UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Anže Sendelbah

Medijska večopravilnost v spletnih anketah
Media multitasking in web surveys

Doktorska disertacija

Ljubljana, 2017

UNIVERZA V LJUBLJANI
FAKULTETA ZA DRUŽBENE VEDE

Anže Sendebah

Mentor: prof. dr. Vasja Vehovar
Somentor: prof. dr. Michael Bosnjak

Medijska večopravilnost v spletnih anketah
Media multitasking in web surveys

Doktorska disertacija

Ljubljana, 2017

Medijska večopravilnost v spletnih anketah

Povzetek

Medijska večopravilnost – sočasne in zaporedne kombinacije medijskih aktivnosti z drugimi medijskimi in nemedijskimi aktivnostmi – je vedno bolj razširjena oblika obnašanja v vsakdanjem življenju. Pri tem raziskave kažejo, da lahko medijska večopravilnost pomembno poslabša kakovost izvedbe ciljno usmerjenih aktivnosti. V kontekstu spletne metodologije so različni avtorji obravnavali tudi večopravilnost anketirancev v spletnih anketah – tj. medijsko večopravilnost, pri kateri se ena izmed aktivnosti nanaša na izpolnjevanje spletne ankete. Avtorji so zlasti zaskrbljeni nad tem, da bi lahko bila večopravilnost anketirancev povezana s slabšo kakovostjo odgovorov. Kljub tem pomislekom je empirična literatura na to temo relativno redka. Opisano področje je zato nadvse aktualno, dodatne raziskave pa so potrebne vsaj iz dveh razlogov. Prvič, spletne ankete so vedno pomembnejši vir podatkov v družboslovju. Drugič, raziskave kažejo, da je medijska večopravilnost pomembno vezana na razvoj, penetracijo in uporabo interneta in/ali modernih medijskih naprav (npr. osebni računalniki in pametni telefoni) – te tehnologije pa so neločljivo povezane tudi s spletnimi anketami.

Pričujoča disertacija naslavlja zgoraj opisano raziskovalno vrzel. V prvi fazi je vzpostavljen teoretični okvir, ki temelji na različnih znanstvenih tokovih: raziskave kognitivnih znanosti o večopravilnosti, interdisciplinarna literatura o medijski večopravilnosti in spletni metodologiji. Na podlagi teoretičnega okvirja identificiram dva ključna izziva raziskovanja večopravilnosti anketirancev.

Prvi izziv se nanaša na merjenje večopravilnosti anketirancev. Pri tem sta aktualna dva glavna pristopa: reaktivni in nereaktivni. Nereaktivni pristop v kontekstu disertacije temelji na parapodatkih – elektronskih sledeh, ki beležijo interakcijo med anketirancem in spletnim vprašalnikom. Reaktivni

pristop pa pomeni zanašanje na samoporočanja anketirancev, ki se merijo prek anketnih vprašanj o večopravilnosti med anketo. Glede na obstoječo literaturo identificiram ključne omejitve vsakega pristopa. Pri parapodatkih je glavna omejitev to, da je težavno meriti določene specifične oblike večopravilnostni anketirancev, pri samoporočanju pa obstaja znatno tveganje, da anketiranci ne bodo poročali o vseh primerih, ko so bili vpeti v druge aktivnosti med anketo.

Drug splošen izziv je raziskovanje odnosa med razširjenostjo in povezanimi dejavniki, pri čemer je zlasti pomemben odnos s kakovostjo odgovorov. Kakovost odgovorov se nanaša na tiste vidike kakovosti podatkov, ki so odvisni od individualnih razlik med anketiranci v smislu dojemanja, razumevanja in interakcije s spletno anketo. V sklopu teoretičnega okvirja ločujem med tremi ravnmi odnosa med večopravilnostjo in kakovostjo odgovorov. Prvič, kakovost odgovorov je lahko vzročno odvisna od večopravilnosti anketirancev. Drugič, večopravilnost anketirancev je lahko v medsebojno odvisnem odnosu s splošno vključenostjo – oziroma zavzetostjo izpolnjevanja – anketirancev. Tretjič, na podlagi empirične literature je mogoče tudi domnevati, da ljudje, ki so v vsakodnevnem življenju pogosteje vpeti v različne intenzivne oblike medijske večopravilnosti, v spletnih anketah podajajo odgovore slabše kakovosti ne glede na to, ali med anketo počnejo še kaj drugega ali ne.

Na podlagi navedenih dveh izzivov sem oblikoval štiri raziskovalna vprašanja in eno glavno hipotezo. Na kratko povzeta hipoteza se glasi: »Pristop s parapodatki poda podobne ključne ugotovitve o razširjenosti in dejavnikih večopravilnosti anketirancev ter o odnosu s kakovostjo podatkov kot pristop merjenja s samoporočanjem.« Štiri raziskovalna vprašanja, ki so povezana s hipotezo, pa se glasijo: kako so indikatorji večopravilnosti anketirancev, ki temeljijo na parapodatkih, povezani z indikatorji večopravilnosti anketirancev, ki temeljijo na samoporočanju; kakšna je razširjenost različnih oblik večopravilnosti anketirancev v spletni

anketi glede na indikatorje, ki temeljijo na parapodatkih in samoporočanju; kakšen je vpliv socio-demografskih in individualnih dejavnikov na razširjenost večopravilnosti anketirancev in kako so indikatorji večopravilnosti anketirancev povezani s kakovostjo odgovorov?

Ob izhajanju iz teoretičnega okvira in raziskovalnih ciljev sem razvil dva nova specifična postopka merjenja večopravilnosti anketirancev, ki temeljita na parapodatkih in samoporočanju. Pristopa sem nato uporabil v empirični študiji – 20-minutni spletni anketi, ki je bila izvedena na članih komercialnega spletnega panela. Vključeni so bili samo anketiranci, ki so izpolnjevali anketo prek osebnih računalnikov. Končni vzorec sestavlja 1.366 anketirancev, ki so po demografski strukturi (spol, starost, regija) podobni splošni slovenski populaciji, stari med 15 in 55 let.

Glavne vsebinske ugotovitve glede na meritve obeh postopkov so sledeče. Več kot polovica anketirancev je med anketo počela nekaj drugega, vendar pa je bilo zgolj 20 odstotkov vpletenih v bolj intenzivne oblike večopravilnosti anketirancev. Starost, navade glede vsakodnevne medijske večopravilnosti in pozornost (kot enemu izmed vidikov zavzetosti izpolnjevanja) so najpomembnejši dejavniki večopravilnosti anketirancev. Odnos med večopravilnostjo in kakovostjo podatkov je bil v splošnem šibek oz. v večini primerov ni bil potrjen.

Kar zadeva primerjavo med parapodatki in samoporočanjem, glavna hipoteza ni bila v celoti potrjena. Glede prisotnosti večopravilnosti anketirancev sicer oba pristopa podajata v grobem podobne ugotovitve, vendar se tukaj kažejo tudi pomembne razlike, zlasti pri identifikaciji specifičnih anketirancev, ki so ali niso med anketo počeli kaj drugega. Pri takšni identifikaciji se pristopa ujemata samo v 66 odstotkih primerov. Pristopa se razlikujeta tudi pri ugotovitvah glede vpliva dejavnikov večopravilnosti anketirancev (npr. starosti). Po drugi strani oba pristopa podajata podobne ugotovitve glede odnosa s kakovostjo odgovorov, pri

čemer parapodatki kažejo nekoliko bolj izrazite vsebinsko smiselne povezave.

Disertacija ponuja določene nove vpoglede tudi v širšem kontekstu raziskovanja medijske večopravilnosti (npr. v vsakdanjem življenju). Zlasti je vredno izpostaviti empirično ugotovitev, da ljudje, ki so vsakodnevno vpeti v intenzivne oblike medijske večopravilnosti, podajajo manj kakovostne odgovore na anketna vprašanja ne glede na to, ali so med izpolnjevanjem ankete vpeti v druge aktivnosti ali ne.

Sklenem lahko, da je disertacija razvila nove pristope za raziskovanje problema večopravilnosti v spletnih anketah ter s sistematično obravnavo izpostavila njihove omejitve in prednosti. Pomembno in na nov način je tudi osvetlila razmerje med večopravilnostjo in kakovostjo odgovorov v spletnih anketah.

Media Multitasking in Web Surveys

Abstract

Media multitasking – concurrent and sequential combinations of media activities with other media or non-media activities – is an increasingly prevalent behavior in everyday life. Research shows that media multitasking can hinder quality of performance in goal-oriented activities. Within the context of web survey methodology, different authors have elaborated on respondent multitasking – i.e. media multitasking where one of the activities is related to responding to a web survey. Authors are specifically concerned that respondent multitasking might be related to lower quality of responses. Despite these concerns, empirical literature on respondent multitasking in web surveys is relatively scarce. More research is warranted for two reasons. Firstly, web surveys are an increasingly important survey mode in social sciences. Secondly, research on everyday media multitasking shows that such behavior is notably driven by the Internet and modern media devices (such as personal computers and smartphones), i.e. technologies that are also inherent to web surveys.

This dissertation addresses this research gap. Firstly, I develop a theoretical framework on media multitasking in web surveys that is based on multiple streams of literature. These streams include cognitive sciences research on multitasking, interdisciplinary research on media multitasking, web survey methodology and an overview of published studies on respondent multitasking in web surveys. Based on the theoretical framework, I identify two general challenges in research on respondent multitasking.

The first general challenge is how to measure respondent multitasking where I focus on two main approaches: reactive and non-reactive data collection on respondent multitasking. Within the context of this dissertation, the non-reactive approach is based on paradata – electronic

tracks about respondents' interaction with the web survey. Whereas the reactive approach means relying on respondents' self-reports that are collected via questions prompting them about their multitasking behavior during a survey. Based on existing literature, I identify the main limitations of each approach. Namely, with paradata it is problematic to observe all types of respondent multitasking behavior, while self-reports can be severely affected by underreporting. An overview of literature also shows that so far only a small number of studies have used non-reactive approaches for research on media multitasking (which includes respondent multitasking).

