; N_{Long}=1,486; N_{Short}=944.

Figure 2.1: LPM regression results of the probability of panel consent

Starting with the total sample, the multivariate findings mostly confirm the expectations from the descriptive sample distributions. Sociodemographic characteristics affect the willingness to provide panel consent: Respondents with non-tertiary education are 3.4 percentage points less likely to give panel consent compared to those with tertiary education. Furthermore, having a migration background, in the first or second generation, significantly reduces the probability of giving panel consent. Moreover, we find a significantly lower probability of providing consent among the older birth cohorts compared to younger cohorts.

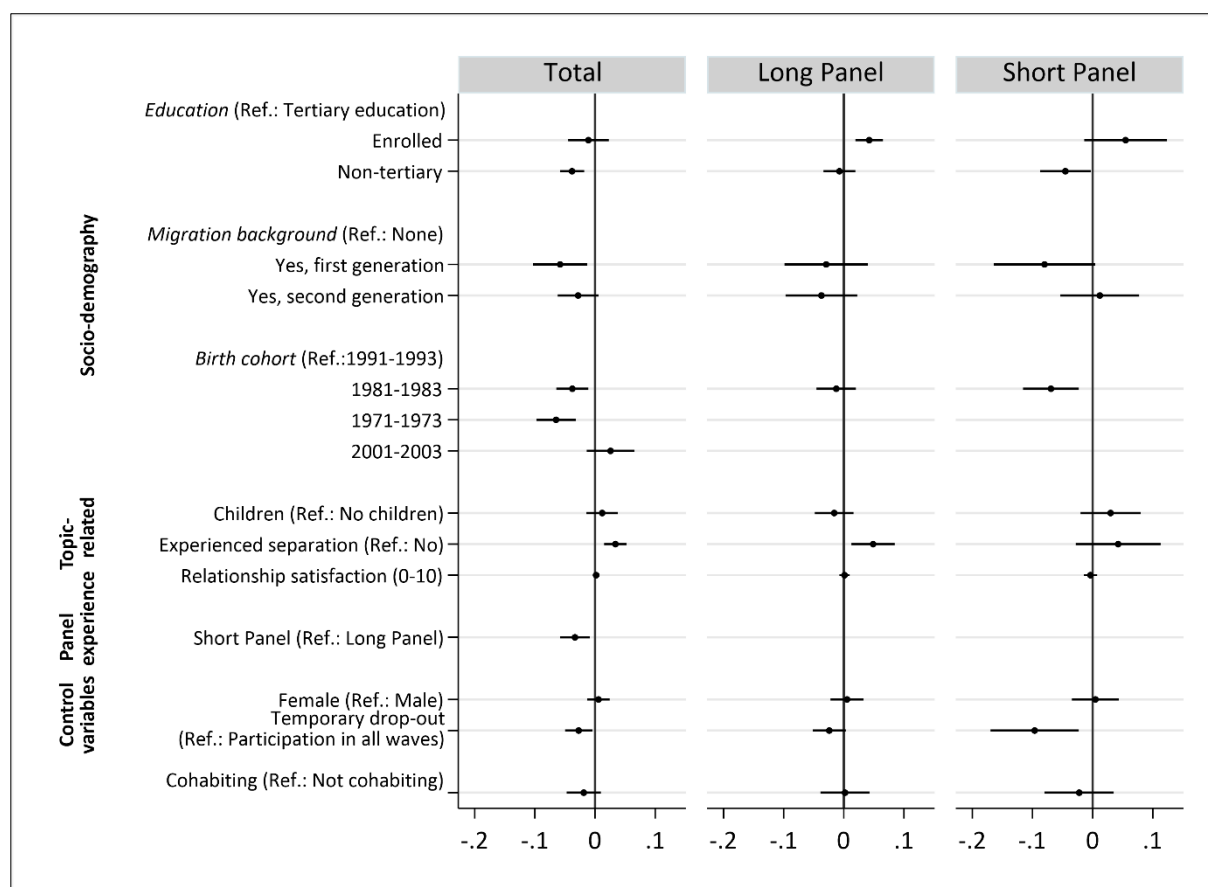
Topic-related characteristics do not consistently affect consent rates, except for having a partner: Respondents with a partner have a 2.2 percentage points higher probability of giving panel consent compared to single respondents. Against our expectations, respondents with children (compared to childless respondents) are not more likely to consent to data transfer and further participation. Similarly, respondents who have experienced a separation since the past wave do not significantly differ from the rest of the sample regarding the likelihood to give panel consent.

As expected, panel experience is found to play a role. Respondents from the pairfam refreshment sample with a shorter panel experience have a 3.7 percentage points lower probability to consent than respondents with longer panel experience from the pairfam original sample.

To evaluate whether the effects differ by panel experience, we run separate models for long and short previous panel participation in Figure 2.1. As described, these models include only the two middle birth cohorts (1981-1983 and 1991-1993) to maintain comparability of the two subsamples. Most effects of age, education, and migration background found in the total sample have significant effects among respondents with short panel experience (i.e., pairfam refreshment sample), while not for respondents with long panel experience (i.e., pairfam original sample), but these differences between are not significant, as shown in the interacted models. In addition, in both samples students are more likely to provide panel consent. All topic-related characteristics, including partnership status, show no effect in both subsamples. The full regression models for the overall sample, the two subsamples, and the interaction model can be found in Table A2.5 in the Appendix.

In a further step, we examine if respondents with higher relationship quality are more likely to give panel consent. As can be seen in Figure 2.2, relationship satisfaction is not associated with

the probability of providing panel consent, neither for the total sample nor for separate samples of respondents who participated longer or shorter in the panel.



Source: pairfam Release 14.1.

Note: Coefficients from linear probability models (LPM) with 95% confidence intervals for the total sample (total), and for respondents who participated longer (long panel) and shorter (short panel) in the panel separately. $N_{Total}=3,018$; $N_{Long}=1,168$; $N_{Short}=757$.

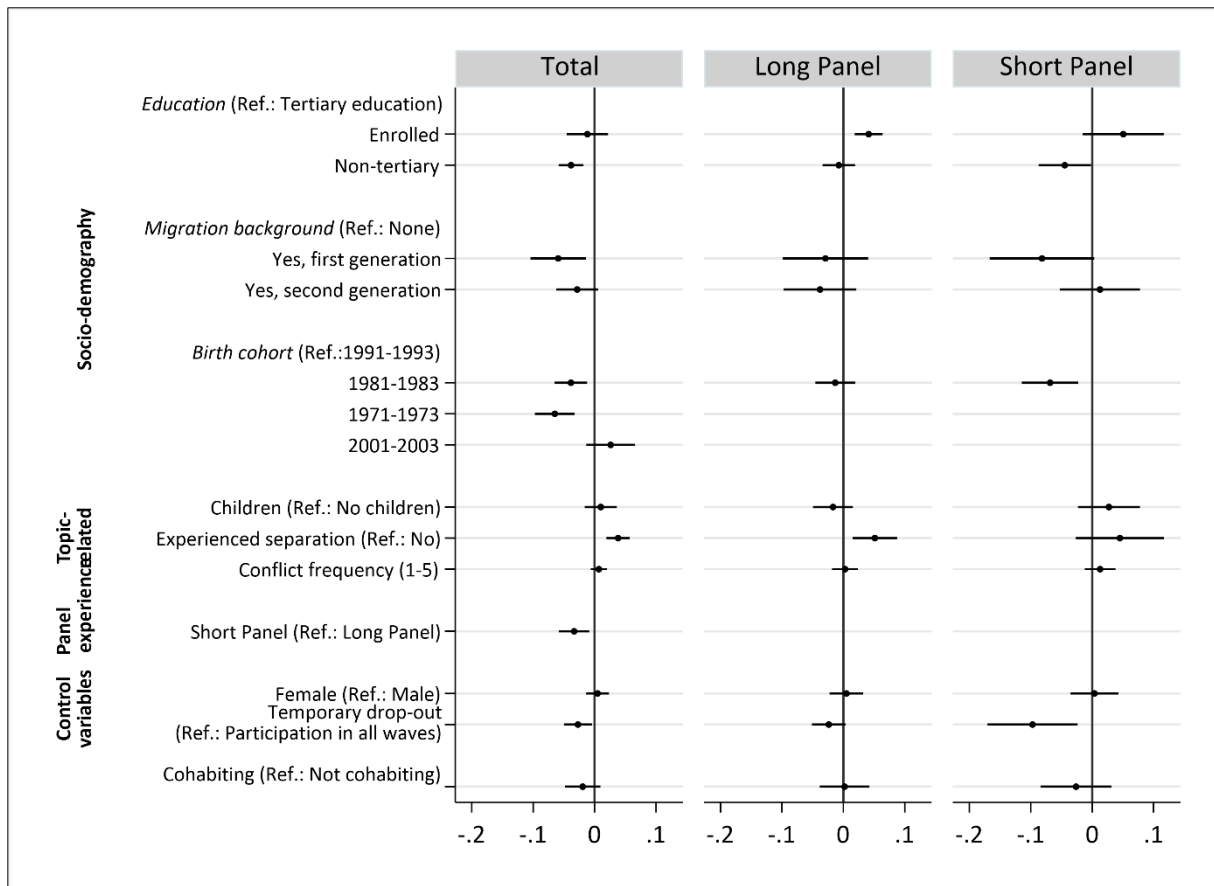
Figure 2.2: LPM regression results of panel consent on relationship satisfaction (subsample: respondents with a partner)

Compared to the full sample, experiencing a separation but already being with a new partner is associated positively with panel consent. A tentative explanation might be that respondents who were willing to participate in wave 14 despite the disruptions in their lives are particularly motivated and committed to the panel while the less committed have already dropped out after such events. Moreover, entering into a new partnership can alleviate the burden of answering questions about an ex-partner. However, the case numbers, especially for the subsample with a

short panel experience, are very low and must therefore be interpreted with caution (see Table A2.4).

The other topic-related and sociodemographic characteristics as well as panel experience reveal similar tendencies for respondents with a partner as for the full sample we have seen in Figure 2.1: Again, respondents with lower education levels, with a migration background, at higher ages, and those with shorter panel experience had a significantly lower probability of giving panel consent, whereas parenthood status showed no effect on panel consent. Looking at the long and short panel participants separately, we again found that the majority of these effects are only evident for respondents who have participated for a shorter period in the panel. As before, differences between short and long panel subsamples tested with interaction effects are not significant (see Table A2.6).

As relationship satisfaction covers only one aspect of relationship quality, we also tested whether respondents with higher conflict frequency in their relationship have a lower probability of giving panel consent in a separate model (Figure 2.3). As relationship satisfaction, conflict frequency does not affect the respondents' willingness to provide panel consent. This applies for the total sample as well as for subsamples including only long and short panel participants, respectively (for full regression models, see Table A2.7).



Source: pairfam Release 14.1.

Note: Coefficients from linear probability models (LPM) with 95% confidence intervals for the total sample (total), and for respondents who participated longer (long panel) and shorter (short panel) in the panel separately. $N_{Total}=3,018$; $N_{Long}=1,168$; $N_{Short}=757$.

Figure 2.3: LPM regression results of panel consent on conflict frequency (subsample: respondents with a partner)

2.5 Discussion

Changes in a running panel survey are sometimes unavoidable. In the present study, we evaluated the consequences of such a change in a long-running panel with focus on family relationships. When the institutional affiliation of the German Family Panel pairfam was changed in 2022, respondents had to be informed and asked for their consent to transfer personal data to the new host institution. It comes as no surprise that some of the respondents opted out of the panel instead of consenting to the transfer of their data and contact information necessary for future participation.

The purpose of the present study was to test whether renewing panel consent is selective: Do sociodemographic characteristics, characteristics related to the survey content, and panel experience play a role in the respondents' decision to provide consent to data transfer and future contact?

Consistent with prior research in other fields regarding general nonresponse in surveys (Radler & Ryff, 2010), we find providing panel consent to be related to sociodemographic characteristics. Older respondents, respondents with a migration background, and those with lower education levels have a significantly lower probability to give panel consent in our sample. This is particularly problematic as individuals with these sociodemographic characteristics are generally underrepresented in surveys and are therefore considered “hard to survey” populations (Tourangeau, 2014). Similar findings are reported by B. Müller and Castiglioni (2015) who also find higher attrition among these subgroups in the pairfam panel. Hence, the same theoretical mechanisms may hold for the decision to give panel consent and the decision to participate in a survey: Individuals who have lower cognitive and technical skills, who are less connected to science and research, and who are more critical of surveys due to higher privacy concerns might be less likely to (further) participate in a survey. For this reason, particular efforts should be made to maintain these subgroups in the panel at critical points such as institutional change.

Due to the focus on topics concerning families and intimate relationships in the pairfam panel, we were especially interested in investigating if such an interruptive announcement leads to a selection of higher-quality relationships and more “conventional” family types. This would be challenging for the panel as a sample bias might affect estimation results, such as misestimations of relationship quality effects. At the same time, it seems plausible that the pairfam interview might be experienced as more pleasant and enjoyable for respondents with higher-quality relationships. We find mixed evidence for this topic-related hypothesis: Having a partner increases the probability of giving panel consent, but all other characteristics we tested do not.

This pattern is difficult to explain and might be caused by mechanisms weighing each other out. On the one hand, the survey topic may be less appealing for respondents without a partner or children; on the other hand, the interview is markedly shorter for those respondents due to routing and filter questions. The first mechanism turns out to be more important in the case of relationship status, and vice versa in the case of parenthood status. Social desirability and burden due to unpleasant questions, in contrast, do not seem to play a major role as neither relationship quality nor separation experiences is linked to the willingness of providing consent. Only a positive association with experiencing a separation and repartnering arises, which is based on a small sample, though. All in all, this is good news for pairfam data users, as nonresponse correlated with subjective or unobservable indicators is more challenging to manage, as these distortions cannot be addressed by weighting based on official statistics (e.g., Lynn, 1996).

Another aspect we focused on is whether panel experience, or more concretely, the time a respondent has already been in the panel, has an impact on panel consent or moderates the effect of our characteristics under study. We find that respondents with short panel experience are less likely to provide panel consent compared to respondents who have participated longer in the panel. It is unclear, however, whether this is a causal effect or due to selection. After many waves, “experienced” panel respondents may have developed commitment and trust in the study, making them more willing to continue participation even after a new host institution takes over. On the other hand, respondents with low commitment and trust may have already dropped out in previous waves. Both explanations may also hold for the moderation effects found: Most effects are only significant for respondents from the refreshment sample with short panel experience, while effects are smaller and not significant in the long panel sample. Although differences between coefficients in the short and long panel are not significant (in the interacted models), the pattern is throughout consistent. Therefore, we conclude that

respondents who joined the panel later (e.g., in the context of a sample refreshment) not only have a higher probability to refuse panel consent, but also show a higher selectivity in consent compared to respondents who have been part of the panel for a longer time.

It should be noted that these selection effects occurred in a long-standing panel, which already experienced greater selection compared to a recruitment survey due to several years of attrition and design changes, such as the mode switch. This might contribute to the explanation why topic-related aspects play a minor role for providing consent as those with poor relationships might already have dropped out earlier. Moreover, the question arises whether the design changes in wave 14 affected the decision to provide panel consent. On the one hand, the consent question concerned only the new institution and the sample under study had already taken part in wave 14, obviously accepting the design change. On the other hand, the consent decision may mirror the respondents' experience with the new survey design; for instance, after having tested the self-administered mode, respondents with lower levels of education may have decided not to continue participation.

These findings have relevant implications for the use as well as for the collection of panel data. As the value of panel data is cumulative, efforts are important to keep respondents in the panel in order to retain a sufficient sample size and keep sample selectivity low (Lynn, 2018; Lynn & Lugtig, 2017). Our findings contribute to a better understanding of how the requirement of obtaining consent for transferring data and contact information in a running panel survey (e.g., in the case of an institutional change, a change of the fieldwork agency, or project consortium) can affect a panel survey, even after several waves. Based on our results, special consideration should be given to groups that are generally more affected by unit nonresponse such as migrants or individuals with low levels of education. These groups are even more at risk when they have been participating in the panel only for a short time. Thus, respondents should at best not be confronted with a “shock” (Lemay, 2009) – here, the announcement of an institutional change

– too early in their panel career as it might cause higher attrition rates, especially for at-risk groups. Although our study reveals no consistent findings of increased selectivity in characteristics related to the survey topic, we find evidence that the respondents’ partnership situation might impact the perceived costs and benefits of the decision to consent for future participation in a survey about families and partnerships such as pairfam. Basically, efforts should be made to address general as well as survey-specific nonresponse throughout the whole survey process, for instance in employing targeted survey designs with special incentives for specific groups, or improving the look and feel of a questionnaire to lower the risk of induced selectivity at interruptive events, such as an institutional change (Lynn, 2017; Singer & Ye, 2012; Vicente & Reis, 2010).

Our study is not without limitations. First, we cannot identify the causal mechanisms leading to lower panel consent rates among hard-to-survey populations and, in particular, do not know whether longer panel experience causally affects the propensity to provide consent (e.g., via increased trust and commitment). To understand this, direct measures for the intervening factors or an experimental setting are necessary.

Second, as the panel consent request included both data transfer to the new institutions and further contact, we have no insights into which aspect actually deterred respondents from providing consent. It might be that respondents simply used the opportunity to exit the panel as it made this option salient, even if they were not bothered by the data transfer to the new institution. To disentangle both aspects, future research may apply an experimental design asking part of the sample only for further contact and the rest for contact and data transfer.

Finally, we focus on panel consent rather than attrition after the institutional change. Respondents who did not consent inevitably dropped out of the pairfam panel, but it is unclear whether those who provided consent to being contacted for the next wave actually participated, as data

for this wave are not yet available. Thus, the analysis of sample selectivity in the first wave at the new project institution is beyond of the scope of this paper, but will be valuable follow-up work for future research.

References

- AAPOR. (2022). *AAPOR standards best practices*. https://aapor.org/wp-content/uploads/2022/11/AAPOR-Standards-best-practices_March-2022.pdf
- Breen, R., Karlson, K. B., & Holm, A. (2018). Interpreting and understanding logits, probits, and other nonlinear probability models. *Annual Review of Sociology*, 44(1), 39–54. <https://doi.org/10.1146/annurev-soc-073117-041429>
- Brüderl, J., Drobníč, S., Hank, K., Neyer, F. J., Walper, S., Wolf, C., Alt, P., Bauer, I., Böhm, S., Borschel, E., Bozoyan, C., Christmann, P., Edinger, R., Eigenbrodt, F., Garrett, M., Geissler, S., Avilés, T. G., Gröpler, N., Gummer, T., . . . Wetzels, M. (2023). *The German family panel (pairfam)* (ZA5678 Data file Version 14.1.0.). GESIS Data Archive, Cologne. <https://doi.org/10.4232/pairfam.5678.14.1.0>
- Brüderl, J., Edinger, R., Eigenbrodt, F., Garrett, M., Hajek, K., Herzig, M., Lorenz, R., Schütze, P., Schumann, N., & Timmermann, K. (2023). *pairfam data manual, Release 14.1.* (ZA5678 Data File Version 14.1.0). LMU Munich, Technical report. GESIS Data Archive, Cologne. <https://doi.org/10.4232/pairfam.5678.14.1.0>
- Brüderl, J., Schmiedeberg, C., Castiglioni, L., Arránz Becker, O., Buhr, P., Fuß, D., Ludwig, V., Schröder, J., & Schumann, N. (2023). *The German family panel: Study design and cumulated field report (Waves 1 to 14), Release 14.1.* pairfam Technical Paper No. 01. <https://doi.org/10.5282/ubm/epub.104749>
- Campbell, A., Perales, F., & Baxter, J. (2020). Sexual minority women in longitudinal survey research: Is attrition a problem? *Archives of Sexual Behavior*, 49, 1443–1461. <https://doi.org/10.1007/s10508-020-01669-z>
- Dillman, D. A., Eltinge, J.L., Groves, R.M. & Little, R. J. A. (2002). Survey nonresponse in design, data collection, and analysis. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.). *Survey nonresponse* (pp. 3-26). Wiley.

- Dollmann, J., Mayer, S. J., Lietz, A., Siegel, M., & Köhler, J. (2023). DeZIM. Panel–data for Germany’s post-migrant society. *Jahrbücher für Nationalökonomie und Statistik*, 243(1), 93–108. <https://doi.org/10.1515/jbnst-2022-0025>
- Esser, H. (1986). Über die Teilnahme an Befragungen. *Zuma Nachrichten*, 10(18), 38–47.
- EU GDPR. (2016). *Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing directive 95/46/EC (General Data Protection Regulation)*. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679&qid=1685616322175>
- Furman, W., & Buhrmester, D. (1985). Children’s perceptions of the personal relationships in their social networks. *Developmental Psychology*, 6(21), 1016–1024.
- Groves, R. M., Cialdini, R., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475–495. <https://doi.org/10.1086/269338>
- Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. Wiley.
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Wiley.
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189. <https://doi.org/10.1093/poq/nfn011>
- Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, 64(3), 299–308. <http://www.jstor.org/stable/3078721>
- Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T., Lück, D., & Neyer, F. J. (2020). The impact of Covid-19 on fieldwork efforts and planning in

- pairfam and FReDA-GGS. *Survey Research Methods*, 14(2), 223–227.
<https://doi.org/10.18148/srm/2020.v14i2.7740>
- Homuth, C., Schmitt, M., Lorenz, C., & Mann, D. (2017). Warum ein erneutes Genehmigungsverfahren im laufenden Längsschnitt weitreichende Folgen für die Datenqualität hat. *Journal for Educational Research Online*, 9(1), 7–31.
<https://doi.org/10.25656/01:12964>
- Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (PAIRFAM): Conceptual framework and design. *Zeitschrift für Familienforschung*, 23, 77–101.
<https://doi.org/10.20377/jfr-235>
- Jackson, J. B., Miller, R. B., Oka, M., & Henry, R. G. (2014). Gender differences in marital satisfaction: A meta-analysis. *Journal of Marriage and Family*, 76(1), 105–129.
<https://doi.org/10.1111/jomf.12077>
- Kalmijn, M. (2023). Are national family surveys biased toward the happy family? A multiactor analysis of selective survey nonresponse. *Sociological Methods & Research*, 52(2), 561–586. <https://doi.org/10.1177/0049124120986208>
- Kalmijn, M., & Liefbroer, A. C. (2011). Nonresponse of secondary respondents in multi-actor surveys: Determinants, consequences, and possible remedies. *Journal of Family Issues*, 32(6), 735–766. <https://doi.org/10.1177/0192513X10390184>
- Laurie, H., Smith, R., & Scott, L. (1999). Strategies for reducing nonresponse in a longitudinal panel survey. *Journal of Official Statistics*, 15(2), 269–282.
- Lemay, M. (2009). *Understanding the mechanisms of panel attrition error* [Doctoral dissertation, University of Maryland]. College Park ProQuest Dissertations & Theses.
- Lutig, P. (2014). Panel attrition. *Sociological Methods & Research*, 43(4), 699–723.
<https://doi.org/10.1177/0049124113520305>

- Lynn, P. (1996). Weighting for non-response. In R. Banks (Ed.), *Survey and statistical computing 1996* (pp. 205–214). Chesham: Association for Survey Computing.
- Lynn, P. (2008). The problem of nonresponse. In E. D. de Leeuw, J. Hox & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 35–55). Taylor & Francis.
- Lynn, P. (2009). Methods for longitudinal surveys. In P. Lynn (Ed.), *Methodology of longitudinal surveys* (pp. 1–20). Wiley.
- Lynn, P. (2017). From standardized to targeted survey procedures for tackling non-response and attrition. *Survey Research Methods*, 11(1), 93–103.
<https://doi.org/10.18148/srm/2017.v11i1.6734>
- Lynn, P. (2018). Tackling panel attrition. In D. L. Vannette & J. A. Krosnick (Eds.), *The Palgrave handbook of survey research* (pp. 143–153). Springer.
- Lynn, P., & Lugtig, P. (2017). Total survey error for longitudinal surveys. In P. P. Biemer, E. D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, N. C. Tucker, & B. T. West (Eds.), *Total survey error in practice* (pp. 279–298). Wiley.
- Müller, B., & Castiglioni, L. (2015). Stable relationships, stable participation? The effects of partnership dissolution and changes in relationship stability on attrition in a relationship and family panel. *Survey Research Methods*, 9(3), 205–219.
<https://doi.org/10.18148/srm/2015.v9i3.6207>
- Müller, B., & Castiglioni, L. (2017). Do temporary dropouts improve the composition of panel data? An analysis of “gap interviews” in the German family panel pairfam. *Sociological Methods & Research*, 49(1), 193–215. <https://doi.org/10.1177/-0049124117729710>
- Müller, G. (2011). Change of the fieldwork agency in the IAB household panel study PASS: A practical report. *methods, data, analysis*, 5(2), 207–226.
<https://doi.org/10.12758/mda.2011.011>

- Radler, B. T., & Ryff, C. D. (2010). Who participates? Accounting for longitudinal retention in the MIDUS national study of health and well-being. *Journal of Aging and Health*, 22(3), 307–331. <https://doi.org/10.1177/0898264309358617>
- Rhoades, G. K., Stanley, S. M., & Markman, H. J. (2008). Couples' reasons for cohabitation: Associations with individual well-being and relationship quality. *Journal of Family Issues*, 30(2), 233–258. <https://doi.org/10.1177/0192513X08324388>
- Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., & Weir, D. R. (2012). Linking survey data and administrative records: Mechanisms of consent. *Sociological Methods & Research*, 41(4), 535–569. <https://doi.org/10.1177/0049124112460381>
- Satherley, N., Milojev, P., Greaves, L. M., Huang, Y., Osborne, D., Bulbulia, J., & Sibley, C. G. (2015). Demographic and psychological predictors of panel attrition: Evidence from the New Zealand attitudes and values study. *PloS One*, 10(3), e0121950. <https://doi.org/10.1371/journal.pone.0121950>
- Schmid, L., Kunz, T., & Naumann, E. (2023). Questionnaire design in the FReDA panel recruitment: Challenges in transitioning from a face-to-face to a self-administered mixed-mode design. *Survey Methods: Insights from the Field*, retrieved from <https://surveyinsights.org/?p=17948>
- Schneider, N. F., Bujard, M., Wolf, C., Gummer, T., Hank, K., & Neyer, F. J. (2021). Family research and demographic analysis (FReDA): Evolution, framework, objectives, and design of “The German Family Demography Panel Study”. *Comparative Population Studies*, 46. <https://doi.org/10.12765/CPoS-2021-06>
- Schräpler, J.-P. (2004). Respondent behavior in panel studies: A case study for income nonresponse by means of the German Socio-Economic Panel (SOEP). *Sociological Methods & Research*, 33(1), 118–156. <https://doi.org/10.1177/0049124103262689>

- Singer, E., Mathiowetz, N. A., & Couper, M. P. (1993). The impact of privacy and confidentiality concerns on survey participation. The case of the 1990 U.S. Census. *Public Opinion Quarterly*, 57(4), 465–482. <https://doi.org/10.1086/269391>
- Singer, E., & Ye, C. (2012). The use and effects of incentives in surveys. *The ANNALS of the American Academy of Political and Social Science*, 645(1), 112–141. <https://doi.org/10.1177/0002716212458082>
- Stadtmüller, S., Silber, H., Daikeler, J., Martin, S., Sand, M., Schmich, P., & Zabal, A. (2019). *Adaptation of the AAPOR final disposition codes for the German survey context*. GESIS–Survey Guidelines, GESIS–Leibniz-Institute for the Social Sciences, Mannheim. https://doi.org/10.15465/gesis-sg_en_026
- Tourangeau, R. (2014). *Hard-to-survey populations*. Cambridge University Press.
- Tourangeau, R. (2017). Presidential address: Paradoxes of nonresponse. *Public Opinion Quarterly*, 81(3), 803–814. <https://doi.org/10.1093/poq/nfx031>
- Uhrig, S. N. (2008). *The nature and causes of attrition in the British Household Panel Survey*. ISER Working Paper Series No. 2008-05, University of Essex, Institute for Social and Economic Research (ISER), Colchester, retrieved from <https://hdl.handle.net/10419/92025>
- Vicente, P., & Reis, E. (2010). Using questionnaire design to fight nonresponse bias in web surveys. *Social Science Computer Review*, 28(2), 251–267. <https://doi.org/10.1177/0894439309340751>
- Yan, T., & Williams, D. (2022). Response burden – Review and conceptual framework. *Journal of Official Statistics*, 38(4), 939–961. <https://doi.org/10.2478/jos-2022-0041>

Appendix

Table A2.1: LPM regression results of panel consent for models including only middle birth cohorts (model 1), DemoDiff sample as separate category (model 2); including survey mode (model 3)

	Model 1	Model 2	Model 3
	Coef. (SE)	Coef. (SE)	Coef. (SE)
Education (<i>Ref.: Tertiary education</i>)			
Enrolled	0.055*** (0.011)	-0.000 (0.014)	-0.015 (0.014)
Non-tertiary education	-0.027* (0.011)	-0.036*** (0.009)	-0.027** (0.009)
Migration background (<i>Ref.: None</i>)			
Yes, first generation	-0.085** (0.026)	-0.081*** (0.021)	-0.082*** (0.021)
Yes, second generation	-0.024 (0.020)	-0.027 (0.014)	-0.030* (0.013)
Birth cohort (<i>Ref.: 1991-1993</i>)			
1981-1983	-0.029* (0.012)	-0.040** (0.012)	-0.033** (0.012)
1971-1973	-	-0.069*** (0.015)	-0.056*** (0.014)
2001-2003	-	0.015 (0.016)	0.015 (0.016)
Partner (<i>Ref.: No partner</i>)			
Has partner	0.020 (0.014)	0.024* (0.010)	0.019* (0.010)
Children (<i>Ref.: No children</i>)			
Has children	0.001 (0.013)	-0.008 (0.011)	0.003 (0.011)
Experienced separation since last wave (<i>Ref.: No</i>)			
Yes	0.026 (0.024)	-0.027 (0.015)	0.024 (0.016)
Panel experience (<i>Ref.: Long</i>)			
Short	-0.038** (0.012)	-	-0.036** (0.011)
Sample (<i>Ref.: Pairfam original sample</i>)			
DemoDiff sample	-	0.084*** (0.018)	-
Refreshment sample	-	-0.033** (0.011)	-

Table A2.1: LPM regression results of panel consent for models including only middle birth cohorts (model 1), DemoDiff sample as separate category (model 2); including survey mode (model 3) (*continued*)

Gender (Ref.: Male)			
Female	0.017 (0.010)	0.012 (0.008)	0.015 (0.008)
Temporary dropout (Ref.: Participation in all waves)			
Yes	-0.045** (0.013)	0.062*** (0.013)	-0.033** (0.011)
Mode (Ref.: PAPI)			
CAWI	-	-	0.109*** (0.016)
Observations	2,430	4,321	4,321
R-squared	0.029	0.031	0.045
Adjusted R-squared	0.025	0.028	0.043

Source: pairfam Release 14.1.