The second general challenge is related to getting more insight on prevalence and factors of respondents multitasking and especially its relationship with response quality. Response quality refers to those aspects of data quality that are influenced by respondents' individual differences in perception, understanding, and interaction with a questionnaire. Generally, I propose that the relationship between multitasking and response quality can manifest on three levels. Firstly, quality of responses can be hindered directly due to respondent multitasking. Secondly, respondent multitasking can be in an interdependent relationship with the respondents' overall survey engagement, which can also worsen response quality. Thirdly, since empirical literature on everyday media multitasking behavior shows that people who are habitually engaged in intensive forms of such behavior (i.e. high media multitaskers) perform worse in cognitive tests that require a sustained amount of attention, I propose that high media multitaskers could also have a lower response quality, regardless if they are multitasking during a survey or not.

Based on these two challenges, I have developed four research questions and one main hypothesis. To briefly sum up the hypothesis: I state that the paradata-based approach can give general findings on prevalence and factors of respondent multitasking as well as on its relationship with

response quality that are comparable to the approach based on self-reports. The four research questions are related to the hypothesis and can be summarized as follows: How are specific paradata observations of respondent multitasking associated with self-reports? What is the prevalence of different types of respondent multitasking? What are the socio-demographic and individual factors of respondent multitasking? What is the relationship between respondent multitasking and response quality indicators?

Following the theoretical framework and research objectives, I have developed new implementations of approaches based on paradata and self-reports. These new approaches are used in an empirical study – a 20-minute web survey conducted on respondents of a commercial Internet panel. Only respondents who used a personal computer to complete the survey are included in the analysis. The final sample includes 1,366 respondents and mimics the general Slovenian population (15-55 years old) in terms of age, gender and region.

The main findings based on both data approaches are as follows: More than half of respondents have multitasked at least once during a survey. However, only about 20% of the sample was engaged in more intensive types of respondent multitasking. Age, everyday media multitasking habits and attention (which can be understood as one of the core elements of respondents' engagement in the survey) were all important factors of respondent multitasking behavior. The relationship between indicators of respondent multitasking and response quality was weak to non-existent.

With regards to comparison between paradata and self-reports, the main hypothesis was not fully supported. In terms of prevalence, both approaches gave similar figures on overall presence of respondent multitasking in the survey. However, there were important differences between them. A typical problem was identifying specific respondents who

had at least one instance of respondent multitasking. Here, the data collection approaches yielded the same results only in 66% of cases. Moreover, approaches gave some similar but also some notably different findings in relation to the factors of respondent multitasking (e.g. age). In terms of the relationship with response quality, there were no notable differences between the approaches; however, paradata indicators did produce slightly more meaningful results.

This dissertation also includes an overview of limitations and directions for further research. Overall, the dissertation contributes new methodological and substantive findings to the scarcely researched topic on the role of media multitasking in web surveys. One of the main contributions is the development of new reactive and non-reactive methods for measuring respondent multitasking as well as a systematic comparison between them. Moreover, the dissertation also expands on the relationship between multitasking and response quality with an overarching theoretical elaboration as well as with new empirical evidences.

In a broader context of media multitasking research, the dissertation also provides new empirical findings on the role of everyday media multitasking habits; here empirical findings indicate that high media multitaskers are more likely to have a lower response quality regardless of whether they are multitasking during the survey or not.

In conclusion, the dissertation develops new approaches for researching respondent multitasking in web surveys and systematically highlights their limitations and advantages. Moreover, it also elaborates on new aspects of the relationship between multitasking and response quality in web surveys.

Contents

Index of tables

Index of figures

List of abbreviations

P_ALL	frequency of all focus-out and time-out events per respondent
P_ALL Rs	respondents who have at least one focus-out or time-out event.
P_FOCUS	frequency of all focus-out events per respondent
P_FOCUS Rs	respondents who have at least one focus-out event
P_TIME	frequency of all time-out events per respondent
P_TIME Rs	respondents who have at least one time-out event
S_ALL	reflects whether respondent has provided at least one self-report on respondent multitasking in either Q1 or Q2
S_ALL Rs	respondents with at least one self-report on respondent multitasking in either Q1 or Q2
S_RM	range of all (media and non-media) activities
S_RM Rs	respondents who reported on at least one activity
S_RM_MM	range of media activities
S_RM_MM Rs	respondents who reported on at least one media activity
S_RM_NM	range of non-media activities
S_RM_NM Rs	respondents who reported on at least one non-media activity
S_SEQ_ALL	number of all interruptions
S_SEQ_ALL Rs	respondents with at least one self-reported interruption

S_SEQ_NM	number of self-reported non-media interruptions
S_SEQ_NM_Rs	respondents with at least one self-reported other-device interruption
S_SEQ_OD	number of self-reported other-device interruptions
S_SEQ_OD_Rs	respondents with at least one self-reported other-device interruption
S_SEQ_SD	number of self-reported same-device interruptions
S_SEQ_SD_Rs	respondents with at least one self-reported same-device interruption

1 Introduction

People are increasingly multitasking in their everyday lives, particularly with modern media devices. Activities on personal computers, smartphones and tablets are often combined with each other as well as with other media and non-media activities. Literature shows that such behavior can have an important relation to performance in goal-oriented activities.

In this dissertation I will deal with specific media multitasking in the context of web survey methodology. While placed in a broader perspective of multitasking and media multitasking literature, the focus is on measuring how people are multitasking while responding to web surveys on personal computers and whether this is related to the quality of their responses.

I will first introduce the key concepts, the state of the research, and the main background of this dissertation (Section 1.1). Next, I will present the aims of the dissertation (Section 1.2). Finally, I will provide an overview of the research design and methods, explaining how they are reflected in the structure of this dissertation (Section 1.3).

1.1 Background

In Section 1.1 I will first provide a general background on multitasking and media multitasking. Particular emphasis will be given on the methodological challenges related to these phenomena. Secondly, I will provide an overview of the current respondent multitasking literature. Here, particular emphasis will be given on respondent multitasking in web surveys, which is the main focus of this dissertation. All theoretical concepts introduced in Section 1.1 are further elaborated and expanded in the theoretical framework (Chapter 2).

1.1.1 Multitasking and media multitasking

"If you chase two rabbits, you will not catch either one."

This popular quote is well over two thousand years old. Various authors attribute its origins to different cultures, from Russia to the Roman Empire and China. In any case, it seems civilizations understood very early that multitasking is somehow bad for human performance.

In 2017, this message resonates through our society stronger than ever. The rabbits we are chasing have partly escaped from the analog to the digital world. New prominent multitasking practices with media have emerged, commonly referred to as *media multitasking* (Foehr 2006; Ophir et al. 2009; Wallis 2010).

Scientific research is moving along with these trends. Cognitive scientists produced the first examples of research on multitasking in the 1930s (Meyer and Kieras 1997). In the recent decades, the advent of media multitasking has triggered a considerable increase of multitasking literature across scientific disciplines (Spink et al. 2008).

On the theoretical level, authors have provided several different conceptualizations of multitasking, as well as different models on how humans are (un)able to multitask (e.g. Meyer and Kieras 1997; Wickens 2008; Salvucci and Taatgen 2011). Empirical literature shows that multitasking is increasingly present in our everyday lives. Modern devices, such as personal computers (PCs), smartphones and tablets, have been particularly important technological drivers of such behavior (e.g. Rideout et al. 2010; Voorveld and van der Goot 2013; Ofcom 2016). While some multitasking occurs due to involuntary reasons (e.g. due to the nature of work), research reveals that people often start to combine multiple activities voluntarily in order to be entertained and/or as "a response to an emotional state such as loneliness or boredom [of being engaged in a single particular activity]" (Baron 2008, 7; see also Jeong and Fishbein 2007; Hwang et al. 2014; Kononova and Chiang 2015). Some authors

even report on addiction or "chronic media multitasking", i.e. "habitual, routine media multitasking, which goes beyond […] control" (Kononova and Chiang 2015, 50; see also Baron 2008; Ophir et al. 2009; Bardhi et al. 2010).

A substantial amount of research, as I fully elaborate in Section 2.2.4, confirms that multitasking does hinder our performance or is at least associated with it. At the very least, multitasking prolongs the time needed to complete tasks. However, it can lower the quality of performance in other aspects too. This has been observed in a variety of settings, ranging from laboratory experiments to in-situ studies; from short-term to long-term associations; from driving to classrooms, working environment, and living rooms (e.g. Wickens 2008; Salvucci and Taatgen 2011; Junco and Cotten 2012). Several authors note that media multitasking could also have a latent relationship with performance. An influential study by Ophir et al. (2009) reported that chronic media multitaskers perform worse in activities that require attention and filtering of distraction, even if they are not multitasking during these activities. This has been confirmed later by several other studies (e.g. Cain and Mitroff 2011; Ralph et al. 2014; Uncapher et al. 2016). However, as several authors stress, we should be careful in making claims that multitasking has inevitable direct or latent negative effects, since the existence and strength of these relationship depends on numerous factors (e.g. Ophir et al. 2009; Wallis 2010; Salvucci and Taatgen 2011).

There is another important reason why we should not be too quick to assume that multitasking is in a general negative causal relationship with the quality of human performance as many of studies on this topic have been conducted in experimental laboratory settings. While several studies in more applied settings have also reported on strong associations between multitasking and quality of performance, proving a causal relationship in such settings is of course always a demanding research task. In the case of multitasking, this presents an additional

methodological challenge because multitasking involves a complex and multifaceted set of behaviors (e.g. Ophir et al. 2009; Wallis 2010; Salvucci and Taatgen 2011).

Overall, multitasking presents a methodological challenge for social science research on the following two levels:

1. The first level consists of methodological issues that are relevant for researchers who explicitly investigate multitasking and/or media use. In this context, the challenge is to properly define, conceptualize, measure, and analyze the complex set of human behaviors that constitute (media) multitasking (e.g. Rideout et al. 2010; e.g. Wallis 2010; Salvucci and Taatgen 2011).

   One specific methodological challenge here involves selection, development, and utilization of data collection methods for (media) multitasking research. They can be broadly divided into *reactive* and *non-reactive* approaches (Lee et al. 2008). The former relates to methods where an active participation by research subjects is required in a sense that they provide information on their multitasking behavior guided by specific prompts from researchers (e.g. surveys, focus groups, qualitative interviews). The latter relates to methods for which no such active participation of research subjects is required. Instead, information on multitasking behavior is obtained either via direct, in-situ observation or via electronic tracks that log activities on media devices.

   Generally, both data collection approaches (i.e. reactive and non-reactive approach) have specific advantages and disadvantages that can importantly influence research findings on multitasking behavior. However, several authors (Greenberg et al. 2005; Wallis 2010; Möller et al. 2013) have stressed that reactive approaches should be more prominently used in multitasking and/or media use research; in particular, because studies have shown that people tend to severely

under-report on how often they combine different media activities (e.g. Kaye and Sapolsky 1997; Iqbal and Horvitz 2007; Brasel and Gips 2011).

2. The second level consists of challenges that are relevant for a broader context of social science research (i.e. not only for studies that explicitly investigate multitasking and other related phenomena), particularly for studies that use self-administered data collection methods. One frequently-mentioned advantage of such collection methods is that participants typically have control over the time, location, and pace of their partaking in the study (e.g. De Leeuw 2008; Bethlehem and Biffignandi 2011; Callegaro et al. 2015). However, this advantage could turn into a disadvantage if participants engage in other activities that are not related to the study. Drawing on the general multitasking literature, it is reasonable to assume that such multitasking behavior could hinder the quality of their participation in the study and possibly increase the measurement error.

   Moreover, multitasking could also indirectly endanger data quality in social science research. As described earlier, several studies have shown that people who are intensively engaged in everyday media multitasking have a general problem with focusing on a single task. These problems may affect their ability to focus sufficiently on participation in social science research.

   While this direct and indirect relationship between multitasking and data quality has always existed for self-administered and other data collection methods (as I elaborate in Section 2.3.4), reports on progressively larger prevalence of everyday media multitasking in the last two decades indicate that this could become an increasingly important methodological challenge for social science research.

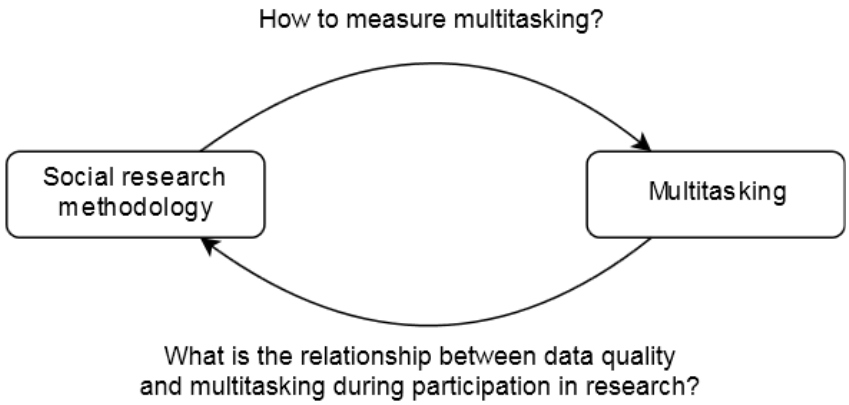
In Figure 1.1, I broadly summarize the relationship between social science research methodology and multitasking. Considering the close interaction

and the two levels of challenges, it is somewhat hard to explain why the scientific field of survey methodology is rarely considerably involved in multitasking research. Nevertheless, I believe a greater involvement is necessary for the following two reasons:

- First, this field of applied statistics can help to address the first level of methodological challenges by providing supplementary knowledge and tools related to measuring and analyzing multitasking with surveys, one of the most common data collection methods in social science research.

- Second, considering the second level of methodological challenges, it is important for survey methodology to investigate whether people multitask while responding to surveys and whether such behavior is related to a larger survey measurement error. For example, engagement in other activities during a survey could mean that respondents do not dedicate enough resources in order to fully understand the questions and/or formulate their responses.

Figure 1.1: Relationship between social science research methodology, multitasking and related challenges



This dissertation is primarily focused on the described second level of methodological challenges. However, it is my belief that some parts of this work will be also useful in the future for addressing the challenges related to the first level.

1.1.2 Respondent multitasking in web surveys

Due to the emergence of media multitasking, scholars have lately called for more research on *respondent multitasking* (e.g. Holbrook et al. 2003; Lynn and Kaminska 2012; Callegaro et al. 2015). Respondent multitasking occurs when a respondent is engaged in other, *secondary activities* while responding to a survey that is considered the *primary activity*. Within the context of this dissertation I understand respondent multitasking as a specific form of media multitasking.

Considering that responding to a survey is a cognitively demanding activity (Holbrook et al. 2003), some respondents might engage in secondary activities in order to seek entertainment during the survey or simply because such behavior is a prominent part of their everyday routine. Of course, there can be also other, less voluntary reasons such as receiving a phone call or taking care of a child.

Several authors have noted that respondent multitasking could be especially prevalent in web surveys where it is typically harder to assure that respondent are sufficiently engaged throughout the survey (e.g. Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Callegaro et al. 2015). Respondents are participating in web surveys via modern media devices such as desktop and portable PCs, smartphones and tablets. As mentioned in Section 1.1.1, research shows that these devices are an important driver of everyday media multitasking.

The main concern over respondent multitasking is its proposed relationship with *response quality* (e.g. Holbrook et al. 2003; Ansolabehere and Schaffner 2015; Callegaro et al. 2015), i.e. those aspects of data quality that are influenced by respondents' individual differences in perception, understanding, and interaction with a questionnaire (Ganassali 2008). In order to provide an optimal, high quality response to a survey question, a survey participant needs to properly complete all stages of the *response process*, i.e. fully

comprehend the question, retrieve and consider the specific information requested by the question, and finally map the judgment to available response options (Tourangeau et al. 2000). Thus, the response process is directly related to response quality.

Based on the general multitasking and survey methodology literature, I suggest that the relationship between multitasking, response process and response quality could be manifesting on three, not mutually exclusive levels.

The first two levels relate to the relationship between respondent multitasking and response quality:

1. Survey methodologists mainly write about a direct causal relationship between respondent multitasking and the response process (e.g. Holbrook et al. 2003; Kennedy 2010; Lynn and Kaminska 2012). This stream of research can be found in numerous studies from cognitive sciences showing that secondary activities generally have negative effects on the quality of performance of primary activities (e.g. Meyer and Kieras 1997; Wickens 2008; Salvucci and Taatgen 2011). Hence, secondary activities during the survey could disturb the necessary cognitive processes needed to adequately complete all stages of the response process. Consequently, this results in a response of lower quality.

2. Multitasking during a survey can also be positively associated with the overall respondents' *engagement* in the survey. In this context, engagement refers to how respondents are focused and motivated to respond to a survey; and a lower respondent's engagement could lead to a non-optimal response process and consequently lower response quality (Callegaro et al. 2015).

Media multitasking behavior could also be in a relationship with response quality. This can be understood as a third, latent level of the general relationship between multitasking and response quality:

3. Namely, drawing on the literature on media multitasking, it is possible to suggest that intensive everyday media multitasking behavior could also be related to response quality. Chronic media multitaskers have problems with focusing on a single activity (Ophir et al. 2009). This could translate into their general ability to remain sufficiently attentive during a response process. When compared to other respondents, chronic media multitaskers could thus produce responses of lower quality, regardless of whether they are actually engaged in secondary activities during the survey.

In light of these concerns, empirical research on respondent multitasking in web surveys is relatively scarce. In 2014, in the proposal of this dissertation, I underscored that there is practically no published empirical study focused on web surveys, while there has already been some research done on telephone surveys (e.g. Pew Research Center 2006; Kennedy 2010; Lavrakas et al. 2010). Since then, to the best to my knowledge, only three papers have been published that focus on respondent multitasking in web surveys (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Sendelbah et al. 2016). In addition, I also found one dissertation with this topic (Antoun 2015).

At the current state of research on respondent multitasking in web surveys, scarce literature provides some answers, but also raises new questions. Reported shares of respondents who are engaged in secondary activities varies considerably: from 22% in Ansolabehere and Schaffner (2015) to 62% in Sendelbah et al. (2016). Two studies report that younger respondents are more likely to multitask during surveys (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015). Other than that, the empirical evidence investigating the relationship between respondent multitasking, other factors and response quality is inconclusive. There are several reasons for these discrepancies. For instance, studies differ in terms of questionnaires (e.g. their length and topics), samples and

methodological approaches for conceptualizing and measuring respondent multitasking.

Overall, the existing studies (on web surveys and other survey modes) have used a variety of reactive and non-reactive methods for measuring respondent multitasking. In relation to that, it is notable that so far no study has measured respondent multitasking with different methods (e.g. either by using two different reactive methods or by using both, reactive and non-reactive approaches). Consequently, there has been no triangulation or validation of methods for measuring respondent multitasking.

Many of the published studies on respondent multitasking have used the reactive approach. What all studies have in common is that they measure respondent multitasking via *self-reports* of respondents who respond to questions on respondent multitasking at the end of the survey. However, question design varies greatly from study to study and there is little to no discussion on how to design and analyze these questions, even though this could importantly influence research outcomes (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Antoun 2015).

An alternative, non-reactive approach in the context of web surveys is possible with *paradata*. In this dissertation, paradata refers to electronic tracks about the process of surveying, generated by the respondent's interaction with the web questionnaire and automatically collected by the server of the survey platform (Callegaro 2013). While some studies (e.g. Stieger and Reips 2010; Ansolabehere and Schaffner 2015; Sendelbah et al. 2016) have explored how certain types of paradata tracks are related to respondent multitasking, so far this approach has not been fully elaborated and utilized. As in the case of self-reports, there is scarce discussion on how to analyze paradata within the context of respondent multitasking research; and in case of paradata such discussion is perhaps

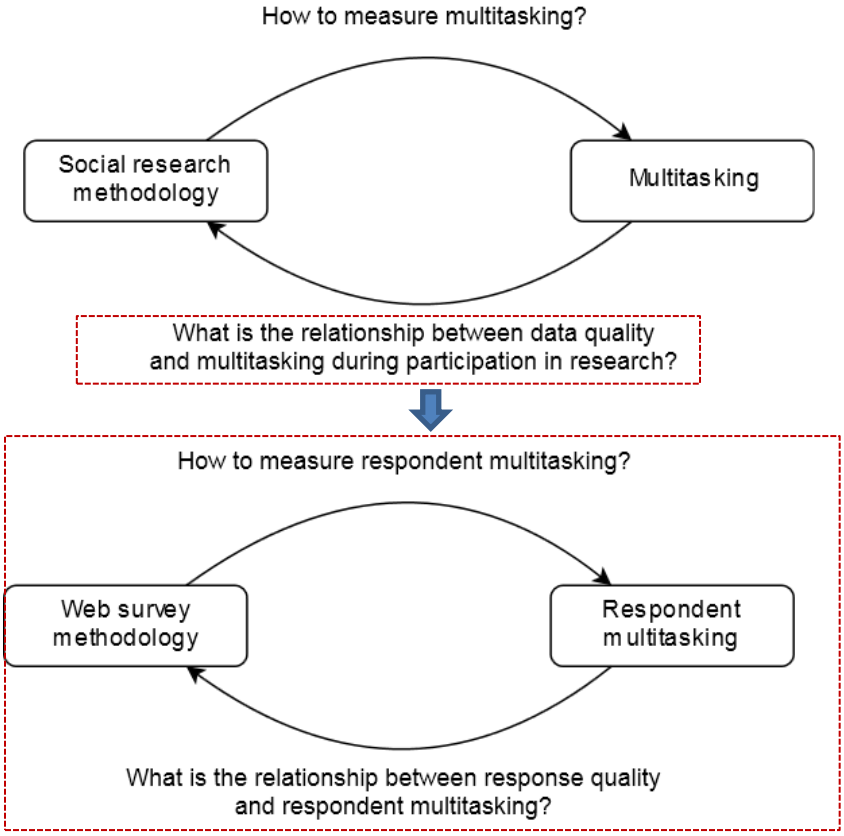
even more warranted since it is a relatively new data source for web survey methodology in general.