Note: Coefficients (Coef.) from linear probability models with robust standard errors (SE) in parentheses. Model 1: The exclusion of the youngest cohort does not lead to different effects, except the effect for the educational level "enrolled". This is probably due to the fact that enrolled persons in the youngest cohort are students at school. In older cohorts, enrolled persons are more likely to be a selective group, such as individuals pursuing a second-chance education or an academic career. *p < 0.05; **p < 0.01; ***p < 0.001.

Table A2.2: Question wording of variables in the questionnaire

Variable	Question text	Answer categories
Panel consent	Next year, the study will no longer be conducted by Kantar on behalf of the universities, but by GESIS - Leibniz Institute for Social Sciences and the Federal Institute for Population Research (BiB). For this reason, we would like to ask for your consent: - To link your previous responses to the survey Relationships and family life in Germany with new data via an identifier - To disclose your address information to the two research institutes mentioned so that they may contact you concerning participation in the survey next year All information you provide will be treated as strictly confidential and in full compliance with all data protection policies. Data protection policies of GESIS and the BiB can be found here: www.gesis.org/www.bib.bund.de. Your participation is voluntary, but nevertheless very important to the success of the study. It is essential that you participate in the survey next year in order for changes over time to be observed. Do you consent?	-Yes -No
Education	Did you complete a school-leaving certificate or some kinds of vocational training since the last interview in [date of last previous interview]? If yes, which?	-No, neither educational certificate nor vocational qualification. -Secondary general school leaving certificate (Hauptschule) -Intermediate school leaving certificate (Realschule, Mittlere Reife) -Entrance qualification for universities of applied sciences (Fachoberschule) -General or subject-specific university entrance qualification (Abitur, EOS)

Table A2.2: Question wording of variables in the questionnaire (*continued*)

		Other school leaving certificate, namely: -Apprenticeship -Vocational school (general, trade, health-care related) -Advanced vocational school (e.g., certified master craftsman, certified technician) -Civil service apprenticeship -University of applied sciences (also Berufsakademie) -University degree -Doctorate -Not mentioned -Mentioned -No answer
Migration background	In what country were you born? Please use the name by which the country is known today.	-Federal Republic of Germany -German Democratic Republic -Turkey -Russian Federation
	In what country was your biological mother born? Please use the name by which the country is known today.	-Poland -Italy -Serbia -Croatia -Greece
	In what country was your biological father born? Please use the name by which the country is known today.	-Romania -Kazakhstan -Bosnia-Herzegovina -In another country, namely: -Don't know -No answer
	Of which country or countries are you a citizen? Please indicate all applicable countries.	
	Of which country or countries is your biological mother a citizen? Please indicate all applicable countries.	
	Of which country or countries is [respondents whose father has passed away your biological father a citizen? Please indicate all applicable countries.	

Table A2.2: Question wording of variables in the questionnaire (*continued*)

Cohort	When were you born? Please tell me the day, month, and year.	-Day: ____ -Month: ____ -Year: ____
Partner	Do you currently have a partner?	-Yes -No -No answer
Children	Do you have children? Please refer to all biological children, regardless of whether they live with you or not. Also adopted children, children of a partner, or foster children.	-No -Yes, one child -Yes, ____ children -No answer
Separation from ex-partner	During the last interview in [interview date of previous wave] you reported having a partner: [name partner]. Are you still together with [name partner]?	-Yes -No -This information is incorrect -No answer
Relationship satisfaction	All in all, how satisfied are you with your relationship?	0 (Very dissatisfied) - 10 (Very satisfied) -Don't know -I don't want to answer that
Conflicts with the partner	How often do the following things happen in your partnership?	-Never -Rarely -Sometimes -Often -Always -Don't know -I don't want to answer that
Gender	Please indicate your gender.	-Male -Female -Other -No answer

Source: pairfam Release 14.1.

Note: Retrieved from Anchor Codebook (CAWI/PAPI), Wave 14 (2021/2022) for time-varying variables and Anchor Codebook, Wave 1 (2008/2009) for time-constant variables.

Table A2.3: Average marginal effects of panel consent

	AME	SE
Education (<i>Ref.: Tertiary education</i>)		
Enrolled	0.010	(0.013)
Non-tertiary education	-0.033***	(0.009)
Migration background (<i>Ref.: None</i>)		
Yes, first generation	-0.080***	(0.020)
Yes, second generation	-0.030*	(0.014)
Birth cohort (<i>Ref.: 1991-1993</i>)		
1981-1983	-0.032**	(0.012)
1971-1973	-0.069***	(0.019)
2001-2003	0.009	(0.010)
Partner (<i>Ref.: No partner</i>)		
Has partner	0.025*	(0.011)
Children (<i>Ref.: No children</i>)		
Has children	-0.002	(0.011)
Experienced separation since last wave (<i>Ref.: No</i>)		
Yes	0.024	(0.015)
Panel experience (<i>Ref.: Long</i>)		
Short	-0.040**	(0.012)
Gender (<i>Ref.: Male</i>)		
Female	0.013	(0.008)
Temporary dropout (<i>Ref.: Participation in all waves</i>)		
Yes	-0.038***	(0.011)
Observations	4,321	
AIC	2134.461	
BIC	2223.658	

Source: pairfam Release 14.1.

Note: Average marginal effects (AME) from logistic regression model with robust standard errors (SE) in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001.

Table A2.4: Summary statistics of variables under study for respondents with a partner

	Total	Consent	No consent
Gender			
Male	40.5	40.3	42.3
Female	59.5	59.7	57.7
Birth cohort			
1971-1973	24.1	23.3	34.1
1981-1983	35.9	35.4	43.3
1991-1993	27.9	28.7	17.3
2001-2003	12.2	12.7	5.3
Education			
Enrolled	8.8	9.2	2.9
Non-tertiary	43.7	42.7	56.7
Tertiary	47.5	48.1	39.4
Migration background			
None	83.3	84.0	74.5
First generation	7.4	6.9	13.9
Second generation	9.3	9.2	11.5
Parental status			
No children	41.4	42.2	29.8
Has children	58.6	57.8	70.2
Experienced separation since last wave			
No	98.9	100.0	98.8
Yes	1.1	0.0	1.2
Temporary dropout previous waves			
No	72.9	73.5	65.9
Yes	27.1	26.5	34.1
Panel experience			
Long	62.8	62.6	64.9
Short	37.2	37.4	35.1
Cohabitation status			
Not cohabiting	18.9	19.5	10.6
Cohabiting	81.1	80.5	89.4
Relationship satisfaction (0-10)	7.9 (1.9)	7.9 (1.8)	7.7 (2.2)
Conflict frequency (1-5)	2.4 (0.7)	2.4 (0.7)	2.4 (0.7)
Observations	3,018	2,810	208
In percent	100	93.1	6.9

Source: pairfam Release 14.1.

Note: Standard deviation for means in parentheses. Distribution for categorical variables in percent (%) and means for metric variables for the total sample (total), and separated between those who provided panel consent (panel consent), and those who did not (no panel consent) for the full sample.

Table A2.5: LPM regression results of panel consent for the full model, models separated by panel experience and an interaction model with panel experience

	Full model	Long panel ex- perience	Short panel ex- perience	Interac- tion model
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Education (Ref.: Tertiary education)				
Enrolled	0.003 (0.014)	0.042*** (0.010)	0.008** (0.029)	0.042*** (0.010)
Non-tertiary education	-0.034*** (0.009)	-0.016 (0.012)	-0.042* (0.019)	-0.016 (0.012)
Migration background (Ref.: None)				
Yes, first generation	-0.085*** (0.021)	-0.042 (0.033)	-0.116** (0.040)	-0.042 (0.033)
Yes, second generation	-0.028* (0.014)	-0.028 (0.026)	-0.014 (0.032)	-0.028 (0.026)
Birth cohort (Ref.: 1991-1993)				
1981-1983	-0.031** (0.012)	-0.015 (0.014)	-0.056* (0.022)	-0.015 (0.014)
1971-1973	-0.059*** (0.015)	-	-	-
2001-2003	0.020 (0.016)	-	-	-
Partner (Ref.: No partner)				
Has partner	0.022* (0.010)	0.018 (0.016)	0.030 (0.026)	0.018 (0.016)
Children (Ref.: No children)				
Has children	0.000 (0.011)	-0.019 (0.015)	0.027 (0.022)	-0.020 (0.015)
Experienced separation since last wave (Ref.: No)				
Yes	0.024 (0.016)	0.012 (0.030)	0.046 (0.041)	0.012 (0.030)
Panel experience (Ref.: Long)				
Short	-0.037** (0.011)	-	-	-0.026 (0.034)
Gender (Ref.: Male)				
Female	0.012 (0.008)	0.013 (0.012)	0.021 (0.018)	0.013 (0.012)
Temporary dropout (Ref.: Participation in all waves)				
Yes	-0.038*** (0.011)	-0.024 (0.013)	-0.102** (0.035)	-0.024 (0.013)

Table A2.5: LPM regression results of panel consent for the full model, models separated by panel experience and an interaction model with panel experience (*continued*)

<i>Interaction effects with Panel experience</i>				
Education				
Enrolled X Short panel	-	-	-	0.046 (0.030)
Non-tertiary education X Short panel	-	-	-	-0.026 (0.021)
Migration background				
Yes, first generation X Short panel	-	-	-	-0.073 (0.052)
Yes, second generation X Short panel	-	-	-	0.014 (0.042)
Birth cohort				
1981-1983 X Short panel	-	-	-	-0.040 (0.026)
Partner				
Has partner X Short panel	-	-	-	0.013 (0.031)
Children				
Has children X Short panel	-	-	-	0.047 (0.027)
Experienced separation since last wave				
Yes X Short panel	-	-	-	0.034 (0.051)
Gender				
Female X Short panel	-	-	-	0.008 (0.022)
Temporary dropout				
No X Short panel	-	-	-	-0.078* (0.037)
Observations	4,321	1,486	944	2,430
R-squared	0.026	0.014	0.055	0.037
Adjusted R-squared	0.023	0.008	0.045	0.029

Source: pairfam Release 14.1.

Note: Coefficients (Coef.) from linear probability models (LPM) with robust standard errors (SE) in parentheses.

*p < 0.05; **p < 0.01; ***p < 0.001.

Table A 2.6: LPM regression results of panel consent on relationship satisfaction (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience

	Full model	Long panel ex- perience	Short panel ex- perience	Interac- tion model
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Education (Ref.: Tertiary education)				
Enrolled	-0.011 (0.017)	0.042*** (0.012)	0.055 (0.035)	0.042*** (0.012)
Non-tertiary education	-0.038*** (0.010)	-0.007 (0.013)	-0.045* (0.022)	-0.007 (0.014)
Migration background (Ref.: None)				
Yes, first generation	-0.058* (0.023)	-0.029 (0.035)	-0.078* (0.043)	-0.029 (0.035)
Yes, second generation	-0.028 (0.017)	-0.037 (0.030)	0.012 (0.033)	-0.037 (0.030)
Birth cohort (Ref.: 1991-1993)				
1981-1983	-0.038** (0.014)	-0.013 (0.017)	-0.069** (0.024)	-0.013 (0.017)
1971-1973	-0.065*** (0.017)	-	-	-
2001-2003	0.026 (0.020)	-	-	-
Children (Ref.: No children)				
Has children	0.012 (0.013)	-0.016 (0.016)	0.030 (0.025)	-0.016 (0.016)
Experienced separation since last wave (Ref.: No)				
Yes	0.034*** (0.009)	0.049** (0.018)	0.042 (0.036)	0.049** (0.018)
Relationship satisfaction (0-10)	0.002 (0.003)	0.001 (0.004)	-0.003 (0.006)	0.001 (0.004)
Panel experience (Ref.: Long)				
Short	-0.033** (0.012)	-	-	0.061 (0.067)
Gender (Ref.: Male)				
Female	0.006 (0.009)	0.005 (0.014)	0.005 (0.020)	0.005 (0.014)
Temporary dropout (Ref.: Participation in all waves)				
Yes	-0.027* (0.012)	-0.024 (0.014)	-0.096* (0.038)	-0.024 (0.014)

Table A2.6: LPM regression results of panel consent on relationship satisfaction (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience (*continued*)

Cohabitation Status (<i>Ref.: Not cohabiting</i>)				
Cohabiting	-0.019 (0.015)	0.002 (0.021)	-0.022 (0.029)	0.002 (0.021)
<i>Interaction effects with Panel experience</i>				
Education				
Enrolled X Short panel	-	-	-	0.013 (0.037)
Non-tertiary education X Short panel	-	-	-	-0.038 (0.025)
Migration background				
Yes, first generation X Short panel	-	-	-	-0.051 (0.056)
Yes, second generation X Short panel	-	-	-	0.049 (0.045)
Birth cohort				
1981-1983 X Short panel	-	-	-	-0.057 (0.028)
Children				
Has children X Short	-	-	-	0.046 (0.030)
Experienced separation since last wave				
Yes X Short panel	-	-	-	-0.006 (0.040)
Relationship satisfaction X Short panel	-	-	-	-0.005 (0.007)
Gender				
Female X Short panel	-	-	-	-0.001 (0.024)
Temporary dropout				
No X Short panel	-	-	-	-0.072 (0.040)
Cohabitation status				
Cohabiting X Short panel	-	-	-	-0.024 (0.007)
Observations	3,018	1,168	757	1,925
R-squared	0.024	0.011	0.049	0.032
Adjusted R-squared	0.020	0.002	0.035	0.020

Source: pairfam Release 14.1.

Note: Coefficients (Coef.) from linear probability models (LPM) with robust standard errors (SE) in parentheses.

*p < 0.05; **p < 0.01; ***p < 0.001.

Table A2.7: LPM regression results of panel consent on conflict frequency (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience

	Full model	Long panel ex- perience	Short panel ex- perience	Interac- tion model
	Coef. (SE)	Coef. (SE)	Coef. (SE)	Coef. (SE)
Education (Ref.: Tertiary education)				
Enrolled	-0.012 (0.017)	0.041*** (0.013)	0.051 (0.034)	0.041*** (0.012)
Non-tertiary education	-0.038*** (0.010)	-0.007 (0.013)	-0.045* (0.022)	-0.007 (0.013)
Migration background (Ref.: None)				
Yes, first generation	-0.059* (0.023)	-0.029 (0.035)	-0.082 (0.043)	-0.029 (0.035)
Yes, second generation	-0.028 (0.017)	-0.038 (0.030)	0.013 (0.033)	-0.038 (0.030)
Birth cohort (Ref.: 1991-1993)				
1981-1983	-0.039** (0.013)	-0.013 (0.017)	-0.068** (0.024)	-0.013 (0.017)
1971-1973	-0.065*** (0.017)	-	-	-
2001-2003	0.026 (0.020)	-	-	-
Children (Ref.: No children)				
Has children	0.010 (0.013)	-0.017 (0.017)	0.027 (0.026)	-0.017 (0.017)
Experienced separation since last wave (Ref.: No)				
Yes	0.038*** (0.009)	0.051** (0.018)	0.045 (0.036)	0.051** (0.018)
Conflict frequency (1-5)				
	0.007 (0.007)	0.003 (0.011)	0.013 (0.013)	0.003 (0.011)
Panel experience (Ref.: Long)				
Short	-0.033** (0.013)	-	-	0.000 (0.051)
Gender (Ref.: Male)				
Female	-0.005 (0.009)	-0.005 (0.014)	0.004 (0.020)	0.005 (0.014)
Temporary dropout (Ref.: Participation in all waves)				
Yes	-0.027* (0.012)	-0.024 (0.014)	-0.097** (0.038)	-0.024 (0.014)

Table A2.7: LPM regression results of panel consent on conflict frequency (subsample: respondents with a partner) for the full model, models separated by panel experience and an interaction model with panel experience (*continued*)

Cohabitation status (<i>Ref.: Not cohabiting</i>)				
Cohabiting	-0.019 (0.015)	0.002 (0.021)	-0.026 (0.029)	0.002 (0.020)
<i>Interaction effects with Panel experience:</i>				
Education				
Enrolled X Short panel	-	-	-	0.009 (0.035)
Non-tertiary education X Short panel	-	-	-	-0.037 (0.026)
Migration background				
Yes, first generation X Short panel	-	-	-	-0.052 (0.056)
Yes, second generation X Short panel	-	-	-	0.051 (0.045)
Birth cohort				
1981-1983 X Short panel	-	-	-	-0.055 (0.029)
Children				
Has children X Short panel	-	-	-	0.044 (0.031)
Experienced separation since last wave				
Yes X Short panel	-	-	-	-0.006 (0.041)
Conflict frequency X Short panel	-	-	-	0.010 (0.017)
Gender				
Female X Short panel	-	-	-	-0.001 (0.024)
Temporary dropout				
No X Short panel	-	-	-	-0.073 (0.040)
Cohabitation status				
Cohabiting X Short panel	-	-	-	-0.028 (0.036)
Observations	3,018	1,168	757	1,925
R-squared	0.024	0.011	0.049	0.032
Adjusted R-squared	0.020	0.002	0.035	0.021

Source: pairfam Release 14.1.

Note: Coefficients (Coef.) from linear probability models (LPM) with robust standard errors (SE) in parentheses.

*p < 0.05; **p < 0.01; ***p < 0.001.

3 Understanding the Nonresponse Process of Secondary Respondents: Evidence From a Self-Administered Multi-Actor Survey⁸

Abstract

Multi-actor surveys offer a valuable opportunity to gain insights from both the primary respondent and a secondary respondent, providing a deeper understanding of dynamics and processes within relationships. However, collecting multi-actor data can be challenging due to the multiple steps and actors involved in conducting a secondary respondent's interview, such as obtaining consent from the primary respondent and inviting the secondary respondent to participate. This study is based on the first wave of the German Family Demography Panel Study, FReDA⁹, which applied a self-administered multi-actor design. The research goal of this study is to identify the steps that lead to most secondary respondent dropouts and to investigate whether individual characteristics of the two actors, dyadic aspects of the relationship between them, and design-specific factors related to their contactability predict nonresponse during this process. Results suggest that dyadic characteristics of the relationship, such as a low level of commitment and closeness, result in lower consent and participation rates. Individual aspects, such as sociodemographic characteristics associated with a higher respondent burden, reveal varying effects within the different steps. Furthermore, the way in which contact is established with the secondary respondent is crucial, as sending invitations in later batches and through primary

⁸ This study is planned for submission to a journal.

⁹ This study used data from the German Family Demography Panel Study FReDA, initiated by a consortium consisting of the Federal Institute for Population Research (BiB), GESIS – Leibniz Institute for the Social Sciences, and the pairfam consortium represented by Prof. Dr Karsten Hank (University of Cologne) and Prof. Dr Franz J. Neyer (Friedrich Schiller University Jena). FReDA is funded by the German Federal Ministry of Education and Research (BMBF) (grant numbers 01UW2001A, 01UW2001B, 01UW2001C).

respondents, rather than contacting each secondary sample member directly, leads to higher nonresponse. Overall, the findings help to address nonresponse among secondary respondents and to provide practical implications for reducing sample selectivity in the collection of multi-actor data.

3.1 Introduction

Increasingly, surveys implement multi-actor designs that cover a wide range of relationships, such as those between physicians and their patients (Schäfer et al., 2011), employees and their employers (Chaplin et al., 2005), and mostly between people who are intimate, like partners or spouses, or parents and children (Dykstra et al., 2005; Huinink et al., 2011; Johnson et al., 2009; Mortelmans et al., 2011; Schneider et al., 2021; Sweet & Bumpass, 1996). The majority of these studies are singular multi-actor surveys, as they sample one person, referred to as the primary respondent, and then recruit another person, the secondary respondent, who is related to the first (Pasteels, 2015). Dyadic or multi-actor data enable the understanding of dynamics and interactions within a relationship by gathering information and perspectives from both actors.

The collection of multi-actor data is challenging. In contrast to general survey participation, involving secondary respondents requires a multi-step process that depends on the cooperation of both primary and secondary respondents. Before the secondary respondent can be contacted to be offered the option to participate, the primary respondent must agree to the interview taking place and the target person being contacted. During this process, expectations and negotiations may arise between the two respondents, since they share a relationship with each other in this design. This means that obtaining multi-actor data involves multiple steps and multiple individuals, difficulties each involving the risk of a loss of potential secondary respondents, which may cause selectivity in dyadic data. Accordingly, nonresponse of secondary respondents can be affected by the particular step in the process, by the individuals involved and by relationship dynamics.

There is a large body of literature focusing on sources and reasons for unit nonresponse in general surveys and its consequences for data collection and usage (Dillman et al., 2002; Groves et al., 1992; Groves & Couper, 1998; Groves & Peytcheva, 2008; Lynn, 2008; Peytchev, 2013).

Next to reasons related to data collection or survey design, such as failure in identification or contact (Fuchs et al., 2013; Lynn & Clarke, 2002), nonresponse results mostly from the sample member's individual situation leading to non-cooperation (Beullens et al., 2018). For instance, the socioeconomic status of a person (Radler & Ryff, 2010; Tolonen et al., 2006; Tourangeau, 2014) as well as personal attributes related to the survey topic, such as a political interest in election studies (Brehm, 2009; Groves et al., 2004; Stroud & Kenski, 2007), impact nonresponse by affecting the individual burden to comply with a survey request (Groves et al., 1992).

However, there is still insufficient research on the process through which nonresponse arises for secondary respondents in a multi-actor survey design. Compared to nonresponse in general surveys, multi-actor surveys face further challenges that make additional factors relevant. First of all, the individual situation of two actors has to be considered, as both respondents have to comply with a survey request during the process. In addition to individual characteristics that impact the respondent burden, dyadic characteristics within the relationship can lead to dynamics that either encourage or discourage respondents from proceeding further into the process. Previous studies on selectivity in multi-actor data on family relations have found evidence that dyads with closer relationships are overrepresented (Hünteler & Wetzel, 2020; Kalmijn & Liefbroer, 2011; Schröder et al., 2012). This suggests that positive relationship dynamics may favor the response process of secondary respondents and should be taken into consideration.

Nevertheless, it is unclear at which step these selection processes occur and which aspects are decisive, especially since interpersonal dynamics are often not visible and indicators are rarely measured in surveys. This is particularly relevant, as previous findings on nonresponse in multi-actor data relied on interviewer-administered multi-actor surveys, which differ in data collection methods due to the presence of an interviewer in the respondent's home. In a self-administered setting, the participation process is different because there is no interviewer present to boost motivation and take advantage of the practical benefits of a household interview (Müller,

2017; Schmiedeberg et al., 2016; Schröder et al., 2016). For instance, consent rates for other survey requests, such as data linkage, were found to be lower in self-administered surveys compared to interviewer-administered surveys (Sakshaug et al., 2017). Similarly, response rates for participating in self-administered surveys, such as web surveys, tend to be lower as well (Daiker et al., 2020; Manfreda et al., 2008). Moreover, a self-administered interview setting may lead to a higher disclosure of respondent and dyad characteristics related to nonresponse because interviewer effects, as discussed in previous research (Müller, 2017; Schmiedeberg et al., 2016), do not exist. Given that most large-scale surveys in social science are transitioning from face-to-face to online or mixed-mode designs, due to cost savings, better coverage, and higher response rates (Couper, 2011, 2017; Gummer et al., 2020; Olson et al., 2020; Wolf et al., 2021), this study makes an important contribution to existing research in investigating selection processes in a self-administered multi-actor survey.

Additionally, unlike previous studies that focused on nonresponse at specific steps during the process, this study is the first one to my current state of knowledge that outlines the entire nonresponse process of secondary respondents. This provides important knowledge that enables data collectors to intervene at critical points in the process, those with the highest risk of nonresponse. Moreover, this study contributes knowledge about survey nonresponse prediction, not only by accounting for the sample member's individual situation, but also by considering aspects of the sample member's relationship. This approach allows to capture interpersonal negotiations and dynamics, which are particularly relevant for participation in multi-actor surveys. Identification of all sources and determinants of nonresponse in multi-actor surveys can achieve thorough understanding of when and why secondary respondents become lost in the process. Addressing biases and missing values in dyadic data *ex post* is very challenging (Pasteels, 2015; Young & Johnson, 2013), hence such knowledge is essential for developing

new approaches to data collection and survey design, the ultimate aim being to increase participation rates among secondary respondents in multi-actor surveys.

Therefore, this study explores the nonresponse process of secondary respondents schematically, using the first wave of the German family demography panel study FReDA – a self-administered multi-actor survey that approaches the partners of respondents to participate as secondary respondents (Hank et al., 2024; Schneider et al., 2021). First, I provide a systematic overview of the participation process of secondary respondents, investigating the sources of nonresponse, in order to identify the steps that result in the most dropouts. Furthermore, I explore which characteristics are associated with nonresponse at a particular step, considering individual characteristics known from the literature and characteristics at the dyadic level, to account for the interdependence within the relationship. As contact becomes more important in a self-administered multi-actor survey design, I also examine the influence of different contact approaches on the nonresponse process. The aim of this study is to investigate whether selection effects are similar or whether they differ throughout the nonresponse process, as each of the two individuals may have different motivations. Therefore, my main two main research questions are:

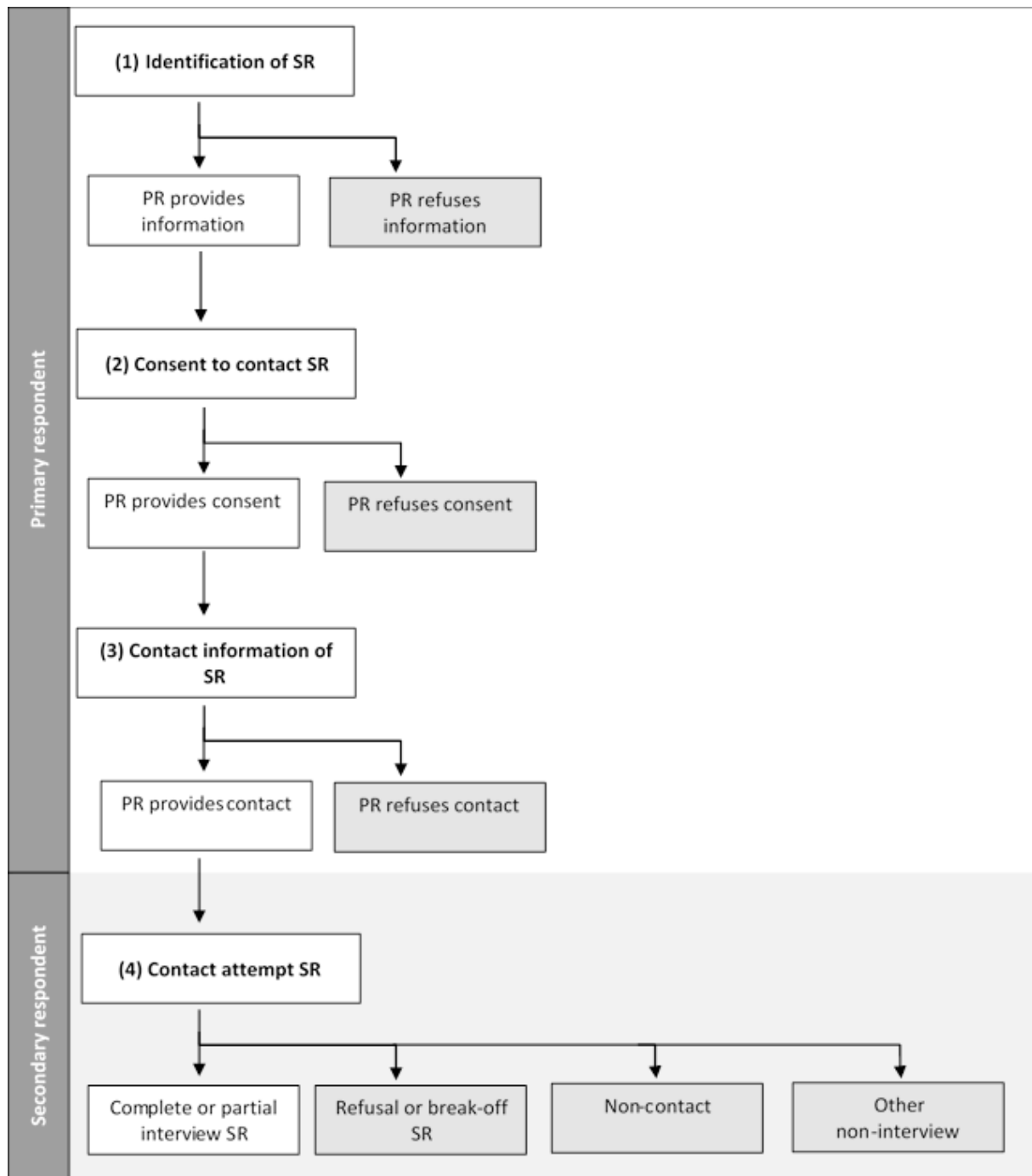
RQ1: Which steps lead to the most dropouts of potential secondary respondents in a self-administered multi-actor survey?

RQ2: Which factors predict nonresponse at a particular step?

3.2 Nonresponse process for secondary respondents

3.2.1 Steps leading to nonresponse

To conduct a secondary respondent interview, several steps must be achieved. Accordingly, Figure 3.1 shows the steps where a loss of potential secondary respondents can happen. The steps outlined in Figure 3.1 are applicable to most multi-actor designs. However, adjustments and adaptations during the process may occur, based on the chosen survey mode, survey design or the prevailing legal conditions. First of all, the target person, i.e., the secondary sample member, has to be identified (1). This information can be obtained by asking a screening question to the primary respondent, e.g. if he or she has a partner, a child or a household member. As using this information is one form of processing data, the data owner, e.g. the survey institute, must obtain consent, via the primary respondent, to interview the target person, for ethical and legal reasons (EU GDPR, 2016) (2). Additionally, the primary respondent must provide contact details for the target person to enable communication between them and the survey institute (3). Once a list of target persons and their contact details is provided, the dyadic sampling frame is specified, and the secondary sample members can be invited to the interview.



Note: SR=secondary respondent; PR=primary respondent. White-colored box means the case remains; gray-colored box means the case drops out.

Figure 3.1: Steps in the nonresponse process for secondary respondents

In the last step, the secondary sample member is contacted and if contact is made successfully, they can decide whether to participate (4). As with nonresponse for general participation in surveys, the required data cannot be collected when the sample member is unwilling to

cooperate and refuses participation, or fails to complete the questionnaire, is unable to cooperate due to non-contact, or when other reasons lead to no interview, such as language barriers or technical problems (AAPOR, 2023; Lynn, 2008).

One can assume that there is no significant dropout or non-disclosure of potential target persons during the first step, as in most cases the primary respondent is not yet aware that this information will prompt an additional request. During the consent step, primary respondents act as gatekeepers, as they have control over whether the secondary respondent receives an invitation to participate or not. As a result, the primary respondent may decline the secondary respondent's participation for various reasons, such as a desire for protection or other concerns. This can lead to a high proportion of non-consenting participants (Carroll-Lind et al., 2006; Slauson-Blevins & Johnson, 2016). Primary respondents who have already given their consent are probably more likely to provide contact details, as they may have anticipated this, resulting in fewer refusals during this step. The refusal of secondary sample members to participate during the last step should also result in a large portion of dropouts. Refusing participation commonly represents the primary source of nonresponse in surveys (Beullens et al., 2018; Williams & Brick, 2018), because engaging in a survey requires great effort (Esser, 1986; Groves & Couper, 1998).

Thus, I expect the loss of potential secondary respondents in this process to take place mainly at two steps:

H1: The refusal to provide consent (step 2) and the refusal to participate (step 4) lead to a higher loss of secondary sample members during the process than the other steps (step 1 and 3).

3.2.2 Predictors of nonresponse

In the following, I discuss factors that might predict nonresponse at the two relevant steps: provision of consent and participation. The predictors apply to individual aspects of each sample member (individual level), dyadic aspects between both sample members (dyadic level), and additionally, for the participation step, I look at design aspects at the contact level.

3.2.2.1 Consent step

The individual level

Providing consent can be seen as one form of complying with a survey request. That means that the cognitive, psychological, or social situation of a primary respondent has an impact on the motivation or burden to further support the survey (Groves et al., 1992). This aligns with findings from self-administered surveys on the obtention of consent for other survey requests, such as future participation. Previous studies have indicated that native respondents, and those with higher levels of education, are more likely to comply with the request for consent to participate in a panel (Sakshaug et al., 2020). Similar results have been observed in interviewer-administered multi-actor surveys: In the Netherlands Kinship Panel Study (NKPS), Kalmijn and Liefbroer (2011) found that parents with lower education levels and those from non-western countries are less likely to give permission to contact their children. Similarly, in the German family panel, pairfam, the participation of partners is lower when the primary respondent has a migration background and lower levels of education (Schröder et al., 2012). For instance, primary respondents with lower levels of education or a migration background might perceive the consent question as being more demanding and burdensome because they may encounter language or comprehension issues. This is especially relevant in a self-administered survey where no interviewer is present to provide additional information, support and guidance at this step.

Additionally, concerns about data privacy may be generally more pronounced in a self-administered web interview.

Women are generally associated with greater openness to support scientific requests (Tourangeau, 2014; Watson & Wooden, 2009), but there has been mixed evidence regarding the association between gender and granting consent for different survey requests (Sakshaug et al., 2012; Sala et al., 2012; Schröder et al., 2012). However, a survey with a thematic focus on fertility and family may be more appealing to women, as traditionally this is considered a female domain (Slauson-Blevins & Johnson, 2016). Similarly, when the interview is perceived as enjoyable and the questions as interesting, such as when parents are asked about their family life and children, there may be a higher motivation to comply with an additional request (Groves et al., 2000).

Although primary respondents should generally be more open to supporting an additional survey request, since they have already agreed to be interviewed in a survey, I expect the individual characteristics of a primary respondent to influence the consent step as follows¹⁰:

H_c2.1: Primary respondents with lower levels of education are less likely to provide consent than primary respondents with higher levels of education.

H_c2.2: Primary respondents with a migration background are less likely to provide consent than primary respondents without such a background.

H_c2.3: Male primary respondents are less likely to provide consent than female primary respondents.

¹⁰ H_c refers to hypotheses related to the consent step, H_p refers to hypotheses related to the participation step.

H_c2.4: Primary respondents who are childless are less likely to provide consent than primary respondents who are parents.

The dyadic level

The decision to consent should also depend on the evaluation of the relationship itself, as primary respondents may not only consider their own concerns, but also anticipate those of the secondary respondent. Previous research on selectivity in multi-actor data has found an association between weak relationship ties and a greater desire for gatekeeping by refusing the permission to interview a secondary respondent: For instance, parent-child dyads are more likely to provide consent for participation when they have a supportive relationship in the pairfam study (Hünteler & Wetzel, 2020), and high levels of contact and relationship quality in the NKPS (Kalmijn & Liefbroer, 2011). However, these findings focus on parent-child dyads, therefore concerns and protective ambitions may be particularly high when it comes to one's own children, especially when they are younger. A study by Müller (2017), based on pairfam data and conducted as personal household interviews, found that co-residence seems to favor the decision to consent for couple dyads, but it is unclear whether this is due to enhanced practical realization because of the interview situation or a higher level of institutionalization within the partnership.

Although previous findings are derived solely from interviewer-administered multi-actor data, they suggest that primary respondents with weak relationships are less inclined to provide consent. This trend should be more pronounced in a self-administered survey, as there is no interviewer present to persuade hesitant respondents. The anonymity of a self-administered interview situation may offer primary respondents in a troubled relationship a viable option to decline consent. This could be because they want to avoid their partner expressing their perspective on the poor relationship, as they may fear additional conflicts or debates that could further

strain their relationship, or they may be concerned about third parties becoming aware of the situation. Additionally, they may already anticipate that the secondary respondent would not be fully committed to comply with such a request, and refuse on their behalf by not consenting. In contrast, a closer relationship between primary and secondary respondents should lead to a stronger commitment and clearer expectations, probably resulting from higher contact frequency and more communication about the request.

Therefore, I expect that dyadic characteristics that express a lower level of closeness and commitment within a relationship will affect the consent stage as follows:

H_c3.1: Primary respondents who are not cohabiting with their partner are less likely to provide consent than those who live together with their partner.

H_c3.2: Primary respondents who are not married to their partner are less likely to provide consent than primary respondents who are married.

H_c3.3: The lower the level of relationship satisfaction, the less likely it is that primary respondents will provide consent.

3.2.2.2 Participation step

The individual level

The decision to participate as a secondary respondent should depend on similar individual characteristics as the decision to participate in general surveys. Individuals with a higher respondent burden, who face difficulties with understanding and language, have lower cognitive skills, or are less interested and open to science – such as those with lower levels of education or those who are non-native speakers – have been found to be less likely to participate (Radler & Ryff, 2010; Tolonen et al., 2006; Tourangeau, 2014). This is particularly the case in self-administered

surveys, suggesting that the digital divide has a greater impact on these harder-to-reach groups (Atkeson et al., 2014). Accordingly, authors have also found this association in interviewer-administered multi-actor surveys and have revealed that poorly-educated persons and those with non-native backgrounds are less likely to participate as secondary respondents (Bastaitis et al., 2015; Kalmijn & Liefbroer, 2011; Schröder et al., 2012).

Studies on general survey participation have provided clear evidence that women are more likely to participate in surveys than men (Korkeila et al., 2001; Watson & Wooden, 2009), especially in surveys that focus on fertility and family (Fokkema et al., 2016; Slauson-Blevins & Johnson, 2016), possibly due to higher levels of pro-sociality and openness to such topics. Since the topic is announced in the invitation letter, or the primary respondent reports about the themes in the questionnaire, secondary sample members are usually aware of what the survey is about. Therefore, asking individuals who are childless or who have had negative experiences in their family life questions about such topics may increase the burden on respondents and reduce their interest in participating in the survey. This is in line with previous research that has found an underrepresentation of divorced respondents (Bastaitis et al., 2015; Kalmijn & Liefbroer, 2011; Mitchell, 2010) and an overrepresentation of individuals with intact family lives (Kalmijn, 2023) in surveys with a focus on families in both multi-actor surveys and surveys of other designs. This suggests that, if a survey topic is less salient and relevant for a person, they do not feel addressed by the survey content and their motivation to participate decreases (Groves et al., 2000).

Therefore, I expect for the individual characteristics of a person to have a similar effect in the participation stage as they do in the consent stage:

H_p2.1: Secondary sample members with lower levels of education are less likely to provide consent than secondary sample members with higher levels of education.

H_p2.2: Secondary sample members with migration backgrounds are less likely to participate than secondary sample members without migration backgrounds.

H_p2.3: Male secondary sample members are less likely to participate than female secondary sample members.

H_p2.4: Secondary sample members who are childless are less likely to participate than secondary sample members who are parents.

The dyadic level

Generally, secondary sample members may be likely to participate, since the primary respondent is one of the issuers of the request. However, those with stronger relationship ties should exhibit a higher motivation and commitment to cooperate. This is supported by findings in interviewer-administered multi-actor surveys on intergenerational relations, which have revealed an overrepresentation of parent-child dyads with closer and more intense family ties in the German family panel, pairfam (Hünteler & Wetzel, 2020; Schröder et al., 2012), the Dutch NKPS (Kalmijn, 2023; Kalmijn & Liefbroer, 2011), and the Divorce in Flanders study in Belgium (Bastaits et al., 2015). Authors who have studied the involvement of partners or spouses as secondary respondents have shown that participation is associated with higher levels of relationship quality and commitment (Barton et al., 2020; Hagedoorn et al., 2015). Others have only identified a relationship between participation and the level of institutionalization of the partnership (Müller, 2017; Schröder et al., 2012).

However, these studies rely on personal interviews and do not account consistently for the steps before participation that might have already caused selectivity. As for consent, a self-administered interview environment may result in a greater disclosure of dyadic selection effects since it is easier to withdraw without others noticing. It can generally be assumed that secondary sample members in closer relationships have a higher willingness, or a perceived obligation, to

support their counterpart with a favor by agreeing to participate, because they already know that the latter consented to the interview. Furthermore, situational closeness may lead to more contact and communication about the survey request, resulting in higher chances of remembering the invitation. Additionally, secondary sample members already know that the survey content relates to the relationship with the primary respondent. Thus, those with positive relationship experiences may feel more comfortable being interviewed about their happy relationships, whereas those experiencing more troubled relationships are more likely to opt out, by not participating at all or by breaking off the questionnaire.

Although those with weak relationships may already dropped out when their partners do not grant consent, I assume that commitment and closeness within the dyad should still play a role during the participation step, and that they should align similarly as for the consent step:

H_p3.1: Secondary sample members who are not cohabiting with their partner are less likely to participate than those who live together with their partner.

H_p3.2: Secondary sample members who are not married to their partner are less likely to participate than secondary sample members who are married.

H_p3.3: The lower the level of relationship satisfaction, the less likely it is that secondary sample members will grant consent.

The contact level

Since successful contact is a requirement for having the opportunity to participate as a secondary respondent, factors related to contactability may also have an impact on nonresponse. In multi-actor surveys, the contact approach is different compared to standardized contact approaches for sample members that belong to a known sampling frame. Therefore, contact information for secondary sample members is obtained mostly through the primary respondent

(Pasteels, 2015). However, the way of establishing contact can vary, depending on the survey design and interview situation in a multi-actor survey. Generally, primary respondents provide contact details to the data collector, be it via an interviewer, a web questionnaire, or a mail questionnaire, who then contacts the secondary sample member. In some cases, the primary respondent can choose to assume the contact approach themselves, opting to contact the secondary sample member personally to deliver the invitation. This can affect the success of participation, because it may be unclear whether the primary respondent in fact does forward the invitation to the secondary respondent, particularly in a self-administered survey, where no interviewer can oversee this process. For example, the primary respondents may change their decision after initially agreeing to interview their partners, or they may simply fail to deliver the invitation.

Furthermore, the timing of the secondary sample member's receipt of the invitation may also affect the chances of participation. Invitations for secondary sample members are often sent in tranches to minimize the time between the primary and secondary respondent interviews. A longer field time suggests that it may have been more difficult to reach and persuade the primary respondent to participate, possibly due to lesser interest in the study. This could also make it equally challenging to convince the secondary sample member. Additionally, most primary respondents complete their interviews in the first few weeks after receiving the study invitation, and the remaining interviews are spread out over several weeks (e.g., Bujard et al., 2023). Therefore, the time between the primary respondent's consent and the receipt of the invitation by secondary sample members is likely to be shorter in the first batch for practical reasons. For example, research has shown that a shorter time span between reminders increases response rates because the study remains more present (Lugtig et al., 2022). A similar mechanism may influence secondary sample members when they receive the invitation earlier.

Therefore, regardless of the individual and dyadic situations of the secondary sample member, nonresponse is likely to be influenced by technical aspects associated with contactability in the following manner:

H_p4.1: Sending the invitation to the primary respondent leads to a lower probability of participation compared to sending the invitation directly to the secondary sample member.

H_p4.2: Sending invitations in later tranches lead to a lower probability of participation compared to sending invitations in earlier tranches.

3.3 Data and methods

3.3.1 Data

Analyses are based on data from the German Family Demography Panel Study FReDA, Release version 3.0.0 (Bujard et al., 2023). FReDA is a survey with a thematic focus on intimate relationships, fertility and family lives (Hank et al., 2024; Schneider et al., 2021). Given the central theme of couple relationships in FReDA, the study design incorporates a dyadic approach in applying a multi-actor-design to interview respondent's partners as secondary respondents. Interviews in FReDA are conducted fully self-administered, respondents can choose to participate in a computer assisted web interview (CAWI) or in a paper-and-pencil interview (PAPI). The first FReDA wave was carried out in 2021, and consists of a randomly drawn sample from municipality registers of 18-49-year-old residents in Germany. The study is designed to conduct interviews of primary respondents twice a year, each with a duration of 20-30 minutes. In the first wave, an additional short recruitment interview of about 10 minutes (W1R) was carried out in advance in order to attract respondents to provide panel consent for future participation. The recruitment survey, conducted from April 7 to June 29, 2021, yielded a response rate 2

(RR2), according to AAPOR standard definitions (AAPOR, 2023), of 34.9 percent, with a net sample size of 37,783. Of these, 22,048 respondents participated in the follow-up subwave, W1A, that was fielded from July 7 to September 22, 2021. In every survey, target persons received up to three reminder letters and in the first invitation letter a €5 unconditional pre-paid incentive was enclosed (Bujard et al., 2023).

In subwave W1A, respondents who reported to have a partner were asked for consent to interview their partners as secondary respondents, and if consent was granted, to provide their partner's contact details¹¹. If primary respondents preferred to handle the contact approach themselves, they could choose to have the invitation letter sent to their own address instead of using the standard method, in which the survey agency contacted the secondary respondent directly, based on the contact address provided by the primary respondent. As did primary respondents, partners invited to participate as secondary respondents received an unconditional €5 prepaid incentive. Additionally, each invitation included a paper-based questionnaire and a QR code for a web survey, enabling secondary respondents to choose whether to participate in a CAWI or a PAPI survey (for the invitation letter, see Figure A3.1). The questionnaire for secondary respondents was designed to be shorter, but complementary to the primary respondent questionnaire, and covered some 70 questions that took about 18 minutes to complete in CAWI mode. The questionnaire focused on sociodemographic information as well as on aspects related to family and partnership life. To minimize the time between the primary respondent's consent and the secondary respondent's invitation, invitation letters were dispatched in three tranches according to the time when the primary respondent's interview had been completed.

¹¹ Apart from those who denied consent, the fieldwork team refrained from inviting 65 partners residing in areas affected by the flood catastrophe in Germany in July 2021.

The field period for all three tranches of the partner wave (W1Apartner) started on August, 11, 2021 and ended on November, 17, 2021¹² (Bujard et al., 2023).

In order to examine the descriptive numbers of dropouts during the nonresponse process, I added contact protocol data from all eligible secondary respondents that were available in an unreleased gross sample dataset of all contacted partners distributed by the survey agency with W1A data from the primary respondents. Contact protocols including information about an invited secondary respondent but without a corresponding primary respondent interview in the scientific use file W1A were excluded (n=81).

The analytical sample for the multivariate analysis includes only cases that participated in both W1R and W1A (n=22,025), as information on the predictors under study were conducted in both waves. According to my research concern, only cases that reported to have a partner in W1A were kept in the analytical sample (n=16,838). Those who reported having no partner or who omitted this question were dropped, as they could not provide information about their partner by design. After listwise deletion of missing values of all variables under study, the overall analytical sample reached n=15,182 observations. These observations form the analytical sample from which to assess the probability of consent being granted. For assessing the probability of secondary sample member participation, the analytical sample consisted of those observations that granted consent and provided contact information (n=8,414) to ensure that the secondary sample member would receive the chance to participate. Summary statistics of the analytical samples of the primary respondents and the secondary sample members for all variables under study can be found in Table A3.1 and A3.2.

¹² Moreover, 100 interviews of secondary respondents were deleted from the dataset due to various reasons including high item nonresponse or implausible entries in W1A.

3.3.2 Measurements

I use two dichotomous outcome variables in the analysis of the process. The first outcome is granting of consent to interview the secondary respondent, coded as 1 “consent” if primary respondents agreed to interview their partners, and coded as 0 “no consent”, when a primary respondent denied consent or refused to answer the question altogether (for the question text see Figure A3.2). The second outcome refers to whether a secondary sample member participated, given a contact granted approach, and is coded as 1 “complete interview”, when a secondary respondent (partially) completed the interview. Accordingly, those who did not respond or failed to complete the questionnaire at the beginning are coded as 0, “no complete interview”.

For the predictor variables, I was restricted to choosing information available on primary respondents, secondary respondents, and secondary nonrespondents in the data, to analyze which aspects led to (non)-participation of secondary respondents. To do so, I used information provided by the primary respondent about their own situation, about the secondary sample member’s situation and about the relationship between the two. To avoid confounding, I used proxy information for all secondary sample members, even if self-reports were available for the ones who participated. Additionally, as most variables covered basic and objective information, deviations between proxy and self-reports would be low. As a robustness check, I compared proxy information with self-reports from the secondary respondents who participated, for the items that were available in both questionnaires. The differences between the distributions are negligible, thus supporting the approach of using primary respondent responses as proxies (see Table A3.3).

The individual level characteristics include the educational level, distinguishing between non-tertiary education (1) and tertiary education (0), according to ISCED 11 standard classification. Having a migration background is based on a person’s citizenship and country of birth and is

coded as 1 “Has migration background” and 0 “Has no migration background”. Gender is coded as 1 “Male” and 0 “Female”. The gender category “Other” was coded as missing value, as the number of observations is too low for a separate category (for primary respondents $n=41$, for secondary sample members $n=14$). As a topic-related characteristic, the circumstance of having no own child(ren), coded as 1 “Childless” and 0 “Parent”, is added, as parents may have a stronger interest and social responsibility to participate in a survey dealing with topics such as family and fertility.

Dyadic level characteristics cover aspects of a couple’s relationship that measure the degree of commitment and closeness between two partners. To capture the level of situational closeness, I include the indicator of whether a couple is living in the same household or not, coded as 1 “Not cohabiting” and 0 “Cohabiting”. As a measurement of commitment between the partners, I include the marital status, indicating whether the couple is legally married (0) or not (1). To assess the degree of emotional closeness between the two respondents, an indicator for the quality of a relationship is used. The item asks “How satisfied are you with the relationship with your partner?” on a scale from 0 (totally unsatisfied) until 10 (totally satisfied). For a better interpretation of the predicted probabilities, additionally I calculate a dichotomous variable cut off by the median value. That means that individuals with 0 to 8 points are classified as having relationships with lower relationship satisfaction, and those reporting 9 and 10 points are categorized as having relationships with higher levels of relationship satisfaction.

To account for variations at the contact level, the way the secondary sample member is contacted is used: When invitation letters are sent directly to the secondary sample members by the survey agency, the item is coded as 0 “Invitation sent to SR”. When the invitation letter is sent to the primary respondent first, because the latter assumes the contact approach themselves, the item is coded as 1 “Invitation sent to PR”. The timing of the contact is measured by the number of tranches in which the invitation letters are sent to the secondary sample members. The earliest

tranche is coded as 1 “First tranche”, the following tranche is coded as 2 “Second tranche” and the last tranche is coded as 3 “Third tranche”.

3.3.3 Methodological approach

To address RQ1, I illustrate the nonresponse process and identify the steps leading to most dropouts, using descriptive numbers retrieved from the contact protocol data. Furthermore, I calculate cumulated percentages and estimate response rate according to AAPOR standard classifications (AAPOR, 2023).

For answering RQ2, I investigate the impact of the characteristics under study on the respective steps in estimating multivariate logistic regression models. The first model has the outcome of consent being granted (or not) by the primary respondent, and the second model has the outcome of participation or non-participation by the secondary sample member. The first model includes individual level variables of the primary respondent and dyadic variables of the relationship. Accordingly, the second model includes predictor variables at the individual level of the secondary sample member, dyadic variables of their relationship, and variables related to contactability. To avoid rescaling and confounding issues, I calculate average marginal effects to compare the effects within each model (Breen et al., 2018; Karlson et al., 2012). This provides information about the accuracy of my expectations regarding the prediction of nonresponse at each step.

To assess the relative strength of the predictors and to compare the magnitude of individual and dyadic effects between and within the two steps, I calculate predicted probabilities based on logistic regression models that are fitted only with the individual and dyadic level characteristics for both outcomes. The same number of variables and a similar measurement and coding of outcome and predictor variables enhances cross-model comparisons of predicted

probabilities (Mize et al., 2019). I estimate the predicted probability for different scenarios at the individual and dyadic level at each step, by estimating the marginal effect for a combination of predictors that were set to a representative value while setting all other covariates in the model to their mean value. To facilitate interpretation, I further calculate the predicted probability of a baseline scenario in which all variables in the model are set to the mean value, for both outcomes.

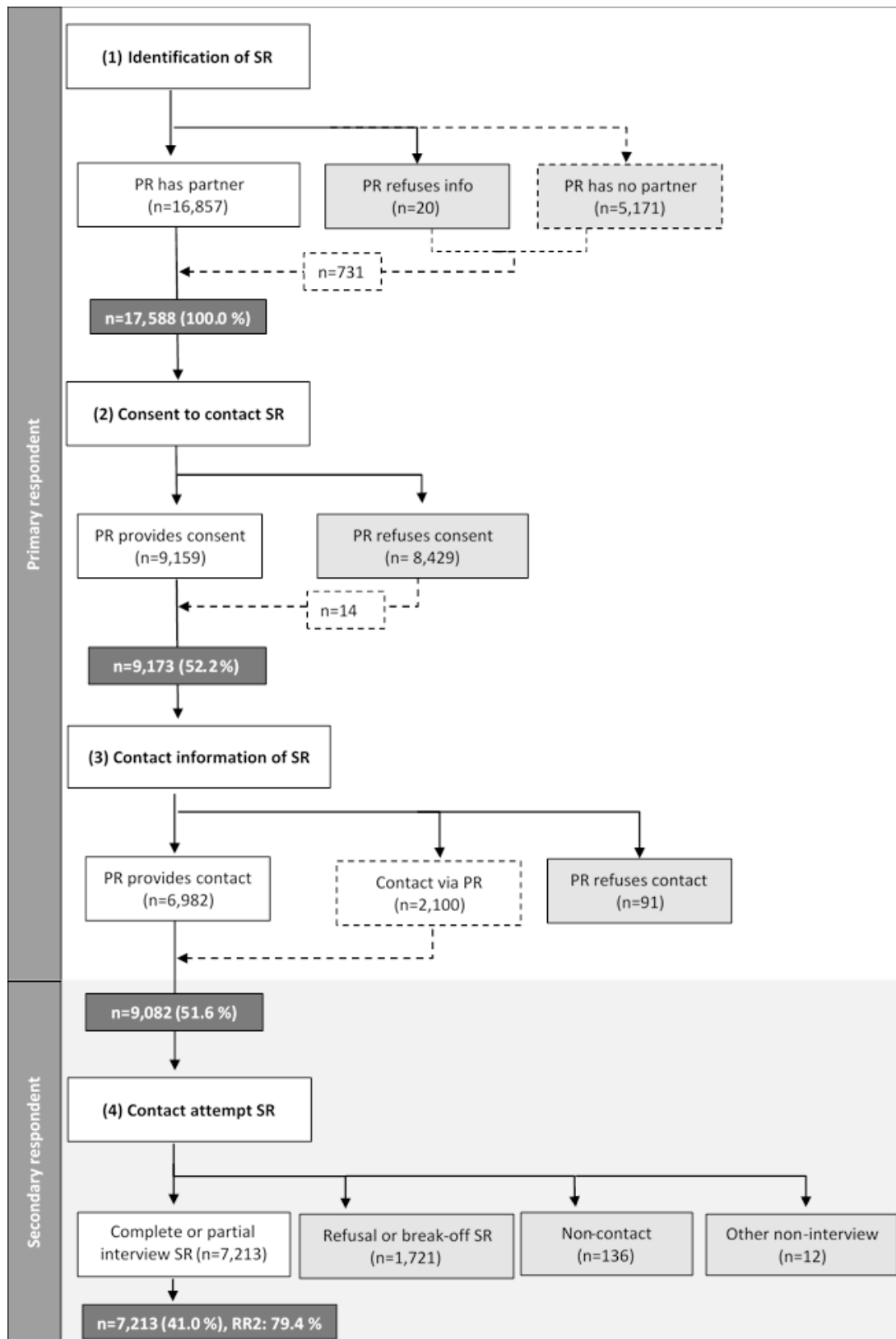
In order to illustrate the impact of variations in contact on participation, I calculate predicted probabilities for different combinations of contact scenarios. For those, I set the individual and dyadic predictors to their optimal representative value to get the best-case scenario in terms of participation likelihood. This allows to determine the extent to which contact variations can enhance or hinder the success of participation under the most favorable conditions.

Robustness check

As not much information is available for secondary nonrespondents, I had to omit including additional control variables that may be relevant for each single step, to ensure comparability between the models. A prior study based on the same dataset revealed that consent probability varies significantly depending on the survey mode (Gummer et al., 2023). However, since primary respondents can choose their preferred survey mode, the survey mode itself may act as a mediator variable. To avoid confounding effects with my individual level items under study and to keep the fitted models comparable, I refrained from including survey mode in my analysis. In order to check for the robustness of my findings, I estimated two separate models for PAPI and CAWI primary respondents on the likelihood to provide consent. Results indicated no consistent differences in selection effects between participants who used PAPI and CAWI (see Table A3.4).

3.4 Results

All in all, 41.0 percent of all potential secondary respondents identified at the beginning of the process completed an interview. When all eligible secondary sample members are defined as secondary sample units for whom we have contact information, and the sampling frame is determined accordingly, the calculated AAPOR RR2 is 79.4 percent (AAPOR, 2023) (see Figure 3.2). However, one has to consider that the participation decision takes place under different conditions than is the case for general participation in surveys, as there is already a pre-selection led by the primary respondent.



Source: FReDA Release v.3.0.0. and contact protocol partner data.

Note: Observations (n) with cumulated percentages (%) and calculated response rate 2 after AAPOR standard definitions (RR2). White-colored box means the case remains; gray-colored box means the case drops out. SR=secondary respondent, PR=primary respondent.