Following Callegaro's (2013) taxonomy of paradata, there are two types of paradata events that have been used for research on respondent multitasking. The first are *response times*, i.e. the time respondents spend to complete an item, question, page or the whole questionnaire. Researchers can define thresholds and response times over such thresholds can be considered as indicators of respondent multitasking. I will refer to such events as *time-out* events. The second group of paradata events relates to *inactivities* within the survey, such as looking at longer periods when there is no activity from input devices (e.g. mouse, keyboard or touchscreen) or whether respondents have switched away from the browser window/tab that contains the web questionnaire. I will refer to the latter as *focus-out* events.

At present, paradata has an untapped potential for measuring respondent multitasking in web surveys. As noted in Section 1.1.1, studies have shown that people tend to under-report their media multitasking behavior and it is reasonable to assume that this problem extends to the specific case of respondent multitasking. The paradata approach is not affected by these issues and offers several other important advantages. For example, it does not increase the respondents' burden, as no additional questions need to be included into the questionnaire. Moreover, with paradata we can observe the multitasking behavior of all respondents, while self-reports lack information from those respondents who did not respond to questions regarding respondent multitasking (either because they chose not to respond or because they dropped out of the survey before seeing these questions). Lastly, once the approach is developed, it can be employed across multiple surveys with relatively small resources compared to self-reports. In fact, it can even be employed retroactively for any survey, as long as its paradata is accessible.

Nevertheless, self-reports can – due to their methodological nature – offer more qualitative insight into respondent multitasking. For example, with paradata, it is impossible to get specific insight into what kinds of secondary activities respondents have been engaged in. This information can be very important for exploring the relationship between respondent multitasking and response quality, given that general multitasking research indicates that the characteristics of secondary activities influence the quality of performance of the primary activity. Moreover, self-reports can generally account for a larger variety of secondary activities. Paradata cannot observe secondary activities that do not result in prolonged response times or different types of inactivities (i.e. either by inactive input devices or by changing the focus from the window or tab with the web questionnaire).

Figure 1.2: Interaction between social science methodology and multitasking (upper figure) and web survey methodology and respondent multitasking (lower figure)



In sum, several authors have actually addressed the methodological challenges regarding the data quality of social science research posed by

increasing multitasking, and some of them have also done that within the context of web surveys. However, the relatively scarce existing empirical research does not provide many conclusive or comparative findings. One of the reasons for that stems from the methodological challenges related to measuring respondent multitasking in web surveys (see Figure 1.2, which expands the issues shown in Figure 1.1). Namely, a mutual interaction exists between social science methodology and multitasking, which complicates any empirical research. A similar interaction with similar challenges is then also present when we narrow the focus on respondent multitasking in web surveys.

1.2 Objectives

I will now define the research problem and provide a description of general objectives, which will then be outlined into four research questions and a hypothesis. These all will be also put in the broader context.

A detailed elaboration of these objectives depends on numerous concepts that need to be thoroughly addressed first, which I will do when developing the theoretical framework in Chapter 2 . Therefore, the research questions (along with corresponding subquestions) and the hypothesis, which I preliminarily outline here, will be further elaborated in Chapter 3 , after conceptual discussion in Chapter 2.

Overall, this dissertation investigates how media multitasking behavior reflects web survey methodology. The main focus is then further narrowed down to respondent multitasking in web surveys. Following suggestions from general media multitasking literature, this could be an increasingly prevalent respondent behavior that is potentially related to lower response quality and higher measurement error. However, as elaborated in Section 1.1, at present, the literature on respondent multitasking is scarce and there are several notable conceptual and empirical gaps.

Based on this, I define the **initial research motivation of the dissertation** as the following: *more research on respondent multitasking in web surveys is warranted in order to determine appropriate methodological measurement approaches, related characteristics and factors, the prevalence of such behavior, as well as the relationship with response quality.*

One the other side, the **initial theoretical objective** of this dissertation is to provide a *systematic integration of knowledge from multitasking and media multitasking literature into the specific context of respondent multitasking* and also to provide a systematic *overview of existing literature on respondent multitasking in the context of web surveys*.

While existing studies on respondent multitasking already draw from general multitasking literature, I will attempt to provide a more extensive elaboration and connect key concepts that exist across the multidisciplinary field of research on different forms of multitasking. In this dissertation the theoretical framework particularly provides guidance on identifying crucial research gaps in the existing literature on respondent multitasking in web surveys; conceptualization and operationalization of respondent multitasking; development of measurement and analytical procedures; interpretation of empirical findings.

I should also articulate the **initial empirical objectives** of this dissertation that are linked to the *developments related to two separate data collection approaches of respondent multitasking – one based on paradata and the other on self-reports* – as well as the *application and comparison of these approaches in an empirical study of respondent multitasking, response quality, and other related factors*.

Following the above motivation, objectives and initial research challenges (Section 1.2) I can outline the preliminary research questions.

1. Paradata approach is based on focus-out and time-out events. Self-reports are obtained based on questions that are designed according to multitasking literature and the specific empirical objectives of this dissertation. Both approaches are used in an empirical study described in the empirical part (Chapter 4). The development of the data collection approach and the empirical study is limited to respondents who used a PC to participate in the survey. However, with some modifications, the developed data collection approaches are also applicable to respondents on other devices. A particular emphasis is given to the paradata approach. On the one hand, this approach has not been fully utilized in current respondent multitasking research, particularly because paradata is not expected to be affected by as many limitations as self-reports (e.g. respondent bias, respondent burden, measurement errors). On the other hand, self-reports can account for a wider variety of secondary activities and offer a greater qualitative insight into respondent multitasking behavior. Since no study thus far has compared the two approaches in the context of research on respondent multitasking, I can formulate the initial research questions and the hypothesis.

   Research question 1. How are paradata indicators of respondent multitasking associated with indicators that are based on self-reports?

The other three research questions relate to the characteristics of respondent multitasking in the empirical study, as observed with indicators from both approaches.

2. One of the most basic questions is how prevalent is respondent multitasking. While all published studies on respondent multitasking in web surveys have addressed this question, authors differ on how they report on these findings. For example, most of the studies (but not all) report on the overall prevalence of respondent multitasking, i.e. the

share of respondents who have multitasked during the survey (Ansolabehere and Schaffner 2015; Sendelbah 2015; Antoun 2015; cf. Zwarun and Hall 2014). Other findings include reports on how many different activities respondent have been engaged in (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Antoun 2015), what the most common secondary activities are (Ansolabehere and Schaffner 2015; Zwarun and Hall 2014), how many times respondents have interrupted the survey due to secondary activities (Zwarun and Hall 2014; Sendelbah 2015), etc. This dissertation will provide a comprehensive report on the prevalence of respondent multitasking that will reflect on all above-mentioned findings and also address additional issues. I can thus formulate the following research question:

Research question 2. What is the prevalence of respondent multitasking?

3. General literature on multitasking shows that numerous factors influence this behavior in terms of how often and intensive it happens, as well as what activities are involved (e.g. younger people tend to multitask more, especially with media activities) (e.g. Foehr 2006; Jeong and Fishbein 2007; Kononova and Chiang 2015). Moreover, general survey methodology literature (e.g. De Leeuw 2008; Bethlehem and Biffignandi 2011; Callegaro et al. 2015), as well as specific studies on respondent multitasking (e.g. Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Antoun 2015), also offer some insight into what factors could be particularly important for multitasking during web surveys. Drawing from both literature streams, this dissertation attempts to provide a new empirical insight into factors of respondent multitasking. At the current state of research on respondent multitasking, these factors are also particularly valuable for providing additional insight into other research questions. Within this context I will mainly be focused on socio-demographic (e.g. age, gender) and individual (e.g. everyday media multitasking habits,

experiences with surveys) factors. The corresponding outline of the research question is as follows:

Research question 3. What are the socio-demographic and individual factors of respondent multitasking?

4. The last research question relates to the main concern regarding respondent multitasking: its relationship with response quality. I investigate two aspects of response quality: non-response behavior (specifically, for how many questions or items a respondent did not provide a response) and satisficing (i.e. specific characteristics of responses that indicate a respondent was not optimally engaged in the response process). In Section 1.1.2, I have elaborated that the relationship between respondent multitasking and response quality could be manifesting on two levels. First, secondary activities could directly interfere with the response process. Researching such causal relationships requires an experimental, controlled research setting and is not possible with the research design used in this dissertation. Second, respondent multitasking could be in an indirect relationship with the response process through an association with the respondents' engagement. In this dissertation, I will explore the relationship between respondent multitasking and response quality on this level. To some degree, all published studies on respondent multitasking in web surveys (as well as for some other survey modes) have investigated the indirect relationship between respondent multitasking and response quality. However, research findings typically show that respondent multitasking is at most marginally related to worse response quality. In contrast with existing literature, I attempt to provide a more elaborate analysis in terms of variety of respondent multitasking and response quality indicators, as well as by accounting for a wide array of factors that could influence this relationship. One of these factors includes everyday media multitasking behaviors. The final research question is thus as follows:

Research question 4: What is the relationship between respondent multitasking and response quality?