Figure 3.2: Nonresponse process for secondary respondents adapted to the FReDA study

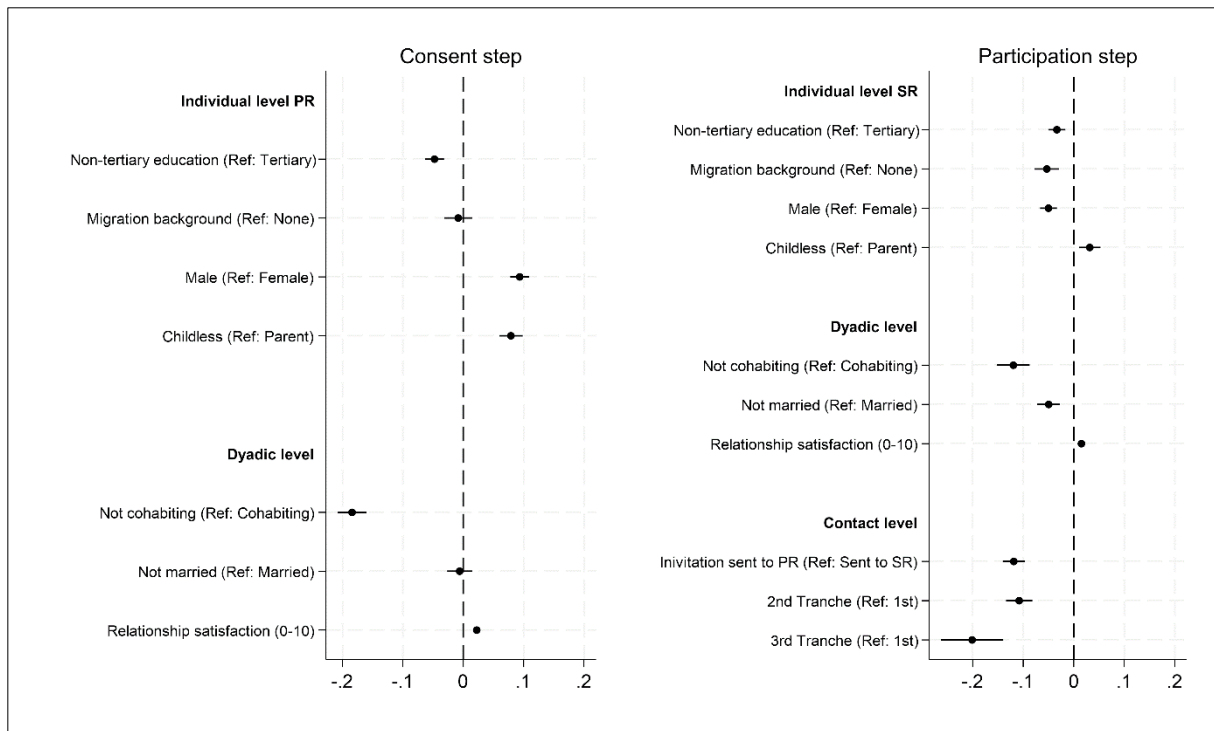
Following my first research question, RQ1, Figure 3.2 displays the nonresponse process adapted to numbers and specifications of the FReDA study. Processes that occur beyond the schematic process illustrated in Figure 3.1 are shown as dashed lines, to account for study-specific features of the FReDA survey. For instance, some primary respondents agreed to a secondary respondent's interview without prior identification, as did some who did not provide consent but shared contact details¹³. Additionally, a further opportunity to establish contact is offered for primary respondents, as outlined previously.

Based on the numbers in Figure 3.2, my first hypothesis, H1, can be supported: Primary respondents' refusal to consent during step 2 and secondary respondents' refusal to participate (or break off) during step 4 lead to the most dropouts of potential secondary respondents, compared to the other steps. Hereby, nearly half of all secondary sample members break away because the primary respondent refuses consent.

To address my second research question, RQ2, I assess which aspects may predict nonresponse at the particular step. Therefore, Figure 3.3 shows the average marginal effects of the predictor variables on the likelihood of granting consent and of participating, to assess whether individual, dyadic and contact-related aspects affect nonresponse at the two steps¹⁴.

¹³ Such inconsistencies may stem from limitations in data collection, especially with paper-based questionnaires lacking automated routing or interviewer support. However, these occurrences are minimal (n=745) and do not significantly affect the overall outcome.

¹⁴ The regression results for the two models can be found in Tables A3.5 and A3.6 in the Appendix.



Source: FReDA Release v 3.0.0.

Note: Average marginal effects (AME) with 95% confidence intervals. SR=secondary respondent; PR=primary respondent. $N_{\text{Consent}}=15,182$; $N_{\text{Participation}}=8,414$.

Figure 3.3: Average marginal effects on the likelihood of providing consent by primary respondents and of participation by secondary respondents

Starting with the consent step (left hand side of Figure 3.3), most effects for characteristics at the individual level contradict my theoretical assumptions. Only hypothesis H_c2.1 can be supported: Primary respondents with lower levels of education show a probability of granting consent that is 4.8 percentage points lower compared to those with tertiary education. Whereas having a migration background does not affect the consent step, men have a higher probability of providing consent than women. Similarly, childless persons are more likely to provide consent than parents. Thus, H_c2.2, H_c2.3 and H_c2.4 are not supported by my findings.

At the dyadic level, most effects are in line with my expectations. The fact that a primary respondent is not living together with the secondary respondent shows the strongest impact, as non-cohabitation reduces the probability of granting consent by about 18.2 percentage points. Furthermore, primary respondents with lower levels of relationship satisfaction are less likely

to agree to an interview for the secondary respondent, which supports $H_{c3.1}$ and $H_{c3.3}$. However, $H_{c3.2}$ must be rejected, as being married with the secondary respondent shows no impact on consent.

For the participation step (right hand side of Figure 3.3), almost all effects at the individual level are in accordance with what I expected. Poorly-educated persons are more likely to refuse participation than highly-educated ones, by about 3.3 percentage points. Whereas having a migration background did not affect consent, secondary sample members with migration backgrounds have a probability of participation that is 5.3 percentage points lower compared to natives. Unlike with the consent step situation, men here are less likely to participate, by about 5.0 percentage points. That means $H_{p2.1}$, $H_{p2.2}$ and $H_{p2.3}$ are supported by my findings. Similar to the consent step, childless persons are more likely to participate than parents, therefore $H_{p2.4}$ must be rejected.

Dyadic aspects also play a role during the participation step: Here again, living apart from the partner shows the strongest association with nonresponse, as it lowers the chances of participation by 11.9 percentage points compared to those living together. Additionally, unmarried secondary sample members and those with lower relationship satisfaction are more likely to be nonrespondents, and respectively, are less likely to participate, supporting all hypotheses at the dyadic level ($H_{p3.1}$, $H_{p3.2}$, and $H_{p3.3}$).

The way contact is established also reveals a strong association with the likelihood to participate. Sending the invitation to the primary respondent significantly reduces the participation probability of secondary sample members, by 11.9 percentage points, compared to sending the invitation directly to the latter, supporting $H_{p4.1}$. In addition to the way in which contact is made, the timing of invitations plays a significant role in the process. Secondary sample members who receive their invitations with the second tranche have a much lower likelihood of

participating, 10.8 percentage points less, compared to those who receive their invitations in the first tranche. The difference increases to 20.1 percentage points if they receive their invitation with the third tranche, supporting $H_{p4.2}$. Table 3.1 provides an overview of all hypotheses and summarizes whether they can be supported by my findings.

Table 3.1: Overview of tested hypotheses and whether they can be supported

RQ1: Which steps lead to the most dropouts of potential secondary respondents in a self-administered multi-actor survey?			
Hypothesis for process		Supported?	
H1: The refusal to provide consent (step 2) and the refusal to participate (step 4) lead to a higher loss of secondary sample members during the process than the other steps (step 1 and 3).		Yes	
RQ2: Which aspects predict nonresponse at the particular step?			
Hypotheses for consent step	Supported?	Hypotheses for participation step:	Supported?
H _{c2.1} : Primary respondents with lower levels of education are less likely to provide consent than primary respondents with higher levels of education.	Yes	H _{p2.1} : Secondary sample members with lower levels of education are less likely to provide consent than secondary sample members with higher levels of education.	Yes
H _{c2.2} : Primary respondents with migration backgrounds are less likely to provide consent than primary respondents without migration backgrounds.	No	H _{p2.2} : Secondary sample members with migration backgrounds are less likely to participate than secondary sample members without migration backgrounds.	Yes
H _{c2.3} : Male primary respondents are less likely to provide consent than female primary respondents.	No	H _{p2.3} : Male secondary sample members are less likely to participate than female secondary sample members.	Yes
H _{c2.4} : Primary respondents who are childless are less likely to provide consent than primary respondents who are parents.	No	H _{p2.4} : Secondary sample members who are childless are less likely to participate than secondary sample members who are parents.	No
H _{c3.1} : Primary respondents that are not cohabiting with their partner are less likely to provide consent.	Yes	H _{p3.1} : Secondary sample members who are not cohabiting with their partner are less likely to participate than those who live together with their partner.	Yes
H _{c3.2} : Primary respondents who are not married to their partner are less likely to provide consent than primary respondents who are married.	No	H _{p3.2} : Secondary sample members who are not married to their partner are less likely to participate than secondary sample members who are married.	Yes
H _{c3.3} : The lower the level of relationship satisfaction, the less likely it is that primary respondents will provide consent.	Yes	H _{p3.3} : The lower the level of relationship satisfaction, the less likely it is that secondary sample members will provide consent.	Yes

Table 3.1: Overview of tested hypotheses and whether they can be supported (*continued*)

	H_p4.1: Sending the invitation to the primary respondent led to a lower probability of participation compared to sending the invitation directly to the secondary sample member.	Yes
	H_p4.2: Sending invitations in later tranches led to a lower probability of participation compared to sending invitations in earlier tranches.	Yes

To further assess the magnitude of the individual and dyadic level impact within and between the two steps, Table 3.2 displays the predicted probabilities of granting consent and of participating at the individual and dyadic level each.

Table 3.2: Predicted probabilities for scenarios at the individual and dyadic level for providing consent and participating while holding other variables at their means

Consent step					
<i>Individual level</i>					
		Woman		Man	
		Low education	High education	Low education	High education
Migrant	Parent	0.44	0.49	0.54	0.59
	Childless	0.53	0.58	0.62	0.67
Native	Parent	0.45	0.50	0.55	0.60
	Childless	0.54	0.58	0.63	0.68
<i>Dyadic level</i>					
		Cohabiting		Not cohabiting	
		Married	Not married	Married	Not married
High satisfaction		0.62	0.61	0.43	0.42
Low satisfaction		0.56	0.55	0.37	0.36
Participation step					
<i>Individual level</i>					
		Woman		Man	
		Low education	High education	Low education	High education
Migrant	Parent	0.74	0.78	0.67	0.72
	Childless	0.79	0.82	0.73	0.77
Native	Parent	0.81	0.83	0.75	0.78
	Childless	0.85	0.87	0.80	0.83
<i>Dyadic level</i>					
		Cohabiting		Not cohabiting	
		Married	Not married	Married	Not married
High satisfaction		0.86	0.81	0.75	0.68
Low satisfaction		0.82	0.76	0.69	0.61

Source: FReDA Release v 3.0.0.

Note: Cells highlighted in dark gray indicate predicted probabilities above the baseline scenario where all variables in the model were held at their means; Bordered cells indicate the best- and worst-case scenarios. N_{Consent}=15,182; N_{Participation}=8,414.

The cells highlighted in dark gray indicate that the predicted probabilities are higher than in the baseline scenario, where all variables in the model are held at their mean values. In the baseline scenario, the predicted probability for providing consent is 0.56, while the predicted probability for participation is 0.81. As already shown in the descriptive numbers in Figure 3.2, the probability of consenting is generally lower than the probability of participating. However, the magnitude of the dyadic and individual effects on the two outcomes is similar.

At the individual level, poorly educated women with children display the lowest probability of consenting, at about 0.44 percent. The difference associated with having a migration background or being native is not significant (see Figure 3.3). The predicted probability of consenting increases by 0.24 percent, to 0.68 percent, for highly educated men without children. In contrast, poorly educated men with a migration background and children show the lowest predicted probability for participation, at about 0.67 percent. The probability increases by 0.20 percent, to 0.87 percent, when the secondary respondent is female, highly educated, native and childless.

Dyadic characteristics under study seem to affect the outcomes more than the individual characteristics under study, for both steps. Although the predicted probability of consent for highly satisfied cohabiting couples (being married shows no significant difference, see Figure 3.2) is 0.62 percent, the probability of consent drops by 0.26 percent, to 0.36 percent, if the respondents do not live together and are less satisfied in their relationship. Something similar happens with participation: Secondary respondents in relationships where both partners are cohabiting, are married and have high levels of relationship satisfaction, have the highest probability of participation, at about 0.86 percent. However, this probability decreases by 0.25 percent, to reach 0.61 percent, if secondary respondents experience weaker dyadic ties, i.e., when they do not live together with their partner, are not married to the latter and have lower relationship satisfaction. This suggests that both the dyadic and individual levels have a comparable and

significant impact on both stages. However, dyadic characteristics appear to have a greater impact on the nonresponse probability at both steps.

As the contact attempt is only relevant in the final step, Table 3.3 displays the predicted probabilities for different contact scenarios when both individual and dyadic characteristics are optimized. In the best-case scenario at both the individual and dyadic levels, the predicted probability of participation is 0.91 percent for highly educated women without a migration background who have no children and are in a relationship where both partners live together, are married, and have a high level of relationship satisfaction.

Table 3.3: Predicted probabilities of participation, while holding other variables at their optimal values, for scenarios at the contact level

Participation step	Contact level	
	Invitation sent to SR	Invitation sent to PR
First tranche	0.93	0.87
Second tranche	0.87	0.77
Third tranche	0.82	0.69

Source: FReDA Release v 3.0.0.

Note: Cells highlighted in dark grey indicate predicted probabilities above the best-case scenario, where variables at the dyadic and individual level in the model were held at their optimal values. $N_{\text{Participation}}=8,414$.

It is evident that the probability of participation can be increased further, to 0.93 percent, if the invitation is sent directly to the secondary respondent and is received with the first tranche. Under the same conditions, the probability of participation decreases by 0.11 percent if the invitation arrives with the last tranche. If the contact is initiated by the primary respondent, there is a 0.87 percent probability of participation when the invitation arrives with the first tranche. However, despite the individual and dyadic predictors, the probability drops by almost 0.20 percent to only 0.69 percent when the invitation is sent out with the third tranche. This illustrates that contact conditions can significantly diminish the chances of participation, regardless of the initial individual and dyadic situations.

3.5 Discussion

Conclusion

When and why do we lose secondary respondents when collecting multi-actor data? The purpose of this study was to approach these two questions in investigating the nonresponse process of secondary respondents, on the basis of a self-administered multi-actor survey.

My findings suggest that obtaining consent from the primary respondent, and inviting the partner to participate as a secondary respondent, are found to be the most pivotal stages during the process, whereby most secondary sample member dropouts happened because primary respondents refused to provide consent. However, the loss at these two steps did not occur at random, since both the individual characteristics of each person and the dyadic characteristics of the relationship seemed to shape each actor's evaluation of the survey request. Furthermore, the design specifications for contacting secondary sample members in multi-actor surveys favored non-participation additionally.

Characteristics at the individual level affected the two steps in different ways, suggesting that there were disparate reasons and motives at work. For the participation step, sociodemographic characteristics that are generally underrepresented in surveys, and that make people harder to reach, such as being male, poorly educated or non-native (Radler & Ryff, 2010; Tourangeau, 2014), are also found to be associated with a lesser likelihood of participation. Accordingly, I expected that factors associated with a higher respondent burden should also be a predictor for consent refusal (Groves & Couper, 1998; Müller, 2017). However, this was only the case for primary respondents with a low level of education. On the one hand, primary respondents are likely to be already selective when it comes to these attributes, and those who are unmotivated or see too great a burden in the survey request would not participate in the study to begin with. On the other hand, gatekeeping as a driver of non-consent may also be more pronounced among

certain demographic subgroups. For example, women might be more inclined to prevent access to the survey because they either do not want their male partner to take part in a study on family and fertility or they already anticipate that their partner would have no interest in these topics, as family and fertility are traditionally seen as a female domain (Slauson-Blevins & Johnson, 2016). Childless individuals have been found to be more likely to consent and to participate, which contradicts my hypothesis, i.e., that parents would be more interested in the study's topic and accordingly, would be more motivated to engage in such a survey. However, at the same time parents, especially of younger children, might have limited time resources due to a greater care workload. Thus, greater time constraints for persons with children could be a more influential factor for compliance with a survey request than are topic-related motivations.

Dyadic features that express lower levels of commitment, closeness and quality within the relationship seem to foster nonresponse consistently during the consent and participation step. These aspects may be associated with poor communication, more uncertainty, less emotional support, and a lack of shared values, contributing to a lower motivation to support the other person and a greater proclivity towards gatekeeping among actors. In particular, living apart from the partner was found to have the greatest influence at both steps. Previous research attributed the positive association between participation and co-residence primarily to the personal interview situation in the respondent's household, as the setting was expected to favor the decision to participate (Müller, 2017; Schröder et al., 2016). However, present findings are based on a self-administered multi-actor survey, showing for the first time that interviewer effects cannot be considered to be decisive. Instead, it is likely that a closer connection between the two respondents, more frequent contact and greater mutual control over the other person could be reasons behind this association. These findings support my assumption that positive dyadic dynamics result in less nonresponse. However, one must bear in mind that topic-related factors may also play a role. For example, a survey about intimate relationships and family lives

might generally lead to a selection of “happier” and thus more socially desirable types of relationships (Kalmijn, 2023).

At the contact level, design-specifications of the FReDA study revealed that both the type of contact and the timing were strongly associated with nonresponse at the participation step. Establishing contact through the primary respondent, instead of sending the invitation directly to the secondary sample member, seems to increase nonresponse. This may be attributed to the fact that doing the former creates an additional gatekeeping situation that allows the primary respondent to reconsider forwarding the invitation. However, one must consider whether the primary respondent would consent to an interview with the secondary respondent without this option in the first place, as the former may be generally more skeptical toward the survey. The timing of secondary respondent’s receipt of the invitation letter also plays an important role. Findings showed that sending the invitation at a later stage significantly reduced the probability of participation. Probably, the primary respondent is more likely to remember the study and to remain aware of the invitation when the time span between the primary respondent interview and the receipt of the invitation is shorter. However, I cannot determine whether this is due to the elapsed time or to the fact that the primary respondents completed their interview later and therefore may have been less receptive to the study as a whole. This could also have an impact on the evaluation by the secondary respondent. Since both factors were not tested experimentally, this study is unable to draw clear conclusions about the underlying motives. However, these findings strongly suggest that further research on the practical aspects of multi-actor data collection would be conducive to achieving higher response rates altogether.

Taken together, individual and dyadic aspects are associated significantly with consent and participation, although there are variations in the magnitude and direction of effects, especially for sociodemographic characteristics. Selection effects at the dyadic level seem to accumulate throughout the process, increasing the selectivity of the dyadic sample towards closer and more

committed relationships. However, the impact of individual aspects varies between the two stages, indicating divergent selection effects of population subgroups during the process. Apart from the individual and dyadic situations, design-specific features of contactability in multi-actor surveys can strongly diminish the chances of participation and can lead to an additional loss of secondary sample members.

Practical implications

These findings show that there are different steps in collecting multi-actor data that need to be considered. Therefore, developing methods and approaches that are tailored to the steps and persons involved in this process is advisable.

Generally, it is important to reduce possible obstacles that may prevent the primary respondent from consenting, as this seems to be the most crucial step when conducting dyadic data gathering. Once consent is obtained, secondary respondents' response rate is much higher than response rates in other self-administered surveys in Germany (Daikeler et al., 2020; Wolf et al., 2021). However, despite the lower nonresponse rate during this step, there are still pronounced selection effects that lead to lower data quality of the dyadic sample.

Firstly, the consent step serves as a gatekeeping situation because access to the target person and accordingly, the information of interest, is only granted if the primary respondent provides his or her permission (Singh & Wassenaar, 2016). Therefore, data collectors should devise ethically acceptable measures that alleviate the gatekeeping tendencies and the concerns of the primary respondents in order to reduce barriers to obtaining their consent. Accordingly, the survey agency can address questions regarding participants' privacy and data security and emphasize the importance of confidentiality and anonymity. This may also increase the willingness of primary respondents to share the contact information of secondary respondents, rather than assuming the task of contacting secondary respondents themselves, which has shown to end in

a decreased participation by secondary respondents. Furthermore, data collectors could also emphasize the benefits of granting consent. For example, the consent request could be framed more strongly as a request for help in better exploring relationship dynamics and problems. At the same time, this approach may encourage primary respondents in relationships that do not conform to socially desirable norms to consent, thereby increasing participants' sense of social responsibility.

Secondly, data collectors should make the cooperation of all actors more attractive by addressing the needs and considerations that might affect dyadic processes and negotiations between the two actors. In that respect, monetary incentives or incentives tailored to the specific dyadic target group could be offered when requesting consent. This leverages the mutual influence of both actors, as the incentive is conditional on the cooperation (involving consent and participation) of both parties. For instance, in the context of a couple relationship, couples could be given the opportunity to win or receive vouchers for a romantic dinner, movie tickets, or access to relationship workshops to promote relationship health and communication skills. By providing incentives tailored to the dyads' interests and preferences, data collectors can encourage active support from both parties, strengthen relationships, and demonstrate appreciation for their joint participation in surveys and research initiatives.

Thirdly, although secondary respondents are generally more likely to participate than primary respondents are likely to consent, the findings indicate that participation is notably lower among hard-to-survey populations, suggesting that the burden for particular demographic subgroups is too high (Tourangeau, 2014). Therefore, testing additional measures that are commonly employed in surveys to enhance the response rate of these particular groups might be useful. For instance, invitation letters should be concise and easy to understand. Additionally, if the budget allows, a higher incentives scheme can be considered, as literature shows that it increases participation among such groups as ethnic minorities and migrants (Fomby et al., 2017; Martinez-

Ebers, 1997). However, it is important to take ethical considerations into account when inventing methods to motivate secondary sample members to participate (see Wittenborn et al., 2013 for a discussion). The high response rate could indicate that secondary respondents feel obligated to participate because the primary respondent has already given consent. It is crucial to make clear that participation is still voluntary and that they are not forced to engage in the survey because their counterpart already agreed to the interview. Therefore, it is necessary to evaluate the ethical justifiability of additional measures to increase participation rates. One way to monitor this is by administering a questionnaire to evaluate the survey process and the experience for secondary respondents.

Lastly, it is recommendable to send out the invitation letters closer to the primary respondent's interview date in order to minimize the risk of the survey being forgotten too quickly. Moreover, whenever feasible, it is preferable to send the invitation directly to the secondary respondent, rather than through the primary respondent.

Limitations

This study is not without limitations. First of all, separating effects between the individual steps is challenging because various selection processes are at work and can influence each other. Therefore, this study cannot draw any causal conclusion about the mechanisms, but it can show correlations that support my theoretical assumptions. Selection processes may start as early as with the participation of the primary respondent, since specific sociodemographic groups and individuals with negative relationship experiences may be less likely to engage. This, in turn, can also influence the composition of the secondary sample, as it is likely that primary respondents have partners that share similar characteristics, such as ethnicity, and may show educational homogamy (Domański & Przybysz, 2007). However, this confirms that additional selection effects are present, suggesting that my findings still underestimate such selection processes.

Second, the dataset used provided only information on secondary nonrespondents reported by the primary respondents. At the same time, FReDA data consists of a high number of observations and covers a range of questions on intimate relationships due to its thematic focus. This enabled me to retrieve dyadic information and provided an opportunity to conduct a thorough analysis of nonresponse. Generally, surveys often lack detailed information beyond sociodemographic characteristics about individuals who choose not to participate and have to rely on, say, register data (Groves & Peytcheva, 2008). Nevertheless, statements from third parties may deviate from self-reports and may have a lower validity, especially for assessing subjective items, like relationship quality. Although it is likely that both primary and secondary respondents have similar experiences in their relationship, I cannot dismiss the possibility that the secondary respondent may perceive the relationship differently, which could influence their participation behavior. However, my findings suggest that nonresponse is not only related to the sociodemographic situation of a person, but also to relationship characteristics which are particularly relevant to data users in the field of family research. Therefore, data collectors of such surveys should make efforts to track nonresponse that is specific or sensitive to the survey topic, for example, by implementing follow-up questionnaires for those who did not participate, to assess more in-depth information on nonrespondents.

Third, this study aimed to compare the selection effects between the primary respondents' consent and the secondary sample members' participation, in order to trace selection processes. Therefore, in my analysis I was limited to the use only of variables that are available for secondary nonrespondents and thus, I focused solely on a set of predictors to determine whether and to what extent similar effects occur during the process. Due to the limited number of variables available for the nonrespondents, I was unable to include additional items to control for possible confounding effects at each step in my models.

Thus, future research should delve deeper into the respective steps and, based on the survey topic and research question, investigate reasons and motives of possible dyadic predictors that might have an impact on nonresponse in the process. However, this is beyond the scope of this paper. The aim of this study was to gain an initial understanding of the nonresponse process and to identify relevant sets of indicators to adjust at each step where selection can occur.

References

- AAPOR. (2023). *AAPOR's standard definitions, version 10*. <https://aapor.org/response-rates/>
- Atkeson, L. R., Adams, A. N., & Alvarez, R. M. (2014). Nonresponse and mode effects in self- and interviewer-administered surveys. *Political Analysis*, 22(3), 304–320. <https://doi.org/10.1093/pan/mpt049>
- Barton, A. W., Lavner, J. A., Stanley, S. M., Johnson, M. D., & Rhoades, G. K. (2020). “Will you complete this survey too?” Differences between individual versus dyadic samples in relationship research. *Journal of Family Psychology*, 34(2), 196–203. <https://doi.org/10.1037/fam0000583>
- Bastaitis, K., Pasteels, I., Ponnet, K., & Mortelmans, D. (2015). Adult non-response bias from a child perspective. Using child reports to estimate father's non-response. *Social Science Research*, 49, 31–41. <https://doi.org/10.1016/j.ssresearch.2014.07.004>
- Beullens, K., Loosveldt, G., Vandenplas, C., & Stoop, I. (2018). Response rates in the European Social Survey: Increasing, decreasing, or a matter of fieldwork efforts? *Survey Methods: Insights from the Field*, 1–12. <https://doi.org/10.13094/SMIF-2018-00003>
- Breen, R., Karlson, K. B., & Holm, A. (2018). Interpreting and understanding logits, probits, and other nonlinear probability models. *Annual Review of Sociology*, 44(1), 39–54. <https://doi.org/10.1146/annurev-soc-073117-041429>
- Brehm, J. O. (2009). *The phantom respondents: Opinion surveys and political representation* [Doctoral dissertation, University of Michigan]. University of Michigan ProQuest Dissertations & Theses.
- Bujard, M., Gummer, T., Hank, K., Neyer, F. J., Pollak, R., Schneider, N. F., Spieß, C. K., Wolf, C., Bauer, I., Börlin, S., Bretsch, D., Brüggemann, K., Christmann, P., Edinger, R., Eigenbrodt, F., Frembs, L. C., Groß, K., Jost, C., Kunz, T., . . . Weih, U. (2023). *FReDA – Das familiendemografische Panel*. <https://doi.org/10.4232/1.14080>

- Carroll-Lind, J., Chapman, J. W., Gregory, J., & Maxwell, G. (2006). The key to the gatekeepers: Passive consent and other ethical issues surrounding the rights of children to speak on issues that concern them. *Child Abuse & Neglect*, 30(9), 979–989. <https://doi.org/10.1016/j.chiabu.2005.11.013>
- Chaplin, J., Mangla, J., Purdon, S., & Airey, C. (2005). *The workplace employment relations survey (WERS) 2004*. Technical report, National Centre for Official Research, London.
- Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75(5), 889–908. <https://doi.org/10.1093/poq/nfr046>
- Couper, M. P. (2017). New developments in survey data collection. *Annual Review of Sociology*, 43(1), 121–145. <https://doi.org/10.1146/annurev-soc-060116-053613>
- Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8(3), 513–539. <https://doi.org/10.1093/jssam/smz008>
- Dillman, D. A., Eltinge, J.L., Groves, R.M. & Little, R. J. A. (2002). Survey nonresponse in design, data collection, and analysis. In R. M. Groves, D. A. Dillman, J. L. Eltinge & R. J. A. Little (Eds.). *Survey nonresponse* (pp. 3-26). Wiley.
- Domański, H., & Przybysz, D. (2007). Educational homogamy in 22 European countries. *European Societies*, 9(4), 495–526. <https://doi.org/10.1080/14616690701314119>
- Dykstra, P. A., Kalmijn, M., Knijn, T. C. M., Komter, A. E., Liefbroer, A. C., & Mulder, C. H. (2005). *Codebook of the Netherlands Kinship Panel Study*. The Netherlands Interdisciplinary Demographic Institute (NIDI), The Hague.
- Esser, H. (1986). Über die Teilnahme an Befragungen. *Zuma Nachrichten*, 10(18), 38–47.
- EU GDPR. (2016). *Regulation (EU) 2016/679 of the European parliament and of the council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC*

- (General Data Protection Regulation). <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679&qid=1685616322175>
- Fokkema, T., Liefbroer, A. C., Emery, T., Hiekel, N., & Kveder, A. (2016). Generations and Gender Programme wave 1 data collection: An overview and assessment of sampling and fieldwork methods, weighting procedures, and cross-sectional representativeness. *Demographic Research*, 34(18), 499–524.
- Fomby, P., Sastry, N., & McGonagle, K. A. (2017). Effectiveness of a time-limited incentive on participation by hard-to-reach respondents in a panel study. *Field Methods*, 29(3), 238–251. <https://doi.org/10.1177/1525822X16670625>
- Fuchs, M., Bossert, D., & Stukowski, S. (2013). Response rate and nonresponse bias - Impact of the number of contact attempts on data quality in the European Social Survey. *Bulletin of Sociological Methodology/ Bulletin De Méthodologie Sociologique*, 117(1), 26–45. <https://doi.org/10.1177/0759106312465547>
- Groves, R. M., Cialdini, R., & Couper, M. P. (1992). Understanding the decision to participate in a survey. *Public Opinion Quarterly*, 56(4), 475-495. <https://doi.org/10.1086/269338>
- Groves, R. M., & Couper, M. P. (1998). *Nonresponse in household interview surveys*. Wiley.
- Groves, R. M., & Peytcheva, E. (2008). The impact of nonresponse rates on nonresponse bias: A meta-analysis. *Public Opinion Quarterly*, 72(2), 167–189. <https://doi.org/10.1093/poq/nfn011>
- Groves, R. M., Presser, S., & Dipko, S. (2004). The role of topic interest in survey participation decisions. *Public Opinion Quarterly*, 68(1), 2–31. <https://doi.org/10.1093/-poq/nfh002>
- Groves, R. M., Singer, E., & Corning, A. (2000). Leverage-saliency theory of survey participation: Description and an illustration. *Public Opinion Quarterly*, 64(3), 299–308. <http://www.jstor.org/stable/3078721>

- Gummer, T., Christmann, P., & Kunz, T. (2023). Gaining consent to survey respondents' partners: The importance of anchors' survey experience in self-administered modes. *Comparative Population Studies*, 48. <https://doi.org/10.12765/CPoS-2023-12>
- Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T., Lück, D., & Neyer, F. J. (2020). The impact of Covid-19 on fieldwork efforts and planning in pairfam and FReDA-GGS. *Survey Research Methods*, 14(2), 223–227. <https://doi.org/10.18148/srm/2020.v14i2.7740>
- Hagedoorn, M., Hein, F. L., Schulz, T., van der Heide, J. J. H., Niesing, J., Westerhuis, R., Ploeg, R. J., & Ranchor, A. V. (2015). Are patient and relationship variables associated with participation of intimate partners in couples research? *Health Psychology*, 34(3), 270. <https://doi.org/10.1037/hea0000141>
- Hank, K., Gummer, T., Bujard, M., Neyer, F. J., Pollak, R., Spieß, C. K., ... & Thönissen, C. (2024). A new data infrastructure for family research and demographic analysis: The German Family Demography Panel Study (FReDA). *European Sociological Review*, jcae019. <https://doi.org/10.1093/esr/jcae019>
- Huinink, J., Brüderl, J., Nauck, B., Walper, S., Castiglioni, L., & Feldhaus, M. (2011). Panel analysis of intimate relationships and family dynamics (PAIRFAM): Conceptual framework and design. *Zeitschrift für Familienforschung*, 23(1), 77–101. <https://doi.org/10.20377/jfr-235>
- Hünteler, B., & Wetzel, M. (2020). Relationship-based selective participation of secondary respondents in a German multi-actor panel study. *Journal of Marriage and Family*, 82(5), 1677–1695. <https://doi.org/10.1111/jomf.12671>
- Johnson, D. R., McQuillan, J., Jacob, M. C., Greil, A. L., Lacy, N., Scheuble, L. K., Johnson, K., Shreffler, K., & Young, R. (2009). *National survey of fertility barriers: Methodology report for wave 1* (Working Paper). <https://digitalcommons.unl.edu/nsfbworking/1/>

- Kalmijn, M. (2023). Are national family surveys biased toward the happy family? A multiactor analysis of selective survey nonresponse. *Sociological Methods & Research*, 52(2), 561–586. <https://doi.org/10.1177/0049124120986208>
- Kalmijn, M., & Liefbroer, A. C. (2011). Nonresponse of secondary respondents in multi-actor surveys: Determinants, consequences, and possible remedies. *Journal of Family Issues*, 32(6), 735–766. <https://doi.org/10.1177/0192513X10390184>
- Karlson, K. B., Holm, A., & Breen, R. (2012). Comparing regression coefficients between same-sample nested models using logit and probit: A new method. *Sociological Methodology*, 42(1), 286–313. <https://doi.org/10.1177/00811750124448>
- Korkeila, K., Suominen, S., Ahvenainen, J., Ojanlatva, A., Rautava, P., Helenius, H., & Koskenvuo, M. (2001). Non-response and related factors in a nation-wide health survey. *European Journal of Epidemiology*, 17(11), 991–999. <https://doi.org/10.1023/A:1020016922473>
- Lutig, P., Toepoel, V., Emery, T., Cabaço, S. L. F., Bujard, M., Naderi, R., Schumann, A., & Lück, D. (2022). *Can we successfully move a cross-national survey online? Results from a large three-country experiment in the Generations and Gender Programme* (Working Paper). Center for Open Science. <https://doi.org/10.31235/osf.io/mu8jy>
- Lynn, P. (2008). The problem of nonresponse. In E. D. de Leeuw, J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 35–55). Taylor & Francis.
- Lynn, P., & Clarke, P. (2002). Separating refusal bias and non-contact bias: Evidence from UK national surveys. *Journal of the Royal Statistical Society Series D: The Statistician*, 51(3), 319–333. <https://doi.org/10.1111/1467-9884.00321>
- Manfreda, K. L., Bosnjak, M., Berzelak, J., Haas, I., & Vehovar, V. (2008). Web surveys versus other survey modes: A meta-analysis comparing response rates. *International Journal of Market Research*, 50(1), 79–104. <https://doi.org/10.1177/147078530805000107>

- Martinez-Ebers, V. (1997). Using monetary incentives with hard-to-reach populations in panel surveys. *International Journal of Public Opinion Research*, 9(1), 77–86.
<https://doi.org/10.1093/ijpor/9.1.77>
- Mitchell, C. (2010). Are divorce studies trustworthy? The effects of survey nonresponse and response errors. *Journal of Marriage and Family*, 72(4), 893–905.
<https://doi.org/10.1111/j.1741-3737.2010.00737.x>
- Mize, T. D., Doan, L., & Long, J. S. (2019). A general framework for comparing predictions and marginal effects across models. *Sociological Methodology*, 49(1), 152–189.
<https://doi.org/10.1177/0081175019852763>
- Mortelmans, D., Pasteels, I., Bracke, P., Matthijs, K., van Bavel, J., & van Peer, C. (2011). *Scheiding in Vlaanderen*. Leuven, Belgium ; The Hague, Netherlands: Acco.
- Müller, B. (2017). Opportunity or affection? Participation of respondents' partners in a multi-actor survey. *Field Methods*, 29(3), 252–265.
<https://doi.org/10.1177/1525822X16671702>
- Olson, K., Smyth, J. D., Horwitz, R., Keeter, S., Lesser, V., Marken, S., Mathiowetz, N. A., McCarthy, J. S., O'Brien, E., Opsomer, J. D., Steiger, D., Sterrett, D., Su, J., Suzer-Gurtekin, Z. T., Turakhia, C., & Wagner, J. (2020). Transitions from telephone surveys to self-administered and mixed-mode surveys: AAPOR Task Force Report. *Journal of Survey Statistics and Methodology*, 9(3), 381–411.
<https://doi.org/10.1093/jssam/smz062>
- Pasteels, I. (2015). How to weight survey data with a dyadic multi-actor design? *Survey Methods: Insights from the Field*, 1–11. <https://doi.org/10.13094/SMIF-2015-00007>
- Peytchev, A. (2013). Consequences of survey nonresponse. *The ANNALS of the American Academy of Political and Social Science*, 645(1), 88–111.
<https://doi.org/10.1177/0002716212461748>

- Radler, B. T., & Ryff, C. D. (2010). Who participates? Accounting for longitudinal retention in the MIDUS national study of health and well-being. *Journal of Aging and Health*, 22(3), 307–331. <https://doi.org/10.1177/0898264309358617>
- Sakshaug, J. W., Couper, M. P., Ofstedal, M. B., & Weir, D. R. (2012). Linking survey data and administrative records: Mechanisms of consent. *Sociological Methods & Research*, 41(4), 535–569. <https://doi.org/10.1177/0049124112460381>
- Sakshaug, J. W., Hülle, S., Schmucker, A., & Liebig, S. (2017). Exploring the effects of interviewer- and self-administered survey modes on record linkage consent rates and bias. *Survey Research Methods*, 11(2), 171–188. <https://doi.org/10.18148/srm/2017.v11i2.7158>
- Sakshaug, J. W., Hülle, S., Schmucker, A., & Liebig, S. (2020). Panel survey recruitment with or without interviewers? Implications for nonresponse, panel consent, and total recruitment bias. *Journal of Survey Statistics and Methodology*, 8(3), 540–565. <https://doi.org/10.1093/jssam/smz012>
- Sala, E., Burton, J., & Knies, G. (2012). Correlates of obtaining informed consent to data linkage: Respondent, interview, and interviewer characteristics. *Sociological Methods & Research*, 41(3), 414–439. <https://doi.org/10.1177/0049124112457330>
- Schäfer, W., Boerma, W. G. W., Kringos, D. S., de Maeseneer, J., Greß, S., Heinemann, S., Rotar-Pavlic, D., Seghieri, C., Švab, I., van den Berg, M. J., Vainieri, M., Westert, G. P., Willems, S., & Groenewegen, P. P. (2011). QUALICOPC, a multi-country study evaluating quality, costs and equity in primary care. *BMC Family Practice*, 12, 1–9. <https://doi.org/10.1186/1471-2296-12-115>
- Schmiedeberg, C., Castiglioni, L., & Schröder, J. (2016). Secondary respondent consent in the German family panel. *Bulletin of Sociological Methodology/ Bulletin De Méthodologie Sociologique*, 131(1), 66–77. <https://doi.org/10.1177/0759106316642707>

- Schneider, N. F., Bujard, M., Wolf, C., Gummer, T., Hank, K., & Neyer, F. J. (2021). Family research and demographic analysis (FReDA): Evolution, framework, objectives, and design of “The German Family Demography Panel Study”. *Comparative Population Studies*, 46. <https://doi.org/10.12765/CPoS-2021-06>
- Schröder, J., Castiglioni, L., Brüderl, J., & Krieger, U. (2012). The influence of relationship quality on the participation of secondary respondents: Results from the German family panel. *Comparative Population Studies*, 37(3–4). <https://doi.org/10.12765/CPoS-2012-07>
- Schröder, J., Schmiedeberg, C., & Castiglioni, L. (2016). The effect of interviewers’ motivation and attitudes on respondents’ consent to contact secondary respondents in a multi-actor design. *Survey Methods: Insights from the Field*, 14. <https://doi.org/10.13094/SMIF-2016-00005>
- Singh, S., & Wassenaar, D. R. (2016). Contextualising the role of the gatekeeper in social science research. *South African Journal of Bioethics and Law*, 9(1), 42–46.
- Slauson-Blevins, K., & Johnson, K. M. (2016). Doing gender, doing surveys? Women’s gatekeeping and men’s non-participation in multi-actor reproductive surveys. *Sociological Inquiry*, 86(3), 427–449. <https://doi.org/10.1111/soin.12122>
- Stroud, N. J., & Kenski, K. (2007). From agenda setting to refusal setting: Survey nonresponse as a function of media coverage across the 2004 election cycle. *Public Opinion Quarterly*, 71(4), 539–559. <https://doi.org/10.1093/poq/nfm042>
- Sweet, J. A., & Bumpass, L. L. (1996). *The national survey of families and households-Waves 1 and 2: Data description and documentation*. Center for Demography and Ecology, University of Wisconsin-Madison.
- Tolonen, H., Helakorpi, S., Talala, K., Helasoja, V., Martelin, T., & Prättälä, R. (2006). 25-year trends and socio-demographic differences in response rates: Finnish adult health

- behaviour survey. *European Journal of Epidemiology*, 21(6), 409–415.
<https://doi.org/10.1007/s10654-006-9019-8>
- Tourangeau, R. (2014). *Hard-to-survey populations*. Cambridge University Press.
- Watson, N., & Wooden, M. (2009). Identifying factors affecting longitudinal survey response. In P. Lynn (Ed.), *Methodology of longitudinal surveys* (pp. 157–181). Wiley.
- Williams, D., & Brick, J. M. (2018). Trends in US face-to-face household survey nonresponse and level of effort. *Journal of Survey Statistics and Methodology*, 6(2), 186–211.
<https://doi.org/10.1093/jssam/smx019>
- Wittenborn, A. K., Dolbin-MacNab, M. L., & Keiley, M. K. (2013). Dyadic research in marriage and family therapy: Methodological considerations. *Journal of Marital and Family Therapy*, 39(1), 5–16. <https://doi.org/10.1111/j.1752-0606.2012.00306.x>
- Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (2021). Conducting General Social Surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*, 85(2), 623–648. <https://doi.org/10.1093/poq/nfab039>
- Young, R., & Johnson, D. (2013). Methods for handling missing secondary respondent data. *Journal of Marriage and Family*, 75(1), 221–234. <https://doi.org/10.1111/j.1741-3737.2012.01021.x>

Appendix

Table A3.1: Summary statistics for primary respondents regarding items under study

	Total	No consent	Consent
<i>Individual level</i>			
Education			
Non-tertiary	46.9	50.5	44.0
Tertiary	53.1	49.5	56.0
Migration background			
No	86.7	86.6	86.8
Yes	13.3	13.4	13.2
Gender			
Male	42.9	37.4	47.3
Female	57.1	62.6	52.7
Parents			
No	45.1	43.8	46.5
Yes	54.9	56.6	53.5
<i>Dyadic level</i>			
Cohabiting			
No	16.4	21.2	12.5
Yes	83.6	78.8	87.5
Married			
No	45.5	47.2	44.1
Yes	54.5	52.8	55.9
Relationship satisfaction			
0-10	8.26 (1.70)	8.10 (1.82)	8.40 (1.59)
Relationship satisfaction (binary)			
Low	46.8	50.3	43.9
High	53.2	49.7	56.1
Observations	15,182	6,779	8,403
In percent	100,0	44,6	55,4

Source: FReDA Release v.3.0.0.

Note: Standard deviation for means in parentheses. Distribution for categorical variables in percent (%) and means for metric variables. PR=primary respondent; SR=secondary respondent.

Table A3.2: Summary statistics for secondary respondents regarding items under study

	Total	No interview	Interview
<i>Individual level</i>			
Education			
Non-tertiary	45.6	51.1	44.2
Tertiary	54.4	48.9	55.8
Migration background			
No	85.0	81.6	85.8
Yes	15.0	18.4	14.2
Gender			
Male	52.9	59.2	51.3
Female	47.1	40.8	48.7
Parents			
No	46.4	48.5	45.9
Yes	53.6	51.5	54.1
<i>Dyadic level</i>			
Cohabiting			
No	12.5	20.8	10.4
Yes	87.5	79.2	89.6
Married			
No	44.1	52.5	42.0
Yes	55.9	47.5	58.0
Relationship satisfaction			
0-10	8.39 (1.59)	8.15 (1.75)	8.46 (1.54)
Relationship satisfaction (binary)			
Low	43.9	50.4	42.2
High	56.1	49.6	57.8
<i>Contact level</i>			
Invitation			
Sent to SR	76.8	65.7	79.6
Sent to PR	23.2	34.3	20.4
Tranche			
First	82.5	73.6	84.7
Second	14.8	21.1	13.2
Third	2.7	5.3	2.1
Observations	8,414	1,695	6,719
In percent	100.0	20.1	79.9

Source: FReDA Release v.3.0.0.

Note: Standard deviation for means in parentheses. Distribution for categorical variables in percent (%) and means for metric variables. PR=primary respondent; SR=secondary respondent.

infas, Postfach 240101, 53154 Bonn

7273/LFD

Anrede
Name
Anschrift
PLZ Ort

Bonn, im Juli 2021

Einladung zur wissenschaftlichen Studie „FReDA – Beziehungen und Familienleben in Deutschland“

Sehr geehrte/r Frau/Herr <Name>,

dieses Jahr ist die wissenschaftliche Studie „FReDA – Beziehungen und Familienleben in Deutschland“ gestartet. Im Rahmen der Studie haben wir vor Kurzem Ihren Partner bzw. Ihre Partnerin befragt. Ihr Partner bzw. Ihre Partnerin hat uns das Einverständnis gegeben, dass wir auch Sie im Rahmen der Studie kontaktieren dürfen. Daher bitten wir nun auch Sie, sich an unserer kurzen Befragung zu beteiligen.

Ihre Meinung ist für uns wichtig

Viele wissenschaftliche Fragen zu dem Thema Partnerschaft lassen sich nur untersuchen, wenn man beide Perspektiven einer Beziehung kennt. Daher möchten wir gerne auch Ihre Erfahrungen und Einstellungen kennenlernen. Im Mittelpunkt dieser Befragung stehen Themen wie Partnerschaft, Einstellungen und Werte, Arbeitsteilung und Haushalt sowie Kinderwunsch und Erziehung. Dazu möchten wir Sie bitten, unseren kurzen Fragebogen zu beantworten. Die Teilnahme ist selbstverständlich freiwillig. Um mehr über die Lebenssituationen von Paaren und das Miteinander in Partnerschaften und Familien zu erfahren, sind wir aber auf Ihre Unterstützung angewiesen.

Als Dankeschön für Ihre Unterstützung erhalten Sie **5 Euro**, die wir diesem Brief bereits beigelegt haben. Wir bitten Sie herzlich, an dieser wichtigen Studie mitzuwirken.

So können Sie teilnehmen

Sie erreichen die Online-Befragung unter <urlcawi>
Geben Sie dort bitte Ihren Zugangscode ein: <pwdblock>

Mit dem rechts abgedruckten QR-Code erreichen Sie die Befragung direkt, ohne Eingabe Ihres persönlichen Zugangscode. Alternativ zu dem Online-Fragebogen können Sie den beiliegenden schriftlichen Fragebogen ausfüllen und in dem beigefügten portofreien Rückumschlag an uns zurückschicken.

infas

infas Institut für angewandte
Sozialwissenschaft GmbH

Postfach 240101
53154 Bonn
kostenfreie Tel. 0800/7384-500
freda@infas.de
www.freda-studie.de

 Bundesinstitut
für Bevölkerungsforschung

gesis
Leibniz-Institut
für Sozialwissenschaften

GEFÖRDERT VON
 Bundesministerium
für Bildung
und Forschung

ERLEBTE VON
 Bundesministerium
des Innern, für Bau
und Heimat

QR-Code
deutsch
<qrcawi>

7273/W2/P/EA/2021

Figure A3.1: Invitation letter for secondary respondents

Was ist die FReDA-Studie

Die Studie „FReDA – Beziehungen und Familienleben in Deutschland“ wird gemeinsam vom infas Institut für angewandte Sozialwissenschaft GmbH, vom GESIS – Leibniz-Institut für Sozialwissenschaften und vom Bundesinstitut für Bevölkerungsforschung (BiB) durchgeführt. Weitere Informationen zur Studie finden Sie auf der Webseite www.freda-studie.de.

Warum gerade Sie?

Im Rahmen der Studie haben wir Ihren Partner bzw. Ihre Partnerin befragt. Ihr Partner bzw. Ihre Partnerin hat bei der Teilnahme angegeben, dass wir Sie gerne kontaktieren können. Nun laden wir auch Sie zu unserer Studie ein. Ihre persönliche Teilnahme ist für uns sehr wichtig, weil wir viele wissenschaftliche Fragen erst dann vollständig untersuchen können, wenn wir beide Perspektiven einer Beziehung kennen. Ihre Meinung kann also durch keine andere ersetzt werden! Dennoch ist die Teilnahme an der Befragung natürlich freiwillig. Alle Regeln des Datenschutzes werden eingehalten. Die Daten werden anonym ausgewertet – niemand wird hinterher sehen können, was Sie persönlich geantwortet haben. Weitere Informationen zum Datenschutz entnehmen Sie bitte der Datenschutzerklärung.

Wo erhalten Sie weitere Informationen?

Bei Fragen zur Teilnahme an der Studie können Sie sich gerne zu den üblichen Bürozeiten unter der kostenfreien Telefonnummer **0800/7384-500** an uns wenden. Nutzen Sie bei Rückfragen auch unsere E-Mail-Adresse freda@infas.de und wir werden Ihnen umgehend antworten. Besuchen Sie für weitere Informationen zu FReDA auch gerne die Homepage unter www.freda-studie.de.

Machen Sie mit und sagen Sie uns, was Sie denken!

Wir bedanken uns schon heute sehr herzlich für Ihre Mitwirkung und verbleiben mit freundlichen Grüßen

Menno Smid
Geschäftsführer
infas Institut für angewandte
Sozialwissenschaft

Prof. Dr. Norbert F. Schneider
Direktor
Bundesinstitut für
Bevölkerungsforschung

Prof. Dr. Christof Wolf
Präsident
GESIS – Leibniz-Institut für
Sozialwissenschaften

Seite 2

Source: FReDA Release v.3.0.0.

Note: Retrieved from infas Methods Report (2023)

Figure A3.1: Invitation letter for secondary respondents (*continued*)

Um mehr über die Lebenssituation von Paaren und das Miteinander in Partnerschaften und Familien zu erfahren, würden wir gerne auch Ihren Partner oder Ihre Partnerin bitten, einen kurzen Fragebogen auszufüllen.

Der Fragebogen für Ihren Partner oder Ihre Partnerin enthält ähnliche Fragen wie in dieser Befragung. Als kleines Dankeschön erhält Ihr Partner oder Ihre Partnerin 5 Euro.

Die Teilnahme Ihres Partners oder Ihrer Partnerin an der Befragung ist für den Erfolg dieser Studie von großer Bedeutung. Die Teilnahme ist selbstverständlich freiwillig. Für die Befragung gelten alle Regelungen des Datenschutzes, die wir strengstens einhalten.

Sind Sie damit einverstanden, dass wir Ihrem Partner oder Ihrer Partnerin einen Fragebogen zusenden?

Ja ☐

Nein ☐

Damit wir Ihrem Partner oder Ihrer Partnerin den Fragebogen zusenden können, geben Sie bitte die vollständigen Kontaktdaten des Partners oder der Partnerin an:

 *Bitte füllen Sie die Adressfelder gut leserlich aus, am besten in Großbuchstaben.*

Wir versichern Ihnen, dass diese Seite mit den Kontaktdaten bei infas sofort nach Eingang vom Fragebogen getrennt wird. Die Angaben aus dem Fragebogen werden niemals zusammen mit ihren Kontaktdaten gespeichert.

Vorname	Nachname
Partner/in wohnt bei mir im selben Haushalt. ☐	
Straße und Hausnummer 	
PLZ	Ort
E-Mail-Adresse 	

Source: FReDA Release 3.0.0.

Note: Retrieved from Partner Questionnaire (PAPI), Wave W1A (2021).

Figure A3.2: Question text for obtaining consent and contact information to interview a secondary respondent

Table A3.3: Descriptive distributions between self-reports by secondary respondents and proxy information by primary respondents

	Proxy information by PR	Self-report by SR
<i>Individual level</i>		
Occupational degree¹		
Non-tertiary	53.9	53.2
Tertiary	46.1	46.8
Country of birth²		
Germany	88.9	89.0
Other	11.1	11.1
Gender		
Male	51.3	50.8
Female	48.7	49.2
Parents		
No	45.2	44.2
Yes	54.8	55.8
<i>Dyadic level</i>		
Cohabiting		
No	10.4	10.2
Yes	89.6	89.8
Married³		
No	n.a.	n.a.
Yes		
Relationship satisfaction		
0-10	8.46 (1.54)	8.47 (1.53)
Relationship satisfaction (binary)		
Low	42.2	43.5
High	57.8	56.5
Observations	6,686	
In percent	100.0	

Source: FReDA Release v.3.0.0.

Note: Standard deviation for means in parentheses. Distribution for categorical variables in percent (%) and means for metric variables. ¹Highest occupational degree is used as an approximation for education due to limited items in the questionnaire for secondary respondents. ²Country of birth is used as an approximation for migration background due to limited items in the questionnaire for secondary respondents. ³Information about marital status is not asked in the questionnaire for secondary respondents. PR=primary respondent; SR=secondary respondent; n.a.=information not available.

Table A3.4: Average marginal effects on the likelihood of providing consent by primary respondents, separated between CAWI and PAPI interviews

	CAWI		PAPI	
	AME	SE	AME	SE
<i>Individual level</i>				
Education (Ref.: Tertiary)				
Non-tertiary	-0.038***	(0.009)	-0.036	(0.022)
Migration background (Ref.: None)				
Yes	-0.006	(0.012)	-0.044	(0.033)
Gender (Ref.: Female)				
Male	0.080***	(0.008)	0.127***	(0.022)
Parental status (Ref.: Parent)				
Childless	0.069***	(0.011)	0.098***	(0.026)
<i>Dyadic level</i>				
Cohabitation status (Ref.: Cohabiting)				
Not cohabiting	-0.183***	(0.013)	-0.206***	(0.028)
Marital status (Ref.: Married)				
Not married	-0.001	(0.011)	-0.024	(0.027)
Relationship satisfaction (0-10)	0.023***	(0.002)	0.015*	(0.006)
Observations	13,237		1,945	
AIC	17593.62		2511.92	
BIC	17653.55		2556.50	

Source: FReDA Release v.3.0.0.

Note: Average marginal effects (AME) from logistic regression model with robust standard errors (SE) in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001.

Table A3.5: Average marginal effects on the likelihood of providing consent by primary respondents

	AME	SE
<i>Individual level PR</i>		
Education (Ref.: Tertiary)		
Non-tertiary	-0.048***	(0.008)
Migration background (Ref.: None)		
Yes	-0.008	(0.012)
Gender (Ref.: Female)		
Male	0.093***	(0.008)
Parental status (Ref.: Parent)		
Childless	0.079***	(0.010)
<i>Dyadic level</i>		
Cohabitation status (Ref.: Cohabiting)		
Not cohabiting	-0.182***	(0.012)
Marital status (Ref.: Married)		
Not married	-0.006	(0.010)
Relationship satisfaction (0-10)	0.022***	(0.002)
Observations	15,182	
AIC	20302.26	
BIC	20363.28	

Source: FReDA Release v.3.0.0.

Note: Average marginal effects (AME) from logistic regression model with robust standard errors (SE) in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001.

Table A3.6: Average marginal effects on the likelihood of participating by secondary respondents

	AME	SE
<i>Individual level SR</i>		
Education (Ref.: Tertiary)		
Non-tertiary	-0.033***	(0.009)
Migration background (Ref.: None)		
Yes	-0.053***	(0.013)
Gender (Ref.: Female)		
Male	-0.050***	(0.008)
Parental status (Ref.: Parent)		
Childless	0.032**	(0.012)
<i>Dyadic level</i>		
Cohabitation status (Ref.: Cohabiting)		
Not cohabiting	-0.119***	(0.016)
Marital status (Ref.: Married)		
Not married	-0.049***	(0.011)
Relationship satisfaction (0-10)	0.015***	(0.002)
<i>Contact level</i>		
Invitation (Ref.: Sent to SR)		
Sent to PR	-0.119***	(0.011)
Tranche (Ref.: First)		
Second	-0.108***	(0.013)
Third	-0.201***	(0.031)
Observations	8,414	
AIC	7971.62	
BIC	8049.03	

Source: FReDA Release v.3.0.0.

Note: Average marginal effects (AME) from logistic regression model with robust standard errors (SE) in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001.

Table A3.7: Average marginal effects on the likelihood of participating by secondary respondents, including only characteristics at the individual and dyadic level

	AME	SE
<i>Individual level SR</i>		
Education (Ref.: Tertiary)		
Non-tertiary	-0.031***	(0.009)
Migration background (Ref.: None)		
Yes	-0.059***	(0.013)
Gender (Ref.: Female)		
Male	-0.051***	(0.009)
Parental status (Ref.: Parent)		
Childless	-0.041***	(0.011)
<i>Dyadic level</i>		
Cohabitation status (Ref.: Cohabiting)		
Not cohabiting	-0.130***	(0.017)
Marital status (Ref.: Married)		
Not married	-0.054***	(0.012)
Relationship satisfaction (0-10)	0.016***	(0.003)
Observations	8,414	
AIC	8209.65	
BIC	8265.95	

Source: FReDA Release v.3.0.0.

Note: Average marginal effects (AME) from logistic regression model with robust standard errors (SE) in parentheses. *p < 0.05; **p < 0.01; ***p < 0.001.

4 Better Ask Online When It Concerns Intimate Relationships?

Survey Mode Differences in the Assessment of Relationship Quality¹⁵

Abstract

Background: The assessment of relationship quality is a key construct in family research and relies on several indicators. As answer behavior for sensitive and subjective questions can be biased by the interview situation, the emerging switch from face-to-face to web mode or mixed-mode designs in surveys challenges the comparability of measurements.

Objective: This study investigates the impact of two modes of data-collection, face-to-face and web mode, on central measurements of relationship quality in quantitative family research.

Methods: In a German experimental pilot study (2018)¹⁶ within the Generations and Gender Programme (GGP), target persons were randomly assigned to conduct the interview face-to-face or online. Mode differences are assessed by comparing distributions for various indicators of relationship quality. To adjust for confounders, post-stratification weighting and multivariate regression analysis are applied.