5. As already mentioned, particular emphasis in the methodological and empirical part of this dissertation is given to the paradata approach. Despite having several limitations relative to self-reports, my suggestion is that paradata can provide similar findings regarding the most crucial aspects of research on respondent multitasking. This does not refer only to the direct comparison of indicators from both approaches as expressed in Research question 1, but it also encompasses topics contained in Research questions 2, 3, and 4. Therefore, the hypothesis of this dissertation is as follows:

   Hypothesis: Paradata-based procedures can measure respondent multitasking in web surveys and provide similar general findings on the prevalence of multitasking behavior, its factors and relationship with response quality, as the approaches based on self-reports.

While the hypothesis encompasses all four research questions, it is less than the sum of its parts. In other words, answers to the research questions are not limited only to obtaining the necessary information for the hypothesis. Instead, in answering research questions I attempt to provide a comprehensive elaboration on prevalence and factors of respondent multitasking, as well as its relationship with response quality in web surveys.

To summarize, the work required to address the above-structured research objectives and the hypothesis to some degree interacts with both challenges in web survey methodology that are related to respondent multitasking (see Figure 1.2). First, two data collection approaches are developed, used and compared which will provide new findings on how to measure respondent multitasking. Here, the hypothesis specifically states that I expect the seemingly underutilized paradata approach is as viable as self-reports, which are otherwise prevailing in research. Second, new

findings on the relationship between respondent multitasking and response quality are also expected to be found, along with detailed reports on the prevalence of different types of respondent multitasking behavior as well as its factors. Overall, as already mentioned, in a broader context of social science research methodology (see Figure 1.1), this dissertation addresses the challenge of how multitasking relates to data quality within the context of web surveys (Figure 1.2).

In conclusion, while the empirical objectives of this dissertation are fixed on respondent multitasking in web surveys, the background is noticeably broader and interdisciplinary. As already explained, the main objective of such theoretical framework is to provide guidance on several crucial stages of theoretical, methodological, and empirical work on this dissertation. Moreover, I hope that such theoretical framework will also establish that some contributions of this dissertation, focused on respondent multitasking in web surveys, are also valuable for other contexts of related research. These contexts include:

- General media multitasking research. An important part of the theoretical framework is an overview of media multitasking literature, with a particular emphasis on conceptual, methodological and empirical aspects. I hope that this overview will be a valuable contribution to an interdisciplinary discussion (e.g. Spink et al. 2008; Wallis 2010; Salvucci and Taatgen 2011) on how to measure media multitasking (see Section 1.1.1).

- Research on everyday media multitasking. This dissertation will investigate the relationship between everyday media multitasking behavior and response quality in a web survey. This can be understood as a small new contribution to an existing body of knowledge on how media multitasking behavior is related to the quality of performance in goal-oriented activities (e.g. Ophir et al. 2009; Cain and Mitroff 2011; Ralph et al. 2014).

- Comparative studies on the use of reactive and non-reactive approaches in research on media behavior. Several authors (e.g. Greenberg et al. 2005; Wallis 2010; Möller et al. 2013) have called for more studies that would use non-reactive approaches in media use and/or media multitasking research; and also compare it with self-reports. Results from this dissertation could be a welcome addition to the scarce empirical studies that explore this topic.

- Analysis of web survey paradata. Several authors have called for "more published reports describing the utility of paradata across a variety of survey applications" (West 2011, 5; see also WEBDATANET 2010; Lynn and Nicolaas 2010; Callegaro et al. 2015). The entire potential of paradata is yet to be discovered and one of the main reasons for this is probably that analysis of paradata is relatively demanding in the context of social sciences. This dissertation will present and elaborate new ways of using paradata to gain insight into the multitasking behavior of web survey respondents.

1.3 Structure of the dissertation

The remainder of this dissertation is split into four chapters. In the theoretical part (Chapter 2), I discuss key concepts and related literature on multitasking (Section 2.1), media multitasking (Section 2.2), and respondent multitasking (Section 2.3). The perspective and focus thus get narrower with each section of this Chapter. In addition to the theoretical and conceptual overview, sections on media multitasking and respondent multitasking also contain a literature overview of key empirical findings on the prevalence of such behavior, its factors, and its relationship with performance. The Section 2.4 concludes with a summary discussion and the identification of key research gaps.

In Chapter 3, which follows the conceptual elaboration, I provide a detailed discussion on theses and research questions related to my

empirical work. The discussion is driven by the integrated understanding that is developed throughout the Chapter 2 , but also as an extension and upgrade of the research questions and the hypothesis already outlined in Section 1.2.

The empirical part is presented in Chapter 4 and is split into three main sections dedicated to procedures, methods and research design (Section 4.1), results (Section 4.2), and discussion (Section 4.3). The core of the empirical work is dedicated to the development of two distinct strategies for measuring respondent multitasking (paradata and self-reports) and using them in an empirical study in order to address the research objectives.

The overview of research design (Section 4.1) contains the description of the study that was based on a web survey, which was designed and conducted specifically for the purpose of this dissertation. The survey includes various specific questions required to elaborate the research questions and the hypothesis, from self-reporting of multitasking to questions related to satisficing. The questionnaire of around 20 minutes was applied to Slovenian respondents from the askGfK online panel. While the panel does not employ probability sampling, respondents included in the analysis generally mimic general Slovenian population between 15 and 55 year olds in terms of age, gender and region, and otherwise presents a standard tool for social and marketing research in Slovenia. The final sample includes 1,366 respondents. We may add that only respondents on PCs were allowed to participate in the survey to avoid potential interactions with mobile devices.

Both measurement approaches (reactive and non-reactive approach), conceptually elaborated already in Chapter 2, are then addressed within this specific implementation.

The specific paradata approach used in the empirical part is described in details in Section 4.1.2. It is in large part determined by the availability of

paradata events provided by the survey platform. In this study, the survey is formally built and hosted on the 1KA open source service for online surveys (although the recruitment was done within an access panel) (EnKlikAnketa - 1KA Web Surveys 2016). 1KA service generally supports a wide array of paradata events that can be collected without any additional software or hardware modifications on respondents' devices. However, the development of the paradata procedure requires a considerable amount of computational and statistical work since paradata input files come in an unstructured format with a substantial amount of records. This procedure is described along with the two obtained paradata events that are used to observe respondent multitasking: *time-out* (as determined via page response times) and *focus-out events* (as determined by whether the respondent switched from the web questionnaire to another internet browser tab or any type of window on his PC)

Measurement via self-reports is described in Section 4.1.3. It is based on two questions that are included at the end of the web questionnaire. These questions are partially based on one of the most prominent questionnaires used in general media multitasking research (Ophir et al. 2009). Furthermore, the design of these questions is also partially influenced by an attempt to provide some comparability of self-reports with the collected paradata measurements.

The next important methodological step has been the construction and computation of specific indicators that measure respondent multitasking. I present and conceptually compare the indicators on respondent multitasking obtained with both approaches in Section 4.1.4.

The other key set of indicators is related to response quality and is elaborated in Section 4.1.5. These indicators are used to assess the respondents' performance in web surveys. Namely, I am interested in two components of response quality that are commonly investigated in survey methodology: non-response behavior and satisficing.

The questionnaire has been designed specifically for the purposes of this dissertation. In addition to questions dedicated to measuring respondent multitasking, other questions fulfill at least one of the two purposes: to measure different aspects of response quality or to gather information on factors that could influence respondent multitasking, response quality or the relationship between them (e.g. survey questions about socio-demographic factors, respondents' everyday media habits, experiences with surveys etc.).

The results (Section 4.2) are organized in the same manner as the overview of key empirical findings on media multitasking (Section 2.2.4) and respondent multitasking (Section 2.3.4) in the theoretical part: prevalence (Section 4.2.1), factors of prevalence (Section 4.2.2), and the relationship with performance, i.e. response quality in the context of web surveys (Section 4.2.3). While Section 4.2 focuses on reporting the results, some brief discussion and elaboration is needed in some parts; especially when analysis takes an exploratory approach. Overall, the statistical tools in this section use descriptive (Section 4.2.1) and confirmatory modeling approaches (Sections 4.2.2 and 4.2.3) for addressing specific research questions. In the latter case, various regression approaches are used.

Chapter 4 ends with Section 4.3 that provides a detailed discussion of the research questions and the hypothesis in light or the empirical findings.

In the conclusion of the hypothesis (Chapter 5), I summarize the dissertation (Section 5.1) and then put some of its discoveries into a broader context of media multitasking and of survey methodology research. I also discuss the limitations and further implications for research (Section 5.2). Finally, I additionally outline the original contribution of the dissertation.

2 Theoretical framework

This chapter consists of three main sections, focused on multitasking, media multitasking and respondent multitasking respectively. While definitions of these three terms are presented in detail in their corresponding sections, it is useful to concisely state them at the beginning of this chapter. Within the context of this dissertation:

- *Multitasking* refers to concurrent or sequential combinations of activities.
- *Media multitasking* refers to multitasking where at least one of the activities is a media activity.
- *Respondent multitasking* refers to multitasking where at least one of the activities is related to responding to a survey. In case of web surveys, any form of respondent multitasking is also a form of media multitasking.

In other words, respondent multitasking in web surveys is a hyponym of media multitasking; media multitasking is a hyponym of multitasking.

Section 2.1 is focused on general multitasking from the perspective of the Threaded Cognition Theory. The main purpose of Section 2.1 is to present the theoretical framework that will drive the understanding of more specific forms of media multitasking and respondent multitasking and their relationship with performance.

In Section 2.2 I overview concepts from media multitasking literature that are especially valuable for my research due to relative scarce literature on the specifics of respondent multitasking. The first two parts (Sections 2.2.1 and 2.2.2) are mainly dedicated to conceptualization of media multitasking. In Section 2.2.3 I overview key reactive and non-reactive measurement approaches. In the case of the former, I am focused on self-

reports obtained through different forms of surveying. In the case of the latter, I am focused on electronic tracking. In Section 2.2.4 I present relevant empirical findings on prevalence, factors and motives of everyday media multitasking, as well as on the relationship between media multitasking and performance in goal-oriented activities.

Section 2.3 is dedicated to respondent multitasking in web surveys. It includes elaboration of the key theoretical and methodological concepts, an overview of empirical findings, and an integration with literature on multitasking and media multitasking. It follows the structure of Section 2.2 as much as possible. First, I present conceptualization of respondent multitasking (Section 2.3.1). Since survey methodology is mainly concerned with respondent multitasking because of its possible relationship with respondents' performance in survey as expressed through response quality indicators, an important part of theoretical elaboration is also this proposed relationship (Section 2.3.2). Next, I overview specifics of reactive and non-reactive measurement approaches for the particular case of respondent multitasking in web surveys in Section 2.3.3, while in Section 2.3.4 I present the findings from existing empirical studies on prevalence, factors and relationship of respondent multitasking and performance. Section 2.3.4 also discusses the current state of research on this specific kind of respondent behavior.