Results: Findings reveal consistent mode effects for almost all indicators of relationship quality even after adjusting for confounders. Respondents in web mode assess their relationship quality substantially lower than respondents in face-to-face mode in thinking more often about breaking up and in reporting lower satisfaction and more conflicts.

¹⁵ This study was conducted in cooperation with Dr. Detlev Lück and has been published in *Demographic Research* (see Schumann & Lück, 2023), DOI: 10.4054/DemRes.2023.48.22

¹⁶ This study used data from the GGP Pilot Study, funded by the Horizon 2020 Research and Innovation Programme under grant agreement n° 739511 for the project “Generations and Gender Programme: Evaluate, Plan, Initiate”, and by the Federal Institute for Population Research (BiB) in Wiesbaden, Germany, in order to co-finance the German data collection.

Conclusion: Web mode seems to support less socially desirable reflections on respondents' relationships, compared to face-to-face mode. Family researchers should be sensitized to consider survey design decisions when evaluating intimate relationships, particularly in longitudinal and cross-national studies.

Contribution: Findings on the assessment of relationships in family research based on self-administered modes, such as web, can be considered as more reliable than those based on interviewer-administered modes.

4.1 Introduction

The assessment of relationship quality is one of the most frequently addressed topics in research on intimate relationships (see Bradbury et al., 2000; Fincham & Beach, 2006; Karney & Bradbury, 2020). Relationship quality is connected with the stability of relationships and therefore often serves as a predictor for processes such as separations, family formations or marriages (Karney & Bradbury, 1995; Lewis & Spanier, 1979). Not only family sociologists and demographers, but also family psychologists frequently use indicators for relationship quality as determinants for various outcomes. Central and frequently analyzed indicators for relationship quality in quantitative family research are subjectively perceived stability (e.g., van Damme & Dykstra, 2018; Wiik et al., 2012), satisfaction with the relationship (e.g., Arránz Becker, 2013; Schmid et al., 2021) as well as certain interactions between partners such as conflict behavior (e.g., Huss & Pollmann-Schult, 2020; Kluwer & Johnson, 2007).

Most of the studies mentioned compare aspects of relationship quality between different countries and cultural backgrounds or between different points in time or life-course phases (e.g., Huss & Pollmann-Schult, 2020; Schmid et al., 2021; Wiik et al., 2012). Whenever different data sources are used for longitudinal or cross-national studies, analyses strongly depend on a consistently high data quality and the comparability of these data. Limitations on reliability and comparability can have many causes. A particularly important determinant is the mode of data collection. Face-to-face interviews have for decades been the most common mode of data collection for large-scale survey programs in the landscape of social science and family research, mainly because of their comparably high response rate and good coverage to achieve population representative samples (Groves et al., 2004; De Leeuw et al., 2008). Notwithstanding, one of the known and well researched downsides of face-to-face interviews is that the personal

interview situation supports the underreporting of sensitive topics, such as illicit or sexual behavior (Aquilino, 1991; Tourangeau & Smith, 1996).

However, this conventional wisdom is currently undergoing a reassessment. For data collectors, the switch from traditional personal interviews to online interviews is becoming increasingly attractive because web interviews are much more cost efficient than face-to-face interviews, especially in countries where labor costs for interviewers are high (Bethlehem & Biffignandi, 2012). Additionally, they facilitate rapid data collection and delivery (Couper, 2011). Moreover, the ubiquity of mobile phones and smartphones is leading to an increasing use of mobile devices to complete web surveys (Gummer et al., 2023), enabling respondents an easier access to the survey, e.g. via QR codes, and allowing to answer while carrying out their daily lives at any time or place they want to (Couper et al., 2017). Last but not least, the COVID-19 pandemic strongly accelerated this transition by forcing established face-to-face studies, such as the Generations and Gender Survey (GGS) or the German family panel pairfam, to switch to web interviews (Gummer et al., 2020), so that we may currently be witnessing the establishment of the self-administered online interview as a new standard mode of data-collection at least in Western Europe and in other countries with high labor costs and appropriate sampling frames.

In view of this development, the potential impact of the mode of data collection on data quality and on substantive analyses has become an even more relevant question for empirical analysis. Therefore, we will examine two survey modes that mark the starting and ending points of the transition described above: face-to-face interviews and web interviews. These two modes show the greatest difference in interviewer involvement: Web surveys are self-administered whereas face-to-face surveys are interviewer-administered. This comes with advantages as well as disadvantages for both modes and with positive as well as negative effects on data quality. On the one hand, an interviewer is able to motivate people to participate and thereby increase response rate (Groves et al., 2004) and the interviewer can support the respondent with questions

requiring a high cognitive effort (Holbrook et al., 2003). On the other hand, the presence of an interviewer increases the normative pressure on respondents to provide a socially acceptable answer, whereas the anonymous interview environment of self-administered interviews allows them to be more honest (Tourangeau & Smith, 1996; Tourangeau & Yan, 2007). Especially with regard to more sensitive questions, the mode of data collection has a strong influence on biases due to social desirability (Chang & Krosnick, 2010; Tourangeau & Yan, 2007). Research has shown that factual measurements of sociodemographic characteristics are less affected by the interview situation, whereas subjective and private questions, that score higher on sensitivity, elicit a stronger mode effect depending on whether they are self-administered or interviewer-administered (Burkill et al., 2016; Christensen et al., 2013). In terms of contents that are important for family research, previous findings on social desirability bias between modes have often concentrated on traditional or obvious sensitive items, that are strongly normative, like attitudes towards gender and family roles (Liu, 2017; Liu & Wang, 2016), questions regarding sexual behavior and sexual experiences (Burkill et al., 2016; Kelly et al., 2013), and questions regarding the assessment of mental and physical health (Braekman et al., 2020; Christensen et al., 2013).

The question remains open to what extent other kinds of subjective questions are perceived as sensitive by respondents, might therefore be prone to biases based on social desirability, and thus might also be subject to effects based on the mode of data collection. The research field of relationship quality provides a very good example for analyzing this question. First of all, it is a frequently addressed topic in research about intimate relationships with high relevance for other family-related events, such as childbearing (Rijken & Liefbroer, 2009), union dissolution (Karney & Bradbury, 1995), and might also affect matters in personal life, e.g. the own well-being (Gustavson et al., 2014). Second, the experimental study we use has many suitable indicators for measuring relationship quality comprehensively. These indicators cover a broad

range of domains in a relationship as we study measurements of perceived relationship stability, satisfaction with different aspects in couples' daily life as well as different areas of conflicts and different levels of aggressive and violating conflict styles. Furthermore, these indicators provide good examples of the subjective perceptions for which we lack clarification to what extent they must be considered as sensitive questions. One can assume that the normative expectation regarding the maintenance of a happy relationship in studies about intimate relationships could be strong, which would in turn increase respondents' perception of social pressure in a personal interview situation. The measurement of relationship quality requires subjective assessments by the respondent and may also be perceived as sensitive depending on the individual situation. This may at least be true for some of the various aspects of relationship quality for which we find indicators in our dataset. The broad spectrum of domains of relationship quality covered by the data may even provide a nuanced picture. Therefore, we will investigate whether measurements of this construct differ between a self-administered and an interviewer-administered interview situation.

For assessing differences between the modes of conduction, an experimental pilot study was carried out in Germany within the Generations and Gender Programme (GGP), which compares the traditional face-to-face mode with the upcoming web mode. The GGP is a well-established, large-scale survey program in family research. We profit from a unique experimental setting which allows us to use an existing survey instrument, the GGS, as well as an experimental design to test for differences in the mode of conduction. The aims of this study are to examine whether differences in measurements on frequently used indicators of relationship quality occur between face-to-face and web modes and to assess which survey design provides the most reliable measurement of relationship quality. Our research question is therefore: Do measurements of indicators frequently used for explaining relationship quality conducted in face-to-face mode differ from measurements conducted in web mode? In a first descriptive step, we

compare the mode-specific differences between the distributions of the particular indicators for relationship quality in the two experimental groups. In a second step, we estimate multivariate regression models to adjust for family-demographic confounders to assess the impact of mode on the particular items of relationship quality. Given that in each mode persons with specific characteristics might be more or less likely to participate, the regression models allow us to control for such selective confounders.

In the context of continuing methodological innovations and developments in data collection, this study should sensitize data users to the possibility of distortions for frequently used key variables in substantive analysis due to survey mode decisions. The findings are especially relevant for data analysis as well as for data conduction of cross-national and panel surveys based on different modes of data collection.

4.2 Background and expectations

4.2.1 Face-to-face versus web interviewing

Face-to-face and web surveys differ mostly regarding the degree of interviewer involvement. According to Couper (2011), this has an impact on overall participation in a survey, thus on the response rate and data quality. A meta-analysis revealed that response rates of web surveys are lower compared to other traditional modes of data collection, like face-to-face surveys (Daike-ler et al., 2020). But the last years have also shown that, at least in Western European countries, the response rates of data-collections in face-to-face mode have declined rapidly (Beullens et al., 2018). Nevertheless, interviewers can be helpful to motivate the target person in the recruitment stage to participate. Furthermore, web surveys obviously bear the risk of underrepresenting the offline population (Schonlau et al., 2009). This is particularly relevant for surveys that have to rely on nonprobability samples, e.g. due to a lack of a suitable sampling frame for the

particular target population, which lowers the representativeness needed for large-scale social science surveys (Tourangeau, 2017). Moreover, the interviewer can play a helpful role during the interview by assisting the respondent in the response process. Interviewers are able to support the respondent in answering questions requiring a high cognitive effort. They can motivate and support the respondent verbally as well as through non-verbal communication throughout a long interview (Holbrook et al., 2003). Some studies showed that self-administered web surveys had higher proportions of item nonresponse, higher proportions of choosing “Don’t know” answers and less differentiation on rating scales than in face-to-face surveys (Heerwegh, 2009; Heerwegh & Loosveldt, 2008).

However, data collectors have to consider that the use of interviewers is considerably more expensive than conducting a web interview. Additionally, there are also positive effects resulting from the absence of an interviewer on the data quality. While in face-to-face interviews, interviewers have the locus of control over the whole interview process, in web interviews, respondents have the autonomy to answer the questionnaire at the time and place they prefer, at the speed that suits them best, and with having the option to stop in between the interview and to continue later on (Couper, 2011). Furthermore, web interviews are characterized by a higher degree of privacy and anonymity than face-to-face interviews. This may be expected to modulate the strength of social desirability effects, as discussed in detail below.

4.2.2 Social desirability bias and relationship quality

Social desirability explains most prominently why the mode of data collection plays such an important role for answering questions prone to sensitivity. According to the concept of social desirability, respondents tend to overreport socially desirable answers and underreport socially undesirable answers (Callegaro, 2008). An open question is how strongly particular questions

are affected by social desirability bias. Often, the strength of social desirability bias corresponds with the degree of sensitivity of a question (Krumpal, 2013). But the perceived sensitivity of a question depends strongly on the person who is interviewed and thus also on his or her individual situation and how much emotional stress the respondent would endure by giving an honest answer. This can further vary along cultural, social, and situational contexts (Lee & Renzetti, 1990). Given that these factors vary across and even within studies due to different questionnaire contents and target populations, it is difficult to draw general conclusions about social desirability bias.

Taking the mode of conducting surveys into account, research has shown that respondents tend to answer more truthfully and honestly in an anonymous interview situation, especially for obviously sensitive questions (Chang & Krosnick, 2010; Krumpal, 2013; Tourangeau & Smith, 1996; Tourangeau & Yan, 2007). In other words, interviewer-administered modes, such as face-to-face, lead more often to socially desirable responding because respondents tend to present themselves in a socially favorable manner instead of reflecting the true situation in front of the interviewer. Whereas in self-administered interviews, like web surveys, respondents have a higher level of privacy and a lower level of perceived social pressure and show more honest answer behavior (Heerwegh, 2009). Interviewer characteristics, such as gender or ethnicity, can also affect interview dynamics and impact responses in personal interviews, particularly for questions prone to social desirability biases and for questions related to these characteristics (Davis et al., 2010). Research shows e.g., that the gender of the interviewer can influence response behavior regarding marriage-related questions (Liu & Stainback, 2013). Moreover, the degree of familiarity of the respondent and the interviewer might also impact the respondent's effort for answering sensitive questions as respondents show lower levels of trust and disclosure when the interviewer is a stranger and is not familiar with the local environment (Weinreb et al., 2018).

In order to reduce bias related to social desirability in face-to-face interviews, highly sensitive questions are often surveyed in form of a so-called “computer-assisted self-interview” (CASI) module where the respondent can complete individual question blocks independently. Even though this is a good way to make the respondents’ answers more anonymous, the control still remains with the interviewer. By contrast, in an entirely self-administered mode, such as web, the respondent can show a higher degree on self-disclosure because there is no other person present and therefore no time pressure and a free choice of where to respond to the interview. Additionally, face-to-face interviews are mostly conducted in the respondent’s own household, which means respondents might be influenced not only by the interviewer but also by any third persons present including for example partners, spouses, or children (Schröder & Schmiedeberg, 2023). It might be the case that these so-called bystander effects have an impact on answering questions in CAPI as well as in CASI during a face-to-face interview; research shows that the reporting of less desirable answers in web mode has the highest response accuracy compared to other self-administered modes (Kreuter et al., 2008).

Previous research on mode-related social desirability bias in studies about families and intimate relationships concentrates on a few subjective and objective indicators, but, as far as we know, there has been no study that investigated the impact of social desirability bias on items about relationship quality. An early experiment from the National Survey of Family Growth discloses a higher reported number of abortions in self-reports than in interviewer-administered interviews (Fu et al., 1998). This is one example of highly sensitive information in family research. A more recent experiment of the third British National Survey of Sexual Attitudes and Lifestyles examines the changes of responses from the same respondents first conducted in computer-assisted personal interviews (CAPI) and CASI and afterwards conducted in web mode. Their findings show that not all sensitive questions regarding sexual life revealed a mode effect between CAPI, CASI, and web. But for some questions regarding individual behavior, such as

same-sex experiences, and opinion questions, such as sexual satisfaction, they find a higher level of self-disclosure and more socially undesirable answers in web interviews compared to CAPI and CASI (Burkill et al., 2016), which implies that even a switch to CASI mode cannot fully compensate for the downsides of conducting personal interviews. The anonymous interview situation can play a major role here, as Robertson et al. (2018) find that respondents in online surveys report the highest comfort level in answering questions about non-heterosexual prevalence compared to sixteen other interviewer-administered as well as self-administered survey modes. An explanation is that online interviews are perceived as less intrusive and as having a higher level of anonymity and privacy without creating the feeling of being observed or recorded (Robertson et al., 2018). In the field of public health research, experiments come to similar findings regarding opinion questions and subjective assessments of personal well-being and health status. Answers to factual questions are comparable between face-to-face and self-administered survey modes, but they detect different levels of mode effects in answering more sensitive questions involving subjective assessments (Braekman et al., 2020; Christensen et al., 2013).

As prior research shows, mode effects can vary strongly between studies, because every survey focuses on different topics, uses different questionnaires, and aims at different target populations. This means effects between face-to-face and web surveys regarding biases due to social desirability are hard to generalize (Couper, 2011). Therefore, it is necessary to evaluate such effects always in the context of the particular study. What also can come into play in surveys about intimate relationships is a social expectation about how a relationship should ideally be. In such surveys, respondents might feel even more embarrassed to admit that their own situation does not conform to a norm or expectation about happy relationships. This means that the perceived social pressure might lead to misreported feelings or subjective assessments (DeMaio, 1984). What underlines this assumption is also a selective trend in the general participation in

studies about families and intimate relationships: People with happier and closer relationships within a family are more likely to participate in such surveys (Kalmijn, 2023).

Items which assess relationship quality are subjective questions that require respondents to reflect on their behavior and feelings. Our indicators under study serve to assess the quality in an intimate relationship. However, they cover a broad range of domains in an intimate relationship in addressing different areas in couples' life, like household division or childcaring, feelings and doubts about the own relationship, and ways of dealing with conflict situations. These different topics can seem more or less sensitive for the respondent depending on the individual situation, especially when they touch a sore point in the own relationship. Due to the higher anonymity, we expect that web interviews support more open and probably more honest answer behavior. Further, the locus of control is up to the respondent, which may lead to a higher comfort level in answering probably unpleasant private questions. This means that the respondent can fill out the questionnaire at the time and place of his or her own choice, for example when he or she is alone at home so that interference by a third person, like the partner, can be avoided. Therefore, we expect a mode effect for indicators of relationship quality as follows: Respondents who participate in web mode provide more socially undesirable answers and higher levels of self-disclosure than face-to face respondents. This means that web-respondents should report a lower level of relationship quality and assess the relationship on average more negatively compared to respondents in a face-to-face interview. Nevertheless, it is an open question to what extent which indicators on relationship quality display such a mode effect and to which extent they must be considered as sensitive questions in the context of family research.

4.3 Data and methods

4.3.1 Experimental design and case selection

For answering our research question, we used data from an experimental pilot study within the framework of the Generations and Gender Programme (GGP) (Emery et al., 2018). The GGP is an international family demographic infrastructure which conducts the Generations and Gender Survey (GGS). The GGS is fielded in many European and a few non-European countries and is designed as a three-wave panel study. The study focuses on families, intimate relationships, and life course trajectories of individuals (Gauthier et al., 2018). Aside from two other countries, the GGP pilot study was conducted in Germany in 2018. The aim of this pilot study was to test whether a revised version of the GGS questionnaire as well as a new survey design work well in the field and whether the GGS can be conducted as a mixed-mode or online survey. A “push-to-web” design was applied and compared with the “traditional” GGS mode, which is face-to-face (Lutrig et al., 2022). Push-to-web design means that we conducted web interviews as we would have done in an entirely online survey but contacted the nonrespondents after the web fielding period again and asked them to participate in a personal interview. For our research question, we concentrate on the web respondents of that group. The German pilot study carried out further experiments regarding the timing and amount of incentives. As variation in incentives may affect data quality and response behavior regarding sensitive questions (Medway, 2012), we only compared groups that used identical incentives. Only in this way can we obtain an experimental setting that provides the same initial conditions for both groups, except the mode of conducting the interview, which serves as the treatment.

Respondents in the reference group participated in a CAPI interview, and respondents in the experimental group participated in a computer-assisted web interview (CAWI). Both, the face-to-face group and the web group, received the same Blaise-programmed GGS questionnaire in

terms of question wording, routing, and design. Further, both groups received a prepaid incentive worth five euros. The target persons, aged 18-49 years, were selected with simple random sampling from local registry offices (“Einwohnermeldeämter”) in the German federal state of Bavaria, with a quota of 50% of addresses coming from rural areas and 50% of addresses coming from urban areas. The target persons were randomly assigned to the experimental groups. The size of the gross sample was calculated by the fieldwork institute on the basis of the expected response rate per mode. The aim was to achieve at least 200 cases per experimental group. Based on experience with other German surveys, the gross sample size of the face-to-face group was set lower because the response rate for the face-to-face mode was expected to be higher than for the web mode.

Both groups received an invitation letter with the unconditional incentive in it. For the face-to-face group, the letter announced that an interviewer would come to the household to conduct the interview. For the web group, an URL with a password was provided in the letter, and target persons were asked to go online and fill out the questionnaire on their own. The web group received also two reminder letters, each two weeks after the previous letter. Table 4.1 gives an overview over the design specifications and the case selection. The overall response rate was calculated according to Response Rate 1 following the AAPOR classification of standard definitions (AAPOR, 2016). The response rate showed that we had a higher participation – nearly 10 percentage points – in the face-to-face group. Both response rates are rather low, but other German social science surveys conducted face-to-face yield similar response rates (Wolf et al., 2021).

Table 4.1: Overview of the experimental design and case selection

	Reference group	Experimental group
Mode of data collection	Face-to-face (CAPI)	Web (CAWI)
Country	Germany	Germany
Target population	18-49-year-olds	18-49-year-olds
Incentives	€5 pre-paid	€5 pre-paid
Maximum number of contacts	Invitation letter + five personal contact attempts	Invitation letter + two reminder letters
Gross sample (n)	685	1,365
Net sample (n)	193	261
Response rate (%)	29.5	19.4
Respondents with a partner (n)	146	197
Respondents with a partner (%)	76.0	77.6

Source: GGP Pilot Study 2018.

Note: n=observations; %=percent.

Our research question focuses on the assessment of relationship quality in couples, so we only consider respondents in our analyses who reported that they had a partner for at least three months. That is how the GGS measures partnership status. Only these respondents could actually answer questions regarding their current relationship. Table 4.1 gives an overview of the number of cases in the sample under study. It can be seen that the proportion of people in a relationship does not differ greatly between the two groups, with 76.0% of respondents having a partner in face-to-face and 77.6% of respondents having a partner in web mode. This corresponds closely to the proportion of persons with partners in other German studies about families and intimate relationships (Kantar Public, 2018). By comparing the distribution of family-demographic characteristics in the overall sample and our analytical sample, including only respondents with a partner, for each mode shown in Table 4.2, we see that slightly more women and fewer younger people have a partner. However, these tendencies are evident for both modes and can therefore be ignored. Generally, respondents in the web sample – irrespectively of having a partner – are higher educated and more likely to live in urban areas compared to face-to-

face respondents (see Table 4.2), which is consistent with existing research (e.g., Atkeson et al., 2014).

Table 4.2: Family-demographic distributions for respondents in the entire sample and for respondents with a partner

	Face-to-face		Web	
	Respondents (total)	Respondents with a partner	Respondents (total)	Respondents with a partner
Sex				
Male	47.92	42.47	47.24	45.69
Female	52.08	57.53	52.76	54.31
Education School				
Low	25.52	21.92	12.90	14.66
Middle	26.04	28.08	33.06	35.60
High	48.44	50.00	54.03	49.74
Age				
18-29	37.70	25.52	30.68	23.71
30-39	29.32	35.86	32.27	32.99
40-49	32.98	38.62	37.05	43.30
Citizenship				
German	88.02	87.67	88.54	87.76
Non-German	11.98	12.33	11.56	12.24
Regional setting				
Urban	39.58	41.10	54.33	51.78
Rural	60.42	58.90	45.67	48.22
Child under six years in household				
No	60.94	51.37	66.93	57.87
Yes	39.06	48.63	33.07	42.13
Relationship status				
Married	62.33	62.33	69.79	69.79
Cohabiting	24.66	24.66	17.19	17.19
Living apart together	13.01	13.01	13.02	13.02
Observations	192	146	254	197

Source: GGP Pilot Study 2018.

Note: Distributions in percent (%).

4.3.2 Methodological approach

The design of the experimental study allows us to compare answering patterns in face-to-face mode and web mode to assess the overall impact of one mode compared to the other. We

analyze 15 single items in the univariate analyses and six items in the multivariate regression analyses. All indicators under study relate to the construct of relationship quality.

In a first step, we apply a univariate approach and calculate means and proportions on item level to see how the mode affects point estimators. We test for mode-specific differences in the distributions with a two-sample t-test for mean differences and a Pearson- χ^2 -test for independence between categorical variables. Univariate comparisons should reveal initial ad-hoc findings regarding the extent of distortion due to social desirability and self-disclosure in one mode as compared to the other. Further, surveys often use ex-post weighting to adjust for certain biases due to mode-specific selectivity, nonresponse bias, or coverage bias (Groves et al., 2004). Hence, we additionally apply post-stratification weighting to evaluate whether the measurement equivalence between the univariate distributions of the two modes improves or not (Bethlehem & Stoop, 2007; Schonlau & Couper, 2017). For example, if more young people participate in web than in face-to-face mode, this may reduce the average duration of relationships in one mode, which might have a confounding effect on relationship quality. Therefore, we adjust the entire sample according to population totals for specific demographic characteristics that are available from official German statistics (“Census” and “Micro Census”) for our target population. As auxiliary variables for weighting we use sex, age groups, highest level of school education, nationality, and regional setting. Selective nonresponse can have many sources and might not be based exclusively on demographic characteristics (Schonlau et al., 2007), but we had to rely on the best information available from official statistics, which only includes demographic information for our target population. A detailed list and the sources of information used for post-stratification weighting can be found in Table A4.1 in the Appendix.

In a second step, we pool the experimental and the reference group and apply multivariate regression analysis with mode as the explaining variable and a block-wise adjustment of further confounding variables to test whether the effect of mode is robust or not. Given that

participation in surveys might be selective, regression analysis allows us to control for characteristics which correlate with selective participation in one mode. Indicators of relationship quality are treated as outcome variables and the mode of conducting the survey as a predictor variable. In the first baseline model we estimate the crude mode effect on the respective indicator of relationship quality. In the second model we include the same standard demographic variables that we used for post-stratification weighting, to adjust for selective participation. In the third model we include additionally family-related variables, which are often used as adjustment variables for relationship quality. We then examine whether a possible effect of the mode of data collection on the respective indicators of relationship quality changes between the models or whether the effect is robust after adjusting for further explanatory determinants. For the one binary dependent variable, which is the question whether or not the respondent had thought about breaking up, we calculate linear probability models (LPM) because it facilitates the interpretation of estimates, especially when comparing coefficients across differently specified models. As a robustness check, we further apply logistic regression models and estimate average marginal effects (AME) that yield similar results as with the linear probability approach (see Table A4.4). For the other indicators, we perform linear ordinary least squares (OLS) regression models and show the estimated coefficients.

4.3.3 Measurements

The items evaluated in our study are often used as predictors, mediators, and outcomes in substantive analyses in research about the quality of intimate relationships. They are included in many large-scale surveys about families and relationships, such as the GGS.

In the univariate approach we examine 15 indicators of relationship quality separately to get an impression of mode effects on a broad variety of items which potentially display different

effects. One item relates to subjective stability, three items relate to satisfaction, and the remaining 11 items to conflict frequencies and styles. Starting with subjective instability, we evaluate the question of whether the respondent has thought about breaking up with their partner. The binary indicator is coded with 0 “no” and 1 “yes.” Items on satisfaction are an often-used survey instrument to assess feelings and are mostly measured on point scales. The GGS questionnaire contains three satisfaction scales, which cover three different domains in intimate relationships. The question wording is, “How satisfied are you with the relationship in general, the division of household tasks, and the division of childcare tasks.” The answers have to be assessed separately on a 11-point scale, where 0 means very dissatisfied and 10 means very satisfied. A rating of 5 means medium satisfaction. The question on satisfaction with the division of household tasks was only asked to respondents who have a coresidential partner and the question regarding satisfaction with the division of childcare tasks was filtered for parents. All other questions on the various aspects of relationship quality, including those on conflicts, were asked to all respondents who have a partner, regardless of other criteria.

Whereas the first four indicators relate more to the level of feelings, the last indicators rather concern the behavior in a relationship, specifically conflict behavior within the couple, differentiated into frequency of conflicts and conflict styles. For the univariate approach, we use seven single items regarding the frequency of conflicts on the following issues: household chores, money, leisure time, relations with friends, relations with parents, having children, and child raising issues. The answer categories range from 1 (never) to 5 (very frequently). For the multivariate analysis, we generate one indicator for the frequency of conflicts within the couple in general, summarizing the information of these seven items, in order to reduce complexity. Given that each single item measures the frequency of certain conflicts on the same scale, we are able to directly compare their answer codes. Assuming that one relationship conflict will rarely touch several of the issues represented by the seven items at the same time and that the

conflicts measured by the seven items have little overlap, we consider the addition of answer codes as an appropriate way of constructing such an indicator for frequency of conflict. Accordingly, we generate an additive index (Cronbach's α : 0.70), which ranges from 1 (never reported conflicts) to 29 (very frequently reported conflicts). It is recoded such that the more conflicts reported, the higher the value of the index.

Finally, we look at four single items which cover reactions in conflict situations. These items assess how often respondents avoid discussions by giving in, discuss conflicts calmly, argue heatedly or get loud and refuse to talk. Here again, answers are coded from 1 (never) to 5 (very frequently). These conflict styles are also summarized into one indicator in order to reduce complexity for multivariate analyses. The indicator measures the tendency of choosing inadequate conflict behavior, according to the social norms of a late-modern society in which it is expected that disagreements are resolved by rational exchange of arguments. Accordingly, we recode the item "discuss conflicts calmly" reversely and construct an additive index summing up all four items, which are all recoded in a way that the more inadequate conflict behavior has been reported, the higher is the score (Cronbach's α : 0.54). We are aware that, in this case, we are summarizing information from more heterogeneous items, so that the validity of the generated indicator is lower. Although Cronbach's α of the second index has a lower internal consistency, the scale is sufficient for our purposes to get an additive measure of inadequate conflict behavior. The index ranges from 1 (only inadequate conflict behavior reported) to 17 (only adequate conflict behavior reported). The original wordings of all questions and answer categories can be found in Table A4.2.

As control variables for the multivariate approach, we use the same demographic indicators as for the post-stratification weighting. Sex of a respondent is coded as (0) "male" and (1) "female" and for nationality we distinguish between (0) "German citizenship" and (1) "non-German citizenship." Age is measured in years and ranges from 18 to 49. School education is measured

dichotomously (“highest school education (college entry qualification)” (1) and “lower/ middle school education or less” (0)). Information on the regional setting or community size was provided by the fieldwork institute and distinguishes areas where respondents live into (1) “urban areas” and (2) “rural areas.” These background variables are not only used frequently as standard demographic controls, they can also affect the participation in web mode (Vehovar et al., 2002). To control for family-related determinants which often correlate with relationship quality, we include the variables relationship status and the existence of co-resident children under six years in the household. We decide to choose an indicator for having children of younger ages because we assume that infants affect couples’ daily life stronger than elder children, as they need more care and attention, often at the expense of the relationship quality and time of the young parents. Additionally, parents of younger children have a lower level of mobility and of available time which can impact the participation in the respective survey mode. As a sensitivity check, we also calculate models using an indicator for having children of any age as control variable, instead, and find no differences in the identified mode effects. The measurement of relationship status is distinguished into respondents who are married with their current partner, irrespectively of cohabitation (“married” (1)), respondents, who live together with their partner without being married (“cohabiting” (2)) and respondents who have a partner, but are not living in cohabitation or in marriage with their partner (“living apart together” (3)), often referred to as long-distance relationships. The other indicator distinguishes between “at least one child under six years living most of the time in the same household with the respondent” (1) versus “no children under six years living in the same household of the respondent” (0). Unfortunately, information about the duration of the current relationship is unavailable in the GGS.

4.4 Results

4.4.1 Univariate analyses

We start with a look at the univariate distributions of the indicators of relationship quality in the two experimental groups to examine whether the mode of conduction affects point estimators under the two experimental conditions. The distributions for all 15 single indicators of interest are shown in Table 4.3. Respondents using the web report higher shares of social undesirable answers than face-to-face respondents for almost all items under study.

Table 4.3: Means of indicators or percentages of confirmative answers by mode of data collection with 95% confidence intervals

	F2F	Web	Mode difference	F2F	Web
	Mean or percentage		Δ	n	
Subjective instability¹					
Thought about breaking up (%)	8.97 (5.27-14.86)	17.13 (12.30-23.35)	8.16	145	181
Satisfaction²					
General relationship	9.11 (8.91-9.31)	8.63 (8.43-8.82)	-0.48	145	190
Household tasks	8.46 (8.19-8.73)	7.90 (7.59-8.20)	-0.56	127	166
Childcare tasks	8.66 (8.28-9.03)	8.11 (7.73-8.50)	-0.55	70	81
Conflict frequency³					
Household chores	2.31 (2.13-2.48)	2.45 (2.31-2.59)	0.14	146	196
Money	1.66 (1.53-1.79)	1.97 (1.83-2.12)	0.31	146	195
Leisure time	2.23 (2.07-2.38)	2.39 (2.26-2.51)	0.16	146	196
Relations with friends	1.50 (1.39-1.61)	1.69 (1.57-1.81)	0.19	145	196
Relations with parents	1.67 (1.53-1.81)	1.85 (1.72-1.99)	0.18	146	195
Having children	1.30 (1.17-1.42)	1.28 (1.18-1.38)	-0.02	145	192
Child raising issues	1.80 (1.64-1.96)	1.94 (1.80-2.08)	0.14	142	182

Table 4.3: Means of indicators or percentages of confirmative answers by mode of data collection with 95% confidence intervals (*continued*)

Conflict style⁴					
Avoid discussion by giving in	2.61 (2.45-2.77)	2.66 (2.53-2.79)	0.05	132	184
Discuss conflicts calmly	4.06 (3.91-4.21)	3.86 (3.72-4.00)	-0.19	136	180
Argue heatedly or get loud	1.88 (1.75-2.02)	2.13 (2.00-2.25)	0.25	136	186
Refuse to talk	1.74 (1.57-1.90)	1.88 (1.75-2.01)	0.14	136	183

Source: GGP Pilot Study 2018.

Note: Distribution for categorical variables in percent (%) and means for metric variables, both with 95% confidence intervals in parentheses. ¹Reference category is “Not thought about break-up”; ²0 “Very dissatisfied” - 10 “Very satisfied”; ³1 “Never” - 5 “Very frequently”; ⁴1 “Never” - 5 “Very frequently”. F2F=face-to-face; Δ=mode difference; n=observations.

When we take the different content dimensions of the indicators into account, we can see that especially those items that concern feelings, such as satisfaction and the perceived stability of a relationship, display stronger mode differences. Based on these univariate findings, over 17% of web respondents – nearly twice as many respondents as in face-to-face mode with 9% – confirm that they thought about breaking up with their current partner. Correspondingly, web respondents rate their general relationship satisfaction nearly 0.5 points lower on a 11-point-scale than face-to-face respondents. The same is true for the satisfaction with daily routines in a relationship, like the division of childcare and household tasks. For the content-specific frequencies of conflicts as well as for the different conflict styles, most items show mode differences with a higher reported frequency of conflicts and of inappropriate conflict behavior in online interviews, compared to face-to-face, however, with varying magnitudes of mode differences between the single conflict items. The item on the frequency of conflicts regarding having children shows a very small and therefore negligible mode difference in the opposite direction. One explanation might be that a majority of the respondents already had children so that this topic was not leading to conflicts between parents anymore. The rather aggressive conflict behavior “argue heatedly or get loud” shows larger mode differences, whereas the comparably

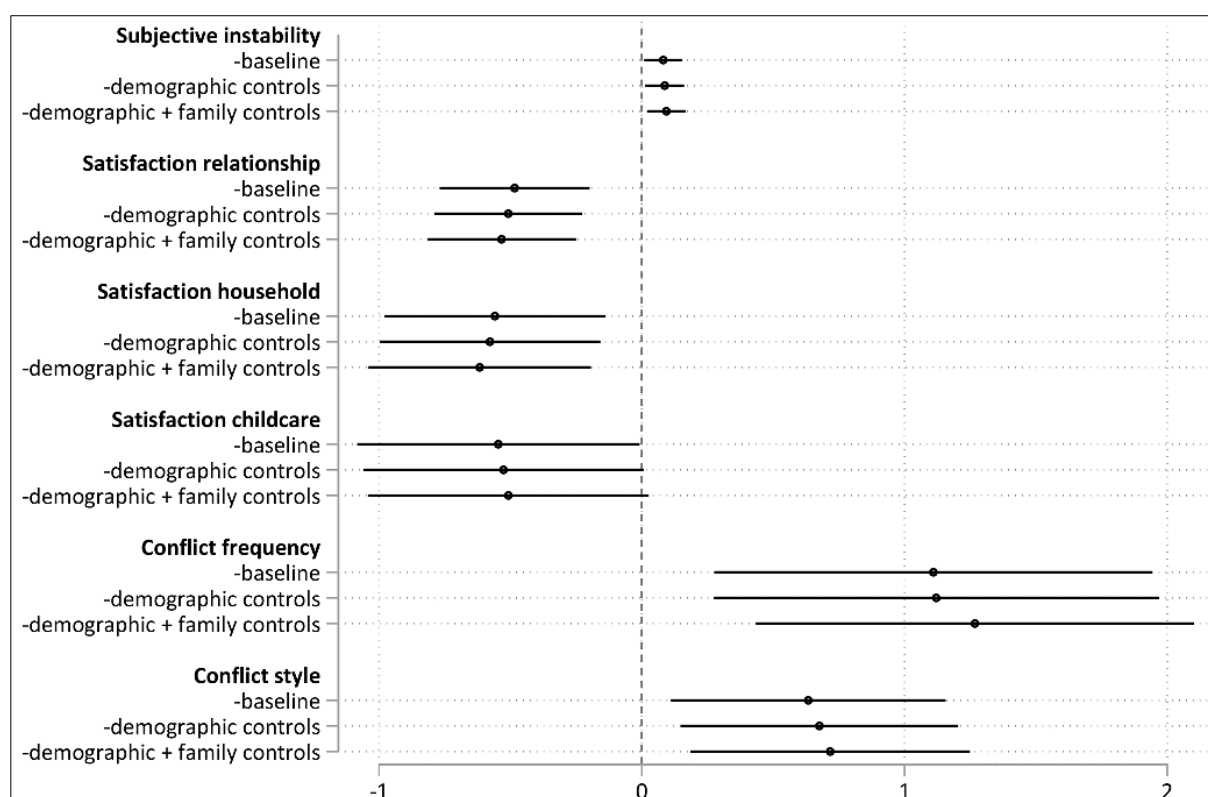
modest conflict style “avoid discussion by giving in” reveals almost no difference between the modes. Because avoiding a discussion might not be a socially undesirable way of dealing with your partner, this conflict behavior is not as clearly indicative of a bad conflict style as the others.

With the help of post-stratification weighting by adjusting sample distributions to the reference distributions of our target sample, we try to control for biases due to selective participation. The weighted distributions of the indicators of relationship quality are, however, very similar (see Table A4.3). This emphasizes that ex-post weighting by the demographic indicators for which reference data are available cannot adjust for the mode-specific differences for our items under study.

4.4.2 Multivariate analyses

We continue to use unweighted data for our analysis as mode differences were the same for unweighted data and for ex-post weighted data. We estimate three models for each of our six outcome variables: a baseline model without control variables, a second model with our demographic confounder variables, and a third model with all demographic and family related control variables. Figure 4.1 displays the effect of web mode on each indicator of relationship quality separately compared to the reference face-to-face mode. Because the focus of this study lies in the evaluation of the mode effect, we refrain from showing the regression results of the control variables. The regression tables of the mode effects can be found in Table A4.4. As described in the methodology section, the single items for frequency of conflict and conflict styles are summed up to two indexes. For each indicator, the effect of web mode is shown, first, as single effect in a baseline model, second, adjusted by standard demographic variables which should control for selective participation, and, third, under additional adjustment of family related

confounders. The size of mode effects can only be compared across the models of the same outcome variable, not between for different outcome variables.



Source: GGP Pilot Study 2018.

Note: Coefficients from linear probability models (LPM) with 95% confidence intervals for dichotomous outcome variables (i.e., subjective instability) and coefficients from linear regression models (OLS) with 95% confidence intervals for metric outcome variables. Demographic controls are sex, nationality, age, education, and regional setting; Family-related controls are relationship status and co-resident children under six years.

Figure 4.1: Effect of web mode on indicators of relationship quality with 95% confidence intervals under block-wise adjustment of demographic and family-related control variables

The multivariate findings in Figure 4.1 confirm that the effect of web mode is robust for all indicators on relationship quality even when we adjust for demographic and family related variables. By comparing the effect of web mode across the three different models for one indicator, we see that the effect is either stable across the models or slightly increases. Therefore, we focus on reporting the findings based on the third model, including demographic and family-related control variables, in the following.

Starting with the indicator on subjective instability, the estimated coefficients based on LPM show that respondents in web mode have a 9.4 percentage points higher probability to report that they thought about a separation than respondents who were asked this question in a face-to-face interview. This means that respondents in the more anonymous web mode are more likely to report that they have thought about breaking up with their current partner than respondents interviewed in a personal interview. The second indicator frequently used for measuring general relationship quality is the overall satisfaction with the relationship. The results confirm the univariate findings: Even under control of demographic and family related variables, respondents in web mode rate their satisfaction with the relationship about 0.5 points lower on a 11-point scale than respondents in face-to-face interviews. The same pattern can be seen for the reporting of satisfaction with specific domains like household and childcare. Respondents in web mode assess their satisfaction with household tasks over 0.6 scale points lower than respondents in face-to-face. Similar for the assessment of satisfaction with childcare tasks: Online respondents rate their satisfaction about 0.5 scale points lower on a scale from 0 to 10 than face-to-face respondents. As can be seen in Figure 4.1 the confidence intervals of the mode effect on satisfaction with childcare touches the zero line slightly which might be explained by the low number of persons who answered this question, as this item was only posed to respondents who have children.

Apparently, feelings and thoughts about the relationship are assessed more negatively in web mode than in face-to-face mode, which speaks for a higher level of self-disclosure and less socially desirable answers in web surveys. In other words, the findings support the assumption that respondents in web interviews are more likely to report that they are less satisfied with their current relationship and are more likely to doubt the stability of the relationship.

The last two indicators focus on the assessment of behavior. For the two indicators, frequency of conflicts and conflict styles, we find a higher reporting of conflicts and of inappropriate

conflict behavior in web mode compared to face-to-face mode. The reporting of the number of conflicts on various topics increases by 1.3 points on a scale from 1 to 29 when respondents answer in web mode compared to face-to-face mode. The effect is similar for the reporting of inappropriate conflict behavior: Compared to respondents in a personal interview, respondents who participate in a web survey report a 0.7 scale points higher level on a scale from 1 to 17.

In summary, web respondents show a consistently higher socially undesirable response behavior than respondents in face-to-face interviews regarding the reporting of feelings as well as of behavior. Considering all indicators of relationship quality examined in this study, respondents in web mode assess their relationship more negatively than those in face-to-face mode.

4.5 Discussion

Our analyses use experimental survey data to assess the existence and the extent of a mode effect, comparing two particularly different modes of data-collection web and face-to-face, on measurements of relationship quality in surveys about families and relationships. Our findings show clear differences for almost all indicators which assess various aspects of the quality of intimate relationships between respondents who are interviewed in a traditional face-to-face design and respondents who participate in a self-administered web interview. Web respondents are more likely to state that they thought about breaking up. They assess a lower relationship satisfaction in general as well as with respect to the distribution of household chores and child-care responsibilities. And they report more conflicts in their partnership as well as higher shares of aggressive or non-constructive conflict behavior. These indicators cover not only different content-related aspects in the context of intimate relationships, but they also rely on a broad range of subjective assessments such as feelings, behavioral patterns, and own experiences. All in all, respondents who participated in web mode report a lower quality in intimate

relationships. These effects are robust, as they remain stable after controlling for other demographic and family-related confounders which correlate with relationship quality and survey participation.

Our results support the assumption that the anonymous and private interview situation of web surveys, compared to traditional face-to-face surveys, leads to a smaller subjectively perceived exposure to social desirability, thereby impacts the responses of interviewees, reduces the bias due to social desirability responding and thus improves the validity of measurements. According to our expectations, web-respondents seem to give more socially and normatively undesirable answers and report a less “rosy” picture of their partnership life than face-to-face respondents. Further, the findings could indicate that respondents who participate online have a higher willingness to self-disclose than respondents who are confronted with an interviewer, which is in line with existing research (Burkill et al., 2016; Robertson et al., 2018). One could assume that measurements on relationship quality conducted in web surveys show a more realistic picture of today’s couple relationships than those conducted face-to-face.

The findings further indicate that the assessment of relationship quality must be considered as highly sensitive and generally biased by effects of social desirability – in web mode only to a lesser degree than in an interviewer-administered mode. However, we cannot prove for a general underreporting of sensitive behavior as we have no reference value of the real situation and can only assess differences in answer behavior between two modes of data collection. This result of relationship quality being a sensitive topic in surveys is relevant in particular for studies about intimate relationships, because surveys in this context are, for the same reason, confronted with the risk of selection biases towards happier and closer relationships (Kalmijn, 2023). One can assume that respondents who are actually less satisfied with their relationship and unhappy with their partner tend to be generally underrepresented in a family survey and are therefore of particular interest.

Depending on the individual situation, some aspects in an intimate relationship might score higher on sensitivity and cause gradually stronger biases based on social desirability than others. As shown in our descriptive findings, single items on conflictual behavior differ in their magnitude of mode differences, which might indicate that the according behaviors are perceived as differently strongly undesirable: For example, refusal to talk may be less undesirable than aggressive and potentially threatening conflict behavior. At the same time, we can assume that similar mode effects could be found for most subjective perceptions and evaluations in other research topics, within family demography and beyond: Many of such indicators might be perceived as sensitive by respondents as they may expect the interviewer or other people to have a certain opinion and according expectations regarding an acceptable answer. The more plausible it is that regarding a certain subjective question a social norm exists, the more likely it is that such an indicator will be biased by effects of social desirability and that it is affected by mode effects.

In our study, we use two extremes of interviewer involvement as an experimental design to sensitize primary researchers as well as data users about the impact of a design decision on data. Nevertheless, there are also mixed-mode designs or hybrid modes of conducting interviews that can be placed on a gradient between face-to-face and web mode, such as CASI modules applied within a personal questionnaire, and these could improve measurement equivalence. A limitation of this study is therefore that it remains an open question of whether the anonymous setting of the web is decisive for the higher degree of disclosure and lower degree of social desirability, or whether the presence of an interviewer or other bystanders might be compensated by such a CASI switch. Even if one could assume less socially desirable answer behavior in CASI than in a face-to-face interview, experimental studies show that web interviews reveal the highest degree of self-disclosure for sensitive questions compared to other self-administered modes (Burkill et al., 2016; Kreuter et al., 2008). Nevertheless, the use of CASI switches should be

used more frequently in personal interviews with intimate and subjective questions, such as those on relationship quality, when a web interview is not possible.

Another limitation is that our experimental study relied on a small number of cases due to budget constraints, as is the case for most experimental studies. Due to the low number of observations, detailed analyses with subgroups, differentiated by gender or age, could not be carried out and thus the methodological approach remained limited. It would be imaginable that, for example, women may be less affected by biases due to social desirability and by mode effects than men since they generally report lower relationship qualities and tend to break up relationships more often than men do. It would be imaginable that parents may be more affected by mode effects than people in childless relationships since maintaining a stable relationship may be stronger socially desired if a child is involved. However, such assumptions require further investigation. While the results in detail must thus therefore be interpreted with caution, the mode effects nevertheless proved robust and revealed a stable pattern across several indicators, which allows us to consider our main findings reliable. It would be highly valuable for family research to analyze whether the measurement of the impact of relationship quality on substantive outcomes, such as break-ups or divorces, is also affected by the mode. Unfortunately, this could not be tested in our study due to small case numbers and due to the lack of a longitudinal design.

We conclude that data users should be aware of the need to control for the mode of data-collection when analyzing data with respect to relationship quality collected in different modes, especially when self-administered as well as interviewer-administered modes were involved. It is not only important to assess data for the representativeness of sociodemographic indicators and, if necessary, weight the data and adjust for these indicators in multivariate analyses, but also to control for and check for interactions with the survey mode when analyzing the data. Particularly in times in which surveys are changing from face-to-face to web mode or to mixed-mode

designs, due to adaption to the COVID-19 pandemic or simply due to cost-efficiency, data users should take the mode of conducting interviews into account. This is especially relevant when central variables measure subjective and sensitive assessments and are prone to social desirability bias. Otherwise, researchers can run the risk of confounding mode effects with substantive effects, for example in terms of cross-national differences or change over time.

References

- AAPOR. (2016). *The American association for public opinion research survey outcome rate calculator 4.1*. <https://www.aapor.org/Education-Resources/For-Researchers/Poll-Survey-FAQ/Response-Rates-An-Overview.aspx>
- Aquilino, W. S. (1991). Telephone versus face-to-face interviewing for household drug use surveys. *International Journal of the Addictions*, 27(1), 71–91. <https://doi.org/10.3109/10826089109063463>
- Arránz Becker, O. (2013). Effects of similarity of life goals, values, and personality on relationship satisfaction and stability: Findings from a two-wave panel study. *Personal Relationships*, 20(3), 443–461. <https://doi.org/10.1111/j.1475-6811.2012.01417.x>
- Atkeson, L. R., Adams, A. N., & Alvarez, R. M. (2014). Nonresponse and mode effects in self- and interviewer-administered surveys. *Political Analysis*, 22(3), 304–320. <https://doi.org/10.1093/pan/mpt049>
- Bethlehem, J., & Biffignandi, S. (2012). *Handbook of web surveys*. Wiley. <https://doi.org/10.1002/9781118121757>
- Bethlehem, J., & Stoop, I. (2007). Online panels – a paradigm theft? In M. Trotman, T. Burrell, L. Gerrard, K. Anderton, G. Basi, M. P. Couper, K. Morris, D. Birks, A. J. Johnson, R. Baker, M. Rigg, S. Taylor, & A. Westlake (Eds.), *The challenges of a changing world: Proceedings of the fifth international conference of the association for survey computing, University of Southampton, 12-14 September 2007* (pp. 113–132). Association for Survey Computing.
- Beullens, K., Loosveldt, G., Vandenplas, C., & Stoop, I. (2018). Response rates in the European Social Survey: Increasing, decreasing, or a matter of fieldwork efforts? *Survey Methods: Insights from the Field*, 1–12. <https://doi.org/10.13094/SMIF-2018-00003>

- Bradbury, T. N., Fincham, F. D., & Beach, S. R. H. (2000). Research on the nature and determinants of marital satisfaction: A decade in review. *Journal of Marriage and Family*, 62(4), 964–980. <https://doi.org/10.1111/j.1741-3737.2000.00964.x>
- Braekman, E., Charafeddine, R., Demarest, S., Drieskens, S., Berete, F., Gisle, L., van der Heyden, J., & van Hal, G. (2020). Comparing web-based versus face-to-face and paper-and-pencil questionnaire data collected through two Belgian health surveys. *International Journal of Public Health*, 65, 5–16. <https://doi.org/10.1007/s00038-019-01327-9>
- Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., Conrad, F., Wellings, K., Johnson, A. M., & Erens, B. (2016). Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles. *PloS One*, 11(2), e0147983. <https://doi.org/10.1371/journal.pone.0147983>
- Callegaro, M. (2008). Social desirability. In P. Lavrakas (Ed.), *Encyclopedia of survey research methods* (pp. 825–826). Sage.
- Chang, L., & Krosnick, J. A. (2010). Comparing oral interviewing with self-administered computerized questionnaires: An experiment. *Public Opinion Quarterly*, 74(1), 154–167. <https://doi.org/10.1093/poq/nfp090>
- Christensen, A. I., Ekholm, O., Glümer, C., & Juel, K. (2013). Effect of survey mode on response patterns: Comparison of face-to-face and self-administered modes in health surveys. *The European Journal of Public Health*, 24(2), 327–332. <https://doi.org/10.1093/eurpub/ckt067>
- Couper, M. P. (2011). The future of modes of data collection. *Public Opinion Quarterly*, 75(5), 889–908. <https://doi.org/10.1093/poq/nfr046>

- Couper, M. P., Antoun, C., & Mavletova, A. (2017). Mobile web surveys. In P. P. Biemer, E. D. de Leeuw, S. Eckman, B. Edwards, F. Kreuter, L. E. Lyberg, N. C. Tucker, & B. T. West (Eds.), *Total survey error in practice* (pp. 133–154). Wiley.
- Daikeler, J., Bošnjak, M., & Lozar Manfreda, K. (2020). Web versus other survey modes: An updated and extended meta-analysis comparing response rates. *Journal of Survey Statistics and Methodology*, 8(3), 513–539. <https://doi.org/10.1093/jssam/smz008>
- Davis, R. E., Couper, M. P., Janz, N. K., Caldwell, C. H., & Resnicow, K. (2010). Interviewer effects in public health surveys. *Health Education Research*, 25(1), 14–26. <https://doi.org/10.1093/her/cyp046>
- De Leeuw, E. D., Hox, J. & Dillman, D. A. (Eds.). (2008). *International handbook of survey methodology*. Taylor & Francis Group.
- DeMaio, T. J. (1984). Social desirability and survey measurement: A review. In C. F. Turner & E. Martin (Eds.), *Surveying subjective phenomena: Volume 2* (pp. 257–282). Russell Sage Foundation.
- Emery, T., Cabaco, S. L. F., Lugtig, P., Toepoel, V., Lück, D., Naderi, R., Schumann, A., & Bujard, M. (2018). *The Generations & Gender Programme: Evaluate, Plan, Initiate: Deliverable 2.1: GGP technical case and e-Needs* (Technical Paper). <https://doi.org/10.31235/osf.io/439wc>
- Fincham, F. D., & Beach, S. R. H. (2006). Relationship satisfaction. In A. L. Vangelisti & D. Perlman (Eds.), *The Cambridge handbook of personal relationships* (pp. 579–594). Cambridge University Press.
- Fu, H., Darroch, J. E., Henshaw, S. K., & Kolb, E. (1998). Measuring the extent of abortion underreporting in the 1995 National Survey of Family Growth. *Family Planning Perspectives*, 30(3), 128–138. <https://doi.org/10.2307/2991627>

- Gauthier, A. H., Cabaço, S. L. F., & Emery, T. (2018). Generations and Gender Survey study profile. *Longitudinal and Life Course Studies*, 9(4), 456–465. <https://doi.org/10.14301/llcs.v9i4.500>
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Wiley.
- Gummer, T., Höhne, J. K., Rettig, T., Roßmann, J., & Kummerow, M. (2023). Is there a growing use of mobile devices in web surveys? Evidence from 128 web surveys in Germany. *Quality & Quantity*, 57(6), 5333–5353. <https://doi.org/10.1007/s11135-022-01601-8>
- Gummer, T., Schmiedeberg, C., Bujard, M., Christmann, P., Hank, K., Kunz, T., Lück, D., & Neyer, F. J. (2020). The impact of Covid-19 on fieldwork efforts and planning in pairfam and FReDA-GGS. *Survey Research Methods*, 14(2), 223–227. <https://doi.org/10.18148/srm/2020.v14i2.7740>
- Gustavson, K., Nilsen, W., Ørstavik, R., & Røysamb, E. (2014). Relationship quality, divorce, and well-Being: Findings from a three-year longitudinal study. *The Journal of Positive Psychology*, 9(2), 163–174. <https://doi.org/10.1080/17439760.2013.858274>
- Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research*, 21(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>
- Heerwegh, D., & Loosveldt, G. (2008). Face-to-face versus web surveying in a high-internet-coverage population: Differences in response quality. *Public Opinion Quarterly*, 72(5), 836–846. <https://doi.org/10.1093/poq/nfn045>
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. (2003). Telephone versus face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly*, 67(1), 79–125. <https://doi.org/10.1086/346010>

- Huss, B., & Pollmann-Schult, M. (2020). Relationship satisfaction across the transition to parenthood: The impact of conflict behavior. *Journal of Family Issues*, 41(3), 383–411. <https://doi.org/10.1177/0192513X19876084>
- Kalmijn, M. (2023). Are national family surveys biased toward the happy family? A multiactor analysis of selective survey nonresponse. *Sociological Methods & Research*, 52(2), 561–586. <https://doi.org/10.1177/0049124120986208>
- Kantar Public (2018). *Beziehungen und Familienleben in Deutschland (pairfam): Methodenbericht Welle 10 (2017/2018)*. Kantar Public, Munich.
- Karney, B. R., & Bradbury, T. N. (1995). The longitudinal course of marital quality and stability: A review of theory, method, and research. *Psychological Bulletin*, 118(1), 3–34. <https://doi.org/10.1037/0033-2909.118.1.3>
- Karney, B. R., & Bradbury, T. N. (2020). Research on marital satisfaction and stability in the 2010s: Challenging conventional wisdom. *Journal of Marriage and Family*, 82(1), 100–116. <https://doi.org/10.1111/jomf.12635>
- Kelly, C. A., Soler-Hampejsek, E., Mensch, B. S., & Hewett, P. C. (2013). Social desirability bias in sexual behavior reporting: Evidence from an interview mode experiment in rural Malawi. *International Perspectives on Sexual and Reproductive Health*, 39(1), 14–21. <https://doi.org/10.1363/3901413>
- Kluwer, E. S., & Johnson, M. D. (2007). Conflict frequency and relationship quality across the transition to parenthood. *Journal of Marriage and Family*, 69(5), 1089–1106. <https://doi.org/10.1111/j.1741-3737.2007.00434.x>
- Kreuter, F., Presser, S., & Tourangeau, R. (2008). Social desirability bias in CATI, IVR, and web surveys: The effects of mode and question sensitivity. *Public Opinion Quarterly*, 72(5), 847–865. <https://doi.org/10.1093/poq/nfn063>

- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025–2047. <https://doi.org/10.1007/s11135-011-9640-9>
- Lee, R. M., & Renzetti, C. M. (1990). The problems of researching sensitive topics: An overview and introduction. *American Behavioral Scientist*, 33(5), 510–528. <https://doi.org/10.1177/0002764290033005002>
- Lewis, R. A., & Spanier, G. B. (1979). Theorizing about the quality and stability of marriage. In W. R. Burr, R. Hill, F. I. Nye, & I. L. Reiss (Eds.), *Contemporary theories about the family: Volume 1* (pp. 268–293). Free Press.
- Liu, M. (2017). Data collection mode differences between national face-to-face and web surveys on gender inequality and discrimination questions. *Women's Studies International Forum*, 60, 11–16. <https://doi.org/10.1016/j.wsif.2016.11.007>
- Liu, M., & Stainback, K. (2013). Interviewer gender effects on survey responses to marriage-related questions. *Public Opinion Quarterly*, 77(2), 606–618. <https://doi.org/10.1093/poq/nft019>
- Liu, M., & Wang, Y. (2016). Comparison of face-to-face and web surveys on the topic of homosexual rights. *Journal of Homosexuality*, 63(6), 838–854. <https://doi.org/10.1080/00918369.2015.1112587>
- Lutig, P., Toepoel, V., Emery, T., Cabaço, S. L. F., Bujard, M., Naderi, R., Schumann, A., & Lück, D. (2022). *Can we successfully move a cross-national survey online? Results from a large three-country experiment in the Generations and Gender Programme* (Working Paper). Center for Open Science. <https://doi.org/10.31235/osf.io/mu8jy>
- Medway, R. (2012). *Beyond response rates: The effect of prepaid incentives on measurement error* [Doctoral dissertation, University of Maryland]. College Park ProQuest Dissertations & Theses.