2.1 Multitasking: Threaded Cognition Theory

Defining multitasking is a difficult task. A long line of research on multitasking spans over several scientific fields. Not surprisingly, there are many different, often mutually incompatible definitions and conceptualizations. Several authors (e.g. Meyer and Kieras 1997; König and Waller 2010; Kenyon 2010; Salvucci and Taatgen 2011; Circella et al.

2012) have warned about the obvious negative implications of such dispersion and inconsistency.

That remains true even if we limit the discussion on the literature from cognitive sciences. Following the review by Meyer and Kieras (1997) that includes research dating back to the 1930s, we can see that studies do not only differ in terms of how multitasking is defined, but also how our cognitive system copes with multitasking. In other words, there are different perspectives on how and if our cognitive system can process multiple stimuli at once, such as: *single channel theories* (our cognitive system as a whole cannot process multiple stimuli at once), *structural bottleneck models* (parts of our cognitive system cannot process multiple stimuli at once), *unitary-resource theories* (multitasking depends on a single resource – e.g. attention – that is limited but partitionable) and *multiple-resource theories* (multitasking depends on multiple resources, each of them limited and partitionable) (Meyer and Kieras 1997).

Drawing from multiple perspectives, several authors have recently adopted cognitive architectures to model and explain multitasking. Briefly put, cognitive architectures understand human cognition as an "information processing system" and describe "the manner in which structures and functions required for human cognitive processes are organized" (Sweller 2012, 370). Numerous cognitive architectures have been developed and used in a variety of applications (Samsonovich 2010). Specific implementations of cognitive architectures in multitasking research can be found in Executive-Process/Interactive Control (EPIC) by Meyer and Kieras (1997), Multiple Resource Model by Wickens (2008), and Threaded Cognition Theory by Salvucci and Taatgen (2011).

Many researchers (e.g. Wang et al. 2012; Van Cauwenberge et al. 2014; Courage et al. 2015) have suggested the latter presents the most applicable and holistic conceptualization of multitasking and the role of our cognitive system. Moreover, Threaded Cognition Theory also attempts to

unify diverse research approaches under a single framework[1]. For these reasons, this theory is the main backbone for understanding multitasking in this dissertation.

Salvucci and Taatgen (2011) describe the study of multitasking by introducing *Unifying Theory of Multitasking*. The core part of this theory consists of three continua: *multitasking continuum*, *abstraction continuum* and *application continuum*.

The multitasking continuum is defined with regards to *task switches*. In multitasking literature, task switch refers to an act of shifting attention from one activity to another. On one end of the continuum there is concurrent multitasking where "each task progresses either simultaneously or with very short interruptions" (Salvucci and Taatgen 2011, 8). In other words, task switches occur every few second or less (e.g. talking while eating). On the other end, there is sequential multitasking where the time between task switches can be expressed in hours and "each task receives focused attention during most of its allocated execution time" (Salvucci and Taatgen 2011, 9). An example would be working on one project in the morning and a different project in the afternoon over a course of a day[2].

The abstraction continuum is based on the Newell's time scales of human action (Newell 1994): the biological band (neural and physiological processes, measured at a sub-second level), the cognitive band (specific actions and unit tasks, measured in seconds), the rational band (tasks, measured in minutes or hours) and the social band (long-term behavior,

---

[1] Several other authors have also attempted to unify multitasking research (e.g. König and Waller 2010; Circella et al. 2012).

[2] A similar distinction can be found in Circella, Mokhtarian, and Poff (2012) where authors position multitasking activities across two dimensions – share of time and share of resources – with four typical examples: monotasking, switching, interleaving, and overlaying of activities.

measured in days, weeks and beyond). Salvucci and Taatgen (2011) argue that it is impossible to account for tasks along the entire abstraction continuum. Instead, it is better to define which time scale is the most appropriate for the topic of interest. E.g. time scale of the talking-while-eating multitasking activity could be set on biological band (if we are interested in neural and physiological processes during these activities) or cognitive band (e.g. if we would be interested in exchanges of bites and uttered sentences). The research of working on multiple projects over a day could be done on rational band (e.g. if we want to know the course of activities during a single day) or social band (e.g. if we would be interested in long-term effects)[3].

Elaboration on multitasking and abstraction continua relates to another fundamental question: What are definitions of a task and an activity? In multitasking literature, there is a surprisingly little theoretical discussion on these concepts. As König and Waller (2010) note, definition depends on the specific research settings (including positioning on the multitasking and abstraction continua). It is outside the scope of this dissertation to account for a wide variety of different understandings of these concepts. Consequently, I use these two terms interchangeably throughout the dissertation (and I predominately use the term activity). I further explain activities within the context of my empirical research in Section 2.3.1.

The final continuum – application continuum – characterizes multitasking research on how much is the research "conceptually close to and relevant for some everyday task scenario" (Salvucci and Taatgen 2011, 11). In this framework, laboratory tasks are placed on the less-applied side (e.g. solving some basic linguistic tests) while real-world tasks (e.g. real-world phone use) are on the more-applied side of the application continuum.

---

[3] Circella, Mokhtarian, and Poff's (2012) concept of time dimension consists of two aspects: length of period of observation and time granularity (size of the time unit). Multitasking and abstraction continuum can be also linked to König and Waller (2010).

Authors mention that positioning on this continuum means dealing with different tradeoffs among practical applicability, the particular investigational focus of the study and research costs (Salvucci and Taatgen 2011).

The second important element of the Threaded Cognition Theory is the adoption of ACT-R (Adaptive Control of Thought—Rational) cognitive architecture (Anderson 2007). In the ACT-R architecture, different perceptual, motor and cognitive resources are connected to and coordinated by the procedural resource. This central resource gathers chunks of information from other resources and sends them new information or instructions. Communication is occurring via resource buffers, i.e. a temporary storage of information, specific for each resource.

Salvucci and Taatgen (2011) use the ACT-R architecture to explain the role and limitations of human processing resources involved in multitasking behavior. Each resource can simultaneously process a limited amount of different tasks and capacities of resource buffers are limited as well. Engagement in multiple activities that require the same resources can thus lead to conflicts where some of the activities are temporarily suspended. Such conflicts may cause delays in execution, while the recall of suspended activities requires additional cognitive processing. Moreover, longer suspension periods and/or more complex activities increase the amount of cognitive resources required when switching between activities. Additional time and resources needed to switch from one activity to another are commonly referred to as *task switch cost*.

Salvucci and Taatgen (2011) note that several factors can influence the described processes. Increased knowledge or experience of specific activities can mitigate the negative effects in multitasking situations that involve such activities. Another difference between individuals can arise from differences in cognitive system capabilities due to age, visual acuity, motor capabilities and/or other reasons.

It should be noted that while other authors approach multitasking from different theoretical perspectives and domains, there is generally a consensus that our capability to multitask is limited by the capabilities of our cognitive system. Authors also generally agree that multitasking is typically associated with inferior performance quality in goal-oriented activities. These limitations have been observed across all three continua also with numerous empirical studies (e.g. see literature reviews in Meyer and Kieras (1997) and Spink, Cole, and Waller (2008)).

However, we should not forget that multitasking is not necessarily related to worse behavioral outcomes. As already noted, Salvucci and Taatgen (2011) state that conflicts arise only if activities require the same resources at the same time. There is considerable empirical evidence in support of that claim (e.g. Kennedy 2010; Wood et al. 2012). In addition, the extent of multitasking behavior can be crucial, since several studies indicate that small amounts of multitasking in certain situations do not impair performance or can even improve it (Adler and Benbunan-Fich 2012; Ie et al. 2012; Tran et al. 2013).

To summarize, in this dissertation I follow the broad definition of multitasking which includes concurrent and sequential combinations of activities. Overall, Threaded Cognition Theory represents a solid backbone for research on respondent multitasking. Firstly, the three continua can be used to put several aspects of diverse literature on different forms of multitasking (including respondent multitasking) on a common denominator. This enables me to include relevant published theoretical and empirical findings into my research, as well as generalize findings of this dissertation into a broader multitasking framework to some extent. Secondly, adaptation of the ACT-R architecture provides a clear and flexible theory on how human cognitive system behaves in multitasking situations that can be also applied on activities associated with respondent multitasking.

2.2 Media multitasking

The most prominent strain of multitasking research in the last two decades has been focused on media multitasking; a form of multitasking that involves at least one media activity. Media activities include consumption of media content, creation of media content, and communication via electronic devices.

Numerous publications have been written on the substantial and still increasing prevalence of media multitasking in everyday life and its specific settings such as schools and work environments (e.g. Iqbal and Horvitz 2007; Jacobsen and Forste 2011; Voorveld and van der Goot 2013). Plentiful are also reports on potential negative consequences of media multitasking, such as on performance and well-being (e.g. Fried 2008; Salvucci and Taatgen 2011; Pea et al. 2012).

However, while empirical findings resonate across academia, media news, and advertising industry, there is relatively little discussion on methodological aspects of media multitasking research. Disparities in conceptualizations in general multitasking literature outlined in Section 2.1 also exist for media multitasking literature. Arguably, in this sense, issues are even worse since this research topic is fairly new and the literature is relatively more spread out across different scientific disciplines. While several authors (e.g. Jeong et al. 2005; Wallis 2006; Tokan et al. 2011) have alerted about these inconsistencies and suggested more holistic approaches, such warnings have yet to be recognized in broader research community.

Before continuing, I will present the structure of Section 2. In Section 2.2.1, I develop a conceptualization of media multitasking, based on a broad understanding of this behavior. In Section 2.2.2, I elaborate on several other important aspects that define media multitasking research, including the multitasking, abstraction and application continua. In Section 2.2.3, I overview reactive and non-reactive empirical approaches for

measuring media multitasking that are relevant for the purposes of this dissertation. In Section 2.2.4, I first present key empirical findings on prevalence, factors and motives of everyday media multitasking. Afterwards, I overview empirical evidence on the relationship between media multitasking and performance in goal-oriented activities. I also discuss factors that influence this relationship.