- Rijken, A. J., & Liefbroer, A. C. (2009). The influence of partner relationship quality on fertility. *European Journal of Population*, 25(1), 27–44. <https://doi.org/10.1007/s10680-008-9156-8>
- Robertson, R. E., Tran, F. W., Lewark, L. N., & Epstein, R. (2018). Estimates of non-heterosexual prevalence: The roles of anonymity and privacy in survey methodology. *Archives of Sexual Behavior*, 47(4), 1069–1084. <https://doi.org/10.1007/s10508-017-1044-z>
- Schmid, L., Wörn, J., Hank, K., Sawatzki, B., & Walper, S. (2021). Changes in employment and relationship satisfaction in times of the COVID-19 pandemic: Evidence from the German family panel. *European Societies*, 23(sup1), S743–S758. <https://doi.org/10.1080/14616696.2020.1836385>
- Schonlau, M., & Couper, M. P. (2017). Options for conducting web surveys. *Statistical Science*, 32(2), 279–292. <https://doi.org/10.1214/16-STS597>
- Schonlau, M., van Soest, A. H. O., & Kapteyn, A. (2007). *Are 'webographic' or attitudinal questions useful for adjusting estimates from web surveys using propensity scoring?* RAND Working Paper Series WR-506. <https://doi.org/10.2139/ssrn.1006108>
- Schonlau, M., van Soest, A., Kapteyn, A., & Couper, M. P. (2009). Selection bias in web surveys and the use of propensity scores. *Sociological Methods & Research*, 37(3), 291–318. <https://doi.org/10.1177/0049124108327128>
- Schröder, J., & Schmiedeberg, C. (2023). Effects of partner presence during the interview on survey responses: The example of questions concerning the division of household labor. *Sociological Methods & Research*, 52(2), 933–955. <https://doi.org/10.1177/0049124120914938>
- Schumann, A. & Lück, D. (2023). Better to ask online when it concerns intimate relationships? Survey mode differences in the assessment of relationship quality. *Demographic Research*, 48, 609–640. <https://www.jstor.org/stable/48728217>

- Tourangeau, R. (2017). Presidential address: Paradoxes of nonresponse. *Public Opinion Quarterly*, 81(3), 803–814. <https://doi.org/10.1093/poq/nfx031>
- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60(2), 275–304. <https://doi.org/10.1086/297751>
- Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin*, 133(5), 859–883. <https://doi.org/10.1037/0033-2909.133.5.859>
- Van Damme, M., & Dykstra, P. (2018). Spousal resources and relationship quality in eight European countries. *Community, Work & Family*, 21(5), 541–563. <https://doi.org/10.1080/13668803.2018.1526776>
- Vehovar, V., Batagelj, Z., Manfreda, K. L., & Zaletel, M. (2002). Nonresponse in web surveys. In R. M. Groves, D. A. Dillman, J. L. Eltinge, & R. J. A. Little (Eds.), *Survey nonresponse* (pp. 229–242). Wiley.
- Weinreb, A., Sana, M., & Stecklov, G. (2018). Strangers in the field: A methodological experiment on interviewer-respondent familiarity. *Bulletin of Sociological Methodology/ Bulletin De Méthodologie Sociologique*, 137-138(1), 94–119. <https://doi.org/10.1177/0759106318761562>
- Wiik, A. K., Keizer, R., & Lappegård, T. (2012). Relationship quality in marital and cohabiting unions across Europe. *Journal of Marriage and Family*, 74(3), 389–398. <https://doi.org/10.1111/j.1741-3737.2012.00967.x>
- Wolf, C., Christmann, P., Gummer, T., Schnaudt, C., & Verhoeven, S. (2021). Conducting General Social Surveys as self-administered mixed-mode surveys. *Public Opinion Quarterly*, 85(2), 623–648. <https://doi.org/10.1093/poq/nfab039>

Appendix

Table A4.1: Distributions and sources of reference information from official statistics used for post-stratification weighting

Indicator	Source	Year of con- duction	Age of population (in years)	Categories	Distribu- tion (in percent)
Sex	Micro Census	2016	18-49	(1) Male (2) Female	(1) 51.63 (2) 48.37
Age	Census	2011	18-49	(1) 18-29 (2) 30-39 (3) 40-49	(1) 32.95 (2) 28.02 (3) 39.02
School education	Census	2011	18-49	(1) Not (yet) graduated/ low school education ("Hauptschulabschluss") (2) Middle school education ("Realschulabschluss") (3) High school education ("Fachhochschulabschluss/ Abitur")	(1) 38.08 (2) 30.55 (3) 31.38
Nationality	Census	2011	18-49	(1) German (2) Non-German	(1) 90.71 (2) 9.29
Regional setting	Registry office	2018	18-49	(1) Rural area (2) Urban area	(1) 50.00 (2) 50.00

Source: Micro Census 2016, Census 2011, Register data 2018.

Table A4.2: Original wording of questions and answers for items measuring relationship quality in the GGP Pilot Study 2018

Item	Question text	Answer Categories
a220	Even people who get along well with their partners sometimes wonder whether their marriage or partnership will work. Over the past 12 months, have you thought about breaking up your relationship?	-Yes -No
a217	How satisfied are you with your relationship with your partner/spouse?	-On a scale from 0 to 10 where 0 means 'not at all satisfied' and 10 means 'completely satisfied' and 5 means 'about average', what number best represents your satisfaction? 0-10
a312	How satisfied are you with the division of household tasks between you and your partner/spouse?	
a314	How satisfied are you with the way childcare tasks are divided between you and your partner/spouse?	
a218	In the last twelve months, how often did you have disagreements with your partner about...	-Never -Seldom -Sometimes -Frequently -Very Frequently
a218a	...household chores?	
a218b	...money?	
a218c	...use of leisure time?	
a218d	...relations with friends?	
a218e	...relations with parents?	
a218f	...having children?	
a218g	...child raising issues?	-Never -Seldom -Sometimes -Frequently -Very Frequently
a219	Couples deal with serious disagreements in very different ways. If you had a serious disagreement with your partner, how often do you...	
a219a	...avoid discussion by giving in?	
a219b	...discuss your disagreement calmly?	
a219c	...argue heatedly or shout?	
a219d	...refuse to talk about it?	

Source: GGP Pilot Study 2018.

Table A4.3: Post-stratification weighted means of indicators or percentages of confirmative answers by mode of data collection

	F2F	Web	Mode difference	F2F	Web
	Mean or percentage		Δ	n	
Subjective instability¹					
Thought about breaking up (%)	6.28 (3.53-10.93)	14.49 (9.55-21.40)	8.21	144	175
Satisfaction²					
General relationship	9.15 (8.96-9.35)	8.71 (8.48-8.94)	-0.44	144	185
Household tasks	8.47 (8.16-8.78)	7.88 (7.50-8.25)	-0.59	126	160
Childcare tasks	8.63 (8.24-9.02)	8.24 (7.78-8.69)	-0.39	70	79
Conflict frequency³					
Household chores	2.30 (2.12-2.49)	2.42 (2.26-2.57)	0.12	145	188
Money	1.68 (1.52-1.83)	2.05 (1.85-2.26)	0.37	145	188
Leisure time	2.25 (2.06-2.44)	2.38 (2.22-2.54)	0.13	145	188
Relations with friends	1.52 (1.39-1.65)	1.71 (1.55-1.86)	0.19	144	188
Relations with parents	1.68 (1.52-1.83)	1.84 (1.69-2.00)	0.16	145	188
Having children	1.30 (1.16-1.44)	1.23 (1.14-1.33)	0.07	144	185
Child raising issues	1.81 (1.63-2.00)	1.89 (1.70-2.08)	0.08	141	174
Conflict style⁴					
Avoid discussion by giving in	2.71 (2.53-2.91)	2.67 (2.50-2.84)	-0.04	131	177
Discuss conflicts calmly	4.03 (3.85-4.20)	3.95 (3.78-4.12)	-0.08	135	175
Argue heatedly or get loud	1.91 (1.77-2.05)	2.06 (1.88-2.23)	0.15	135	179
Refuse to talk	1.71 (1.55-1.87)	1.83 (1.67-1.99)	0.12	135	177

Source: GGP Pilot Study 2018.

Note: Distribution for categorical variables in percent (%) and means for metric variables, both with 95% confidence intervals in parentheses. ¹Reference category is "Not thought about break-up"; ²0 "Very dissatisfied" - 10 "Very satisfied"; ³1 "Never" - 5 "Very frequently"; ⁴1 "Never" - 5 "Very frequently". F2F=face-to-face; Δ =mode difference; n=observations.

Table A4.4: Effect of web mode on separate indicators of relationship quality under block-wise adjustment of demographic and family-related confounders

	Model 1		Model 2		Model 3	
Linear probability models						
Subjective Instability						
	Coef.	SE	Coef.	SE	Coef.	SE
Mode (Ref.: Face-to-face)						
Web	0.082 (0.009- 0.154)	0.037	0.087 (0.014- 0.161)	0.037	0.094 (0.020- 0.167)	0.037
Observations	326		319		317	
Degrees of freedom (df)	1		6		9	
R-squared	0.01		0.04		0.06	
Adjusted R-squared	0.01		0.02		0.04	
Logistic regression models						
Subjective Instability						
	AME	SE	AME	SE	AME	SE
Mode (Ref.: Face-to-face)						
Web	0.082 (0.099- 0.153)	0.037	0.088 (0.016- 0.161)	0.037	0.095 (0.022- 0.168)	0.037
Observations	326		319		317	
Degrees of freedom (df)	1		6		9	
AIC	257.26		256.99		254.38	
BIC	264.84		283.35		291.98	
Linear regression models						
Satisfaction relationship						
	Coef.	SE	Coef.	SE	Coef.	
Mode (Ref.: Face-to-face)						
Web	-0.484 (-0.769- -0.198)	0.143	-0.508 (-0.789- -0.227)	0.143	-0.533 (-0.817- -0.250)	0.144
Observations	335		329		327	
Degrees of freedom (df)	1		6		9	
R-squared	0.03		0.11		0.12	
Adjusted R-squared	0.03		0.09		0.10	
Linear regression models						
Satisfaction household						
	Coef.	SE	Coef.	SE	Coef.	
Mode (Ref.: Face-to-face)						
Web	-0.559 (-0.980- -0.138)	0.214	-0.578 (-0.998- -0.158)	0.213	-0.617 (-1.040- -0.194)	0.215
Observations	293		286		285	
Degrees of freedom (df)	1		6		8	
R-squared	0.02		0.09		0.10	
Adjusted R-squared	0.02		0.07		0.07	
Linear regression models						
Satisfaction childcare						
	Coef.	SE	Coef.	SE	Coef.	
Mode (Ref.: Face-to-face)						
Web	-0.546 (-1.083- -0.009)	0.272	-0.526 (-1.060- -0.007)	0.270	-0.508 (-1.042- 0.027)	0.270
Observations	151		149		149	
Degrees of freedom (df)	1		6		7	
R-squared	0.03		0.11		0.11	
Adjusted R-squared	0.02		0.07		0.07	

Table A4.4: Effect of web mode on separate indicators of relationship quality under block-wise adjustment of demographic and family-related confounders (*continued*)

Linear regression models						
Conflict frequency						
	Coef.	SE	Coef.	SE	Coef.	
Mode (<i>Ref.: Face-to-face</i>)						
Web	1.110 (0.276- 1.944)	0.424	1.121 (0.273- 1.968)	0.431	1.268 (0.434- 2.102)	0.424
Observations	320		313		311	
Degrees of freedom (df)	1		6		9	
R-squared	0.02		0.04		0.10	
Adjusted R-squared	0.02		0.02		0.01	
Linear regression models						
Conflict style						
	Coef.	SE	Coef.	SE	Coef.	
Mode (<i>Ref.: Face-to-face</i>)						
Web	0.634 (0.111- 1.156)	0.266	0.676 (0.148- 1.204)	0.268	0.717 (0.185- 1.249)	0.270
Observations	309		304		302	
Degrees of freedom (df)	1		6		9	
R-squared	0.02		0.04		0.05	
Adjusted R-squared	0.02		0.02		0.02	

Source: GGP Pilot Study 2018.

Note: Coefficients (Coef.) from linear probability models (LPM) and average marginal effects (AME) from lo-gistic regression models for dichotomous outcome variables (i.e., subjective instability) and coefficients from linear regression models (OLS) for metric outcome variables, each with 95% confidence intervals in parentheses and standard errors (SE). Model 1=baseline; Model 2=under adjustment of demographic variables (sex, nationality, age, education and regional setting); Model 3=under adjustment of family-related variables (relationship status, co-resident children under six years).

5 General Discussion and Conclusion

Summary and implications

The primary objective of surveys, family surveys included, is to accurately represent and measure the target population and constructs of interest (see Groves et al., 2004; Vannette & Krosnick, 2017). However, this objective is compromised by nonresponse and measurement errors that can arise throughout the survey process, especially in the case of changes in survey design or fieldwork procedures. Such changes are often necessary due to various factors such as cost considerations, limited funding, or the integration of new technologies into the data collection process (Couper, 2013, 2017).

The three studies presented in this dissertation examined the impact of methodological changes to the collection of data by focusing on the challenges posed by the survey topic (here: families and partnerships) and three different design characteristics often inherent in family surveys: longitudinal design, multi-actor approach, and the measurement of sensitive questions. At the same time, the three studies highlighted relevant processes in surveys including the consent to participate, participation, and response.

The first study, presented in Chapter 2, focuses on the initial condition for participation in a panel study: consent for future participation. This is a crucial survey process and serves as a fundamental prerequisite for collecting panel data, making it vital for data infrastructures that collect longitudinal data. Based on pairfam data, I examined which respondents refused consent for future participation in the 14th panel wave after an institutional change was announced. My expectation was that those who reported negative experiences in their family and partnership during the interview in wave 14 would be more likely to withdraw their consent. However, apart from having a partner, which increased the likelihood of consent, the findings did not

suggest that providing panel consent was related to the study topic. Experiences of separation and negative evaluations of one's own relationship did not appear to influence the decision to participate in future panel waves. Nevertheless, findings indicate that the denial of consent occurred primarily among sociodemographic groups that are generally underrepresented in surveys, such as individuals with lower education levels. Moreover, respondents who had been participating in the panel study for a longer time were more likely to provide panel consent than those who had entered the panel only recently.

These results bear important implications for collectors of panel data. Retaining as many respondents as possible is crucial to maintaining a reasonable sample size across waves, as attrition decreases panel size and the statistical power of data collected with each wave. As the value of panel data accumulates over time, data collectors should consider that the request to consent to future waves can be seen as an easy exit option among respondents. Therefore, the project team should carefully evaluate whether changes in the institution responsible for data collection are truly necessary, as these provoke a renewal of panel consent. Although this did not result in a selection bias towards happier relationships or more socially desirable partnership forms, refusal of consent was not random, with individuals from hard-to-reach groups being less likely to provide consent. Accordingly, data users should consider using calibration weights to mitigate bias due to increased selectivity within the remaining panel sample.

The second study, presented in Chapter 3, builds on this process by examining not only the consent to participate, but also participation as a subsequent survey request. This study explored the recruitment and participation process of partners as secondary respondents in a self-administered multi-actor survey by examining aspects that encourage the two actors to comply with a survey request during the recruitment process. As both respondents are in a relationship, my expectation was that the relationship situation and its assessment would impact their

engagement in a multi-actor survey. The results revealed that relationships characterized by a higher degree of satisfaction and closeness were more likely to comply with survey requests during both the recruitment and participation process. This was evident in both the primary respondent's decision to consent to interviewing their partner and the secondary respondent's decision to participate. Along with sociodemographic characteristics such as higher levels of education and having a migration background, the findings further show that invitations sent directly to the secondary respondent at earlier stages increased the likelihood of participation in a multi-actor survey.

These results suggest that positive relationship characteristics of the dyad under study enhance the participation of secondary respondents in a self-administered multi-actor survey. This underlines the significance of relationship dynamics for the collection of multi-actor data. Hence, data collectors may leverage this insight and employ strategies to counteract selective dropout during recruitment, particularly in a self-administered survey mode with no interviewer involvement as a source of motivation. Accordingly, strategies should focus on providing special incentives tailored to the needs and concerns of the dyad under study. Special attention should also be paid to the invitation of the secondary respondent in a self-administered interview: Quickly establishing a direct contact seemed to increase the chances of participation. These findings also imply a challenge for users of multi-actor data, as results indicate a selection bias towards higher-quality relationships so that couples with poorer relationship quality are likely to be underrepresented in the dyadic sample. Due to the lack of reference information for subjective indicators, weighting procedures cannot be applied to counteract these selection effects. Therefore, data users should bear in mind that the generalizability of their findings based on self-administered multi-actor data is limited, particularly when analyzing relationship quality.

After addressing consent to participate and actual participation, the study presented in Chapter 4 examined the survey response process, switching from the “representation” dimension to the “measurement” dimension of a survey lifecycle. Specifically, it focuses on measuring accurate responses for topics relevant to studying family and partnership life. As the measurement of sensitive and personal topics is often influenced by social desirability, it is crucial for surveys covering sensitive topics to analyze interview conditions that promote the highest response accuracy. I expected respondents to be more open and honest about their relationships in an online setting, as the absence of an interviewer reduces the social pressure to not admit when things are not going well. I therefore tested whether the mode of data collection affects measurements related to the assessment of relationship quality, drawing on experimental data from a pilot study executed as part of the GGP project. The findings revealed that respondents who participated online reported consistently lower levels of relationship quality across various indicators measured in the GGS questionnaire compared to those who participated in face-to-face interviews. This suggests that the more anonymous CAWI mode leads to less socially desirable response behavior when reporting emotions and behaviors within partnerships. These results highlight that measurement effects between data collection modes exist for questions that align with socially desirable concepts or expectations of a survey, such as partnership quality in a family survey.

In line with a large body of literature (e.g., Burkill et al., 2016; Heerwegh, 2009; Liu, 2017; Tourangeau & Smith, 1996), these findings suggest that self-administered data collection modes such as web interviews may be preferable for surveys addressing sensitive topics. Further, this applies to topics that correlate with normative or social expectations and that include subjective questions about personal lives. Moreover, as changes in survey mode are sometimes required, both data collectors and data users should consider mode effects on measurement as a potential source of bias. On the one hand, longitudinal or cross-national studies planning a

mode switch should consider conducting mode experiments or implementing a mixed-mode design as a transitional approach to identify and later account for mode measurement effects. Data users, on the other hand, should be aware of the necessity to account for different modes in their analysis and in interpreting results when comparing data from panel or cross-national surveys that use or mix different modes. Beyond that, these findings also imply that indicators used to measure the quality of a partnership are influenced by social desirability. Consequently, family surveys, regardless of the survey mode, may tend to underestimate this construct.

In summary, the three studies presented in this dissertation provide insights into key processes for collecting panel, multi-actor, and sensitive data. Changes in survey methods can introduce nonresponse and measurement error at each of these stages. This has been shown to lead to increased sample selectivity within both a long-running panel survey and a newly recruited multi-actor survey, affecting both the representation of family and partnership life and the representation of sociodemographic subgroups, prompting topic-related and sociodemographic selection effects. In terms of measurement, response behavior for sensitive questions is closely related to the study's topic, with the assessment of partnership and family life being strongly affected by social desirability bias between interviewer- and self-administered survey modes in family research. This indicates that increased sample selectivity and inaccurate measurements in family surveys are a result of methodological changes in survey design and data collection procedures, undermining a survey's objective.

Limitations and future research

The three studies presented in this dissertation aimed to present examples of changes that may occur in family survey designs and their effects on two particular potential sources of errors affecting data quality. I have focused on measurement and nonresponse errors as the interaction between the target person and the survey topic is particularly likely to affect the survey

processes of participation and responding, influencing how the addressed individual reacts to a survey request. However, it is important to note that there are also other sources of error and determinants that can arise during survey data collection, as illustrated in the total survey error framework shown in Figure 1.1 (Biemer et al., 2017; Groves et al., 2004). For instance, the construct of interest may not be measured validly, or the sampling frame may not adequately cover the target population. Additionally, nonresponse not only depends on a sample member's decision to cooperate, but can also arise from noncontacts (Lynn, 2008; Lynn & Clarke, 2002). Ultimately, all of these factors may impact the survey process to varying degrees, leading to cumulated errors that affect data quality (Tourangeau, 2019).

Another relevant aspect for future development in this field of research is an examination of cumulated error. In this dissertation, I examined two sources of error separately, as certain changes in survey methods exert a stronger influence on either the “representation” or “measurement” dimension of conducting a survey. However, previous research has suggested a relationship between participation probability and response accuracy (Olson, 2006). This is particularly true for sensitive and socially undesirable topics, leading to target individuals who are reluctant to participate and/or report accurately (Tourangeau et al., 2010). This could also pertain to the indicators of partnership situations and quality analyzed in this investigation. Future research should explore the interdependence of these sources of error in family surveys in greater detail, and attempts should be made to disentangle selection from measurement effects in order to examine whether they accumulate.

An additional factor to be mindful of is the feasibility of testing changes in survey design experimentally, as was done in the GGP survey mode experiment presented in Chapter 4. Ongoing survey programs often cannot afford to conduct such experiments due to limited resources or the need for fast decision-making to maintain the data infrastructure. Therefore, two of the three

studies presented here rely on non-experimental data. I used prior respondent information to examine how nonresponse in the sense of refusing panel consent relates to the survey's topic in Chapter 2. In Chapter 3, I analyzed proxy data from primary respondents concerning secondary nonrespondents. However, analyzing topic-related nonresponse in a sample that is already established may underestimate effects, as those who are less interested in the topic are likely to have refused initial participation. At the same time, testing topic-related nonresponse in an experimental setting with a randomly drawn sample is challenging, as information about the personal situation of sample members who do not participate is missing. To grasp the significance of topic-related nonresponse, future research should explore the application of mixed methods approaches to elaborate further on the underlying motives of individuals' (un)willingness to comply with a survey request, such as conducting qualitative interviews or targeted questionnaires for both respondents and nonrespondents.

For data users, it would be useful to examine not only the direct effects on data quality, but also explore how other substantive outcomes such as total fertility rate, marital stability and dissolution, and care arrangements are affected in family demography research. For instance, studies using panel data indicate that individuals experiencing life changes such as separation are more likely to drop out between panel waves (Kapteyn et al., 2006; Müller & Castiglioni, 2015), which may lead to biased conclusions regarding the causes and consequences of separation. However, using multi-actor data on parent-child relationships, Kalmijn and Liefbroer (2011) found evidence for a selection towards closer family ties in multi-actor data, but this had no consistent effects on substantive outcomes such as parental support and child well-being. In that respect, data users should also acknowledge that individual selection effects on sociodemographic characteristics such as educational level and gender may interact, leading to biases in substantive outcomes. For instance, an overrepresentation of highly educated women can lead to distortions of results on fertility or family behavior, as this subgroup is associated with lower

rates of parenthood (Schaeper et al., 2017). Therefore, it would be worthwhile for future research to assess whether weighting procedures or adjusted estimation models can account for such associations or interactions in substantive results in family research and beyond.

Contribution

The contribution of the findings in this dissertation apply not only to data infrastructures in family research, but also beyond. Family surveys are only one example of surveys with a special topic that raise sensitivity to selection and measurement effects. Most large-scale surveys in the social sciences have a thematic focus, such as the European panel SHARE that addresses the challenges and life situations of the older population, the US National Health Interview Survey (NHIS) that focuses on health and well-being, or GLEN, a large-scale panel study for environmental social science research in Germany. Surveys therefore face the same sources of error, but to which extent these errors bias the data cannot be generalized across surveys with different thematic contexts. This dissertation aims to reduce sample selectivity and improve measurement accuracy regarding the thematic focus of surveys. Topic-related norms and ideals may create challenges in measuring and representing key indicators, affecting both survey response rates and data quality. There is clearly an urgent need to consider context-specific factors when analyzing nonresponse and measurement errors as survey methodologies evolve and when survey design and data collection procedures are adjusted.

Furthermore, understanding the impact of such methodological decisions is vital for the establishment, maintenance, and continuation of data infrastructures such as pairfam, FReDA, and the GGP. Accordingly, my findings provide key insights and practical recommendations for data infrastructures to address and manage nonresponse and measurement errors throughout the survey process. This is particularly important for those collecting panel data over extended periods of time, in particular when obtaining consent for future participation. It is also relevant

for those implementing a multi-actor design as part of recruiting target individuals within relationships, as well as for data infrastructures that aim to gather sensitive and personal information prone to social desirability effects, particularly when utilizing different survey modes.

As my findings aim to raise awareness in analyzing survey data and interpreting results based on such data, this dissertation addresses not only primary researchers who collect data, but also secondary researchers who draw upon it for analyses. Although the three empirical cases presented here (i.e., pairfam, FReDA, and the GGS) demonstrate that data collectors need methodological flexibility to adapt to new circumstances, such adaptations have introduced and/or exacerbated biases that may complicate the interpretation of associations and trends in family processes and dynamics. My goal is to improve analyses, emphasize the need for thorough robustness checks, and clarify how methodological choices in data collection can influence data quality, potentially undermining research objectives.

In conclusion, I would like to quote a few lines by Couper (2013) to leave the reader with something to reflect on:

“Like good wine, the provenance of the data we analyze is important, as is quality. We need to educate users on how to consume data. Sometimes I fear this may be a lost cause. Analytic software makes it too easy for people to conduct analyses without concern for where the data come from or how they are produced. The analytic software we use is agnostic as to the source of the data. Also, the sheer volume of data, and the number of people who directly consume data without regard for source, makes this an almost impossible task. But we must try, at least among ourselves – in the papers we present, in the journal articles we submit

and review, in the reports we write. We should take care to point out what we did, and alert readers to the risks of using the data.”
(Couper, 2013, pp. 153-154)

References

- Biemer, P. P., De Leeuw, E. D., Eckman, S., Edwards, B., Kreuter, F., Lyberg, L. E., Tucker, N. C., & West, B. T. (Eds.). (2017). *Total survey error in practice*. Wiley.
- Burkill, S., Copas, A., Couper, M. P., Clifton, S., Prah, P., Datta, J., Conrad, F., Wellings, K., Johnson, A. M., & Erens, B. (2016). Using the web to collect data on sensitive behaviours: A study looking at mode effects on the British National Survey of Sexual Attitudes and Lifestyles. *PloS One*, *11*(2), e0147983. <https://doi.org/10.1371/journal.pone.0147983>
- Couper, M. P. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods*, *7*(3), 145–156. <https://doi.org/10.18148/srm/2013.v7i3.5751>
- Couper, M. P. (2017). New developments in survey data collection. *Annual Review of Sociology*, *43*(1), 121–145. <https://doi.org/10.1146/annurev-soc-060116-053613>
- Groves, R. M., Fowler Jr, F. J., Couper, M. P., Lepkowski, J. M., Singer, E., & Tourangeau, R. (2004). *Survey methodology*. Wiley.
- Heerwegh, D. (2009). Mode differences between face-to-face and web surveys: An experimental investigation of data quality and social desirability effects. *International Journal of Public Opinion Research*, *21*(1), 111–121. <https://doi.org/10.1093/ijpor/edn054>
- Kalmijn, M., & Liefbroer, A. C. (2011). Nonresponse of secondary respondents in multi-actor surveys: Determinants, consequences, and possible remedies. *Journal of Family Issues*, *32*(6), 735–766. <https://doi.org/10.1177/0192513X10390184>
- Kapteyn, A., Michaud, P.-C., Smith, J. P., & van Soest, A. (2006). *Effects of attrition and non-response in the Health and Retirement Study*. IZA Discussion Paper No. 2246, Institute for the Study of Labor (IZA), Bonn.

- Liu, M. (2017). Data collection mode differences between national face-to-face and web surveys on gender inequality and discrimination questions. *Women's Studies International Forum*, 60, 11–16. <https://doi.org/10.1016/j.wsif.2016.11.007>
- Lynn, P. (2008). The problem of nonresponse. In E. D. de Leeuw, J. Hox, & D. A. Dillman (Eds.), *International handbook of survey methodology* (pp. 35–55). Taylor & Francis Group.
- Lynn, P., & Clarke, P. (2002). Separating refusal bias and non-contact bias: evidence from UK national surveys. *Journal of the Royal Statistical Society Series D: The Statistician*, 51(3), 319–333. <https://doi.org/10.1111/1467-9884.00321>
- Müller, B., & Castiglioni, L. (2015). Stable relationships, stable participation? The effects of partnership dissolution and changes in relationship stability on attrition in a relationship and family panel. *Survey Research Methods*, 9(3), 205–219. <https://doi.org/10.18148/srm/2016.v10i1.6207>
- Olson, K. (2006). Survey participation, nonresponse bias, measurement error bias, and total bias. *Public Opinion Quarterly*, 70(5), 737–758. <https://doi.org/10.1093/poq/nfl038>
- Schaeper, H., Grotheer, M., & Brandt, G. (2017). Childlessness and fertility dynamics of female higher education graduates in Germany. In M. Kreyenfeld and D. Konietzka (Eds.), *Childlessness in Europe: Contexts, causes, and consequences* (pp. 209–232). Springer.
- Tourangeau, R. (2019). How errors cumulate: Two examples. *Journal of Survey Statistics and Methodology*, 8(3), 413–432. <https://doi.org/10.1093/jssam/smz019>
- Tourangeau, R., Groves, R. M., & Redline, C. D. (2010). Sensitive topics and reluctant respondents: Demonstrating a link between nonresponse bias and measurement error. *Public Opinion Quarterly*, 74(3), 413–432. <https://doi.org/10.1093/poq/nfq004>

- Tourangeau, R., & Smith, T. W. (1996). Asking sensitive questions: The impact of data collection mode, question format, and question context. *Public Opinion Quarterly*, 60(2), 275–304. <https://doi.org/10.1086/297751>
- Vannette, D. L., & Krosnick, J. A. (Eds.). (2017). *The Palgrave handbook of survey research*. Springer.