2.2.1 Conceptualization of media multitasking

Perhaps most commonly, media multitasking is defined as multitasking of one media activity only with another media activity (e.g. Foehr 2006; Ophir et al. 2009; Voorveld and van der Goot 2013). Several researchers define it as combination of a single media activity with one or more non-media activities (e.g. Jeong and Fishbein 2007; Meng and McDonald 2009; Tokan et al. 2011). Finally, some authors have suggested that media multitasking includes combinations of media activities with both, non-media and/or other media activities (e.g. Jeong et al. 2005; Wallis 2010; Wang and Tchernev 2012).

For the purposes of this dissertation, I follow the later stream. I particularly build upon the definition developed at the 2009 research seminar *The impacts of media multitasking in children's learning and development* held at Stanford University (Wallis 2010). In the definition agreed upon at that seminar, media multitasking encompasses the following combinations: "(a) between medium and face-to-face interaction; (b) between two or more media; and (c) within a single medium" (Wallis 2010, 8).

Most notably, I suggest that the term *device* substitutes the term medium in the above definition. As Wallis (2010, 6) notes "the computer is not a single medium, but a meta-medium in which previous media can be represented, connected, and integrated in new ways" and similar can be also said for tablets and smartphones. Definition based on complex, ever-changing medium manifestations could lead to different interpretations.

On the other hand, device is a tangible object and as such should bring more clarity.

Furthermore, for purposes of this research, I conceptualize the type (a) in a broader context so that it includes all real-life activities (e.g. household chores, eating, walking). I will refer to this type as *external media multitasking*, where the adjective external suggests that media multitasking includes at least one activity outside the media environment.

Following this logic, I refer to type (b) as *internal media multitasking* where multitasking consists of at least two different media activities carried out on at least two different devices inside the media environment. In this type of media multitasking there are no non-media activities (or alternatively, researchers are not interested in any activities outside the media environment)[4].

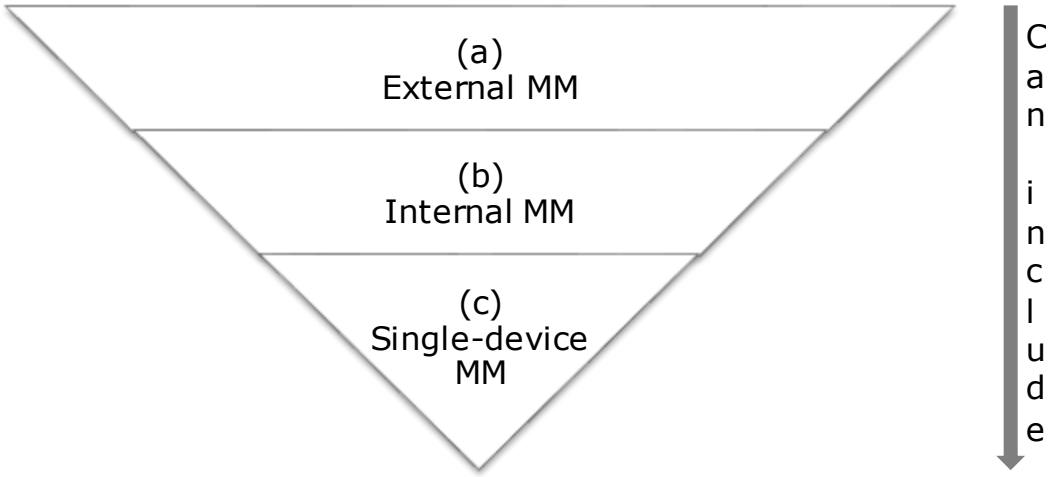
Type (c) is a *single-device media multitasking*. In this case, media multitasking consists of at least two different media activities carried out within a single device. There are no media activities on other devices and there are no non-media activities (or alternatively, researchers are not interested in them).

I present the three types of media multitasking in a top-down hierarchy in Figure 2.1. Such hierarchy suggests that internal media multitasking can also include elements of single-device media multitasking (e.g., different media activities within computer combined with watching TV). Likewise, external media multitasking can include elements of internal and single-device media multitasking.

---

[4] This type of media multitasking has also been referred to as media multitasking (e.g. Foehr 2006), simultaneous media usage (e.g. Pilotta et al. 2004), concurrent media exposure (Holmes et al. 2005), multi-screening (Google 2012), media stacking and media mashing (Ofcom 2013), media multiplexing (Chen et al. 2012), multiple media use (Tokan et al. 2011), connected viewer (Smith and Boyles 2012), etc.

Figure 2.1: Media multitasking (MM) types



While more complex and layered definitions of media multitasking type exist (Jeong et al. 2005; Tokan et al. 2011), the above three-part definition sufficiently and concisely covers all possible media multitasking scenarios in the context of this research.

To conclude, I define media multitasking as any form of multitasking that includes at least one media activity. Media activity refers to media content consummation, media content creation and communication via electronic devices. There are three general types of media multitasking, in a bottom-top order: single-device media multitasking (at least two media activities within a single device), internal media multitasking (at least two media activities on at least two devices) and external media multitasking (at least one media activity and at least one non-media activity).

2.2.2 Research aspects of empirical literature on media multitasking

In addition to differing conceptualizations, research on media multitasking also differs in several other important aspects. It should be noted that Section 2.2.2 does not want to imply that some approaches with regards to these aspects are more appropriate than others. Appropriateness of positioning on these aspects is of course inherently tied to the underlying research goals of each study.

In addition to conceptualization, other key research aspects of media multitasking literature are as follows:

(a) Related but separate aspect to conceptualization of media multitasking is selection of device(s) and/or activities included in research. Several studies have considered a wide range of devices (e.g. Voorveld and van der Goot 2013), a common choice is also to focus on screen-based devices (e.g. Google 2012), and narrower, more specific selections also exist (e.g. IAB Europe 2012). In context of single-device media multitasking, most commonly researched devices are computer and TV (Yeykelis et al. 2014).

(b) Research differs on how they define basic media activities. Some studies are interested only in what types of devices are being combined for media activities, e.g. "using a computer" is being understood as a single activity (e.g. Pilotta et al. 2004), even though a person could be engaged in different media activities on the computer. On the other side, some researchers take a granular, activity-based approach (e.g. Ophir et al. 2009). Most studies take a mixed approach between these device- and activity-based extremes (e.g. Foehr 2006).

(c) Another important difference is whether researchers are equally interested in all activities included in the research or if they make distinctions between *primary* and *secondary* activities. In this sense, perception and/or performance of the primary activity is the main topic of interest, while secondary activities are typically characterized as distractions. Generally, the research approach and subsequent analysis in such studies are adapted to the needs of the specific research goal. For instance, measurement of secondary activities is especially focused on how they are related to the primary activity. Primary activities include school tasks (e.g. Wood et al. 2012), office work tasks (e.g. Iqbal and Horvitz 2007), and perceptions of advertisements (e.g. Chinchanachokchai et al. 2015). Equal treatment of all included activities can be regularly found in general studies on media multitasking (e.g. Rideout et al. 2010).

Research approaches also vary with regards to positioning on all three continua from the Unifying Theory of Multitasking (see Section 2.1).

(d) Multitasking continuum: There are studies that focused solely on concurrent media multitasking (e.g. W. Zhang et al. 2010), while research on a single-device media multitasking is typically bounded to explore sequential combinations of activities (e.g. T. Zhang et al. 2015). Most of research on media multitasking aims to investigate across the whole multitasking continuum.

(e) Abstraction continuum: Media multitasking is rarely investigated on a level of biological or social band; studies can be typically located on a cognitive (e.g. Yeykelis et al. 2014) or rational band (e.g. Foehr 2006).

(f) Application continuum: In general, both sides of the continuum – less-applied side such as research in cognitive laboratories (e.g. Bowman et al. 2010) and more-applied such as research on real-world scenarios (e.g. Judd 2013) – are well represented. However, several authors have noted that most of the empirical work focused on effects of multitasking on performance is done via experiments with simple tasks in contrast of complex real-life situations (e.g. Foehr 2006; Junco and Cotten 2012; Wallis 2010).

Related to the application continuum, studies of course also differ in more "classical" methodological aspects such as the target population, sampling, empirical method, etc. The latter is particularly important for this dissertation and is discussed further in Section 2.2.3.

2.2.3 Reactive and non-reactive approaches for measuring media multitasking and media use

In media multitasking literature, theoretical discussion and especially empirical research on comparison of reactive and non-reactive approaches are relatively rare. Therefore, Section 2.2.3 also integrates relevant findings from literature on survey methodology and media use.

I should also note that while "reactive" and "non-reactive" are umbrella terms for many different methods, the discussion is mostly limited to two specific methods that are important for this dissertation: Self-reports and electronic tracks.

In the first part, I briefly overview usage of reactive approaches, including the corresponding applications in media multitasking research and its typical advantages and disadvantages. Similarly, an overview is then provided also for non-reactive approaches. I conclude Section 2.2.3 with a summary of studies that have empirically compared both approaches.

Reactive approaches (self-reports)

I broadly separate quantitative reactive data collection methods[5] in studies on media multitasking in two categories: *Cross-sectional survey questionnaires* (e.g. Ophir et al. 2009) and *continuous (or repeated) survey* recording based on specific diary format (e.g. Voorveld and van der Goot 2013). Some studies also combine results from different reactive methods. For instance, Kaiser Family Foundation study combines a cross-sectional survey and diaries (Foehr 2006).

Several authors have expressed doubts over accuracy and reliability of self-reporting in media multitasking (e.g. Wallis 2010) and media use research (Greenberg et al. 2005). These concerns are confirmed by findings from survey methodology research. Responding to questions about behavioral frequencies often requires a "complex interplay between memory and judgment" (Tourangeau et al. 2000, 137). Respondents use different response strategies with varying levels of accuracy, ranging from recall of specific individual episodes to estimation based on general impression (see Tourangeau et al. (2000) for an overview and discussion of response strategies).

[5] Qualitative reactive approaches such as focus groups and interviews are rarely used in the media multitasking literature, although some exceptions exist (e.g. Iqbal and Horvitz 2007; Agosto and Abbas 2010; Bardhi et al. 2010).

Response accuracy can also be affected by question design. For example, shorter reference periods (e.g. asking about behavior for the past week instead of for the past six weeks) increase accuracy of responses; and careful question wording (e.g. asking "how many times" instead of "how often") can also have beneficial effects (Blair and Burton 1987; Burton and Blair 1991). Studies on whether to ask about a typical or the most recent time period (e.g. Chang and Krosnick (2003) compare "typical week" and "past week" questions), comparisons of open- and closed-ended question formats (Tourangeau et al. 2000), and effects of decomposition of questions into several subcategories or providing other forms of contextual cues (cf. Stone et al. 1999; Belli et al. 2000; Tourangeau et al. 2000) also provide interesting, although rather inconclusive findings.

Overall, researchers agree that accuracy of responses to behavioral frequency questions can be a considerable issue even for carefully conducted surveys. This is especially true for frequent, unregimented and/or mundane behavior patterns (Blair and Burton 1987; Tourangeau et al. 2000; Chang and Krosnick 2003). Arguably, majority of everyday media behavior, including media multitasking, falls into this category. In addition, other factors that are not specific to questions about behavioral frequencies, such as social desirability bias and privacy issues should also be considered.

While most of these concerns also hold true for diaries, this method generally gives more valid measures of behavioral frequencies; in particular, because shorter reference periods encourage more accurate response strategies (see Bolger et al. (2003) for an overview of diary methods). However, diaries have some specific limitations compared to surveys that might negatively influence participations rates and accuracy. Specifically, participants can feel a heavier burden because of recurrent queries about their behavior. In addition, some research indicates that participants might change their behavioral patterns because keeping a diary influences their perceptions, awareness, and knowledge about

particular behavior (e.g. Bolger et al. 2003; Greenberg et al. 2005; Möller et al. 2013).

To conclude, researchers face many potential problems when using reactive approaches for investigating media multitasking. This is especially true for studies that want to account for a wide array of media multitasking behaviors. These complex and multifaceted behavioral patterns require careful planning from researchers and thorough response from participants. Accuracy can also be particularly questionable in studies interested in task switches. For example, research shows that people switch between activities on PCs very rapidly (e.g. Kraushaar and Novak 2010; Judd 2013; T. Zhang et al. 2015). Even if task-switching frequency in the researched behavior is lower, it is unlikely that respondents would accurately remember all of the numerous task switches between different, often relatively mundane media activities.

Non-reactive approaches (electronic tracks)

In early 1940s, Nielsen published a paper that introduced the Audimeter device (originally developed at the Massachusetts Institute of Technology in 1936). Installed into a radio receiver, the device made records of when the radio was turned on and off, as well to which radio station was tuned in "at every instant of the day or night" (Nielsen 1942, 219). In the following decades, Nielsen and many other researchers have developed this approach further and expanded it to other media devices (e.g. see Hill (2014) for an overview of studies on TV viewing behavior).

Emergence of modern telecommunication services (e.g. mobile phone networks and the Internet) has opened many new possibilities. Researches can now make use of records on traffic generated in telecommunication networks (Weiss 2005) and/or servers (Srivastava et al. 2000). Within specific context of web usage, such data collection is referred to as a *server level data collection*. *Client level data collection* requires additional software implementation on user's device, but in return

can offer enhanced information about user behavior that cannot be captured by server level collection (Srivastava et al. 2000).

Typical end-product of these methods is a data log where every interaction between the user and media device/service is marked with a timestamp. Hence, non-reactive data is generally not influenced by different respondent factors that can contribute to measurement error in reactive approaches such as social desirability bias. For these reasons, different authors encourage the use of such methods in different contexts of media use and media multitasking research (e.g. Greenberg et al. 2005; Wallis 2010; Möller et al. 2013).

However, non-reactive approaches also have some important limitations. Excluding server level collection methods, these methods require software and/or hardware modifications of participants' devices. In such cases, some people could be especially inclined to decline their participation due to privacy or other concerns (Milavsky 1992). Technical specifics of different models and generation of devices might also lead to exclusion of certain subpopulations or segments of devices (Blake and Klimmt 2012). Consequently, it is hard to guarantee large and representative samples. However, we should note that in recent decades reactive methods are facing this issue as well, albeit for different reasons (I discuss non-response behavior in Section 2.3.2).

Increasingly high penetration and usage of different types of media devices pose another challenge for nonreactive methods. Holistic investigation of behavior in modern media environment must account for a wide array of media devices and activities. While non-reactive methodological solutions for single devices and/or services exist, "cross-platform measurement is currently in its infancy" (Hill 2014, 84).

This limitation is especially relevant for media multitasking research on concurrent and/or sequential combinations of media activities on different devices that occur in short time spans. Moreover, measuring non-media

activities with non-reactive methods is another substantial challenge. It is therefore not surprising, that we have yet to see a study that would employ a non-reactive approach for measuring external or internal media multitasking.

Existing non-reactive studies are thus limited to single-device media multitasking. Data in these studies is typically collected on the client side, e.g. by installing special tracking software to measure all activities on participants' PC (e.g. Kraushaar and Novak 2010; Yeykelis et al. 2014; T. Zhang et al. 2015). Studies that are interested in online activities also exist and include servers-side data collection such as proxy server logs (Grace-Martin and Gay 2001) and query logs from web search engine (Spink et al. 2006), as well as client-side level collection (Lottridge et al. 2012).

Typically, these studies are conducted on students (e.g. Mark et al. 2014; Yeykelis et al. 2014) or office workers (e.g. Iqbal and Horvitz 2007; Mark et al. 2012). Observation periods are rarely longer than two weeks and sample size rarely reaches a three-digit number. However, exceptions to some of these constraints exist (e.g. Judd 2013; T. Zhang et al. 2015). I am not aware of any study that would claim to have a representative sample.

More than 70 years since introduction to media use research, non-reactive data collection still has important methodological challenges to resolve. Despite that, these methods increasingly attract researchers from different fields due to several advantages over reactive approaches. In the context media multitasking research, non-reactive approach might be preferable if researchers are interested in multitasking on a single media device and/or online services.

Overview of comparative studies on reactive and non-reactive approaches

Another useful property of non-reactive methods is that it can be used to triangulate and/or validate results from self-reports. Several authors have done this in studies on media use of different technologies, including television (e.g. Ferguson 1994; Kaye and Sapolsky 1997), the Internet (e.g. Lagerstedt et al. 2012; Lottridge et al. 2012) and internal ICT systems at a workplace (e.g. Collopy 1996; Deane et al. 1998). Most commonly, such validation is done in research on smartphone usage (e.g. Kobayashi and Boase 2012; Abeele et al. 2013; Rivron et al. 2015).

These studies often show conflicting results and a thorough overview would exceed limitations of this paper. Arguably the most common finding is that generally people tend to over-report the time spent on media activities (e.g. Deane et al. 1998; Kobayashi and Boase 2012; Berolo et al. 2015). On the other hand, in several studies participants under-reported frequency of specific media activities (e.g. Lagerstedt et al. 2012; Bouwman et al. 2013; Möller et al. 2013).

The literature discusses different factors that contribute to the direction and amount of misreporting. Besides the already elaborated factors, most commonly investigated is the actual intensity of use. Several authors provide evidences of so-called regression to the mean, i.e. "light users" tend to over-report their usage, while "heavy users" under-report it (e.g. Collopy 1996; Timotijevic et al. 2009; Abeele et al. 2013). Bouwman et al. (2013) report that heavy users are more accurate in their assessment than light users, while Berolo et al. (2015) report the opposite. Research on socio-demographic factors is also inconclusive; however, there are some empirical evidences that factors such as age, gender, income, and employment are significantly associated with accuracy of self-reports on media behavior (e.g. Abeele et al. 2013; Bouwman et al. 2013; de Reuver and Bouwman 2015).

Despite the above mentioned issues, different studies show that self-reports are usually significantly correlated with non-reactive measurements (e.g. Deane et al. 1998; Abeele et al. 2013; de Reuver and Bouwman 2015). However, while self-reporting might be accurate enough for observing general trends of media use, several authors warn against relying solely on self-reports as indicators of media use in explanatory models can lead to Type-1 and Type-2 errors (e.g. Collopy 1996; Kobayashi and Boase 2012; de Reuver and Bouwman 2015).

In the remainder of Section 2.2.3, I overview comparative studies on media multitasking behavior.

Ferguson (1994) and Kaye and Sapolsky (1997) have investigated channel changing frequencies on a cable television. Despite different research approaches, both studies report that participants substantially under-reported how many times they switched between media contents on their television. Ferguson (1994) reports that half of participants underestimated the exact count by more than 3 times compared to the data from electronic counter. Kaye and Sapolsky (1997) report that on average, participants underestimated the exact count by a factor of 10.

Iqbal and Horvitz (2007) monitored use of computers at a workplace and were especially interested in reactions to alerts generated by e-mail and instant messaging clients. Authors report that participants "appear to be largely unaware of the amount of time they end up spending on the alerting application, on other tasks they invoke as a result of responding to the alert, and on browsing through other peripheral applications before resuming the suspended task" (Iqbal and Horvitz 2007, 684).

Brasel and Gips (2011) conducted an observational laboratory study to investigate how people combine television and computer use. Authors report that participants switched between the media devices four times per minute. However, in a post-hoc survey participants recalled significantly smaller amount of switches, only 15% of the actual amount.

While we are not aware of any research on smartphone use that would investigate multitasking per se, we can find some relevant findings from studies that investigated use of different smartphone apps. Möller et al. (2013) have compared non-reactive data to three different diary methods with varying reporting intervals. Depending on the reporting interval, self-reports revealed 40% to 70% of actual app usages. In Rivron et al. (2015) study only 43% of participants correctly estimated the number of apps they use daily, while all other participants have underestimated it.

2.2.4 Key empirical findings on everyday media multitasking

This overview of key empirical findings on media multitasking consists of three parts: prevalence, factors, and relationship with performance. Empirical findings on prevalence, factors and motives mainly originate from applied studies. On the other hand, evidences on performance mostly come from less-applied studies, such as cognitive experiments. While it would be valuable to describe in details each included study in terms of other conceptual and empirical aspects presented in Sections 2.2.1, 2.2.2 and 2.2.3, such thorough review is outside the limitations of this dissertation. However, I will attempt to cover the most important information related to these aspects throughout the overview.

Prevalence

I present findings on prevalence separately for (a) external, (b) internal and (c) single-device media multitasking in everyday life. Where possible, I gravitate towards large-scale representative studies in terms of target population and sample; as well as on studies that provide time trends. Moreover, I focus especially on media multitasking combinations that involve the use of computer and/or the Internet since this directly relates to research on respondent multitasking in web surveys.

(a) Large scale studies on prevalence of **external media multitasking** (i.e. multitasking that involves media and non-media activities) are scarce

and mostly limited to research on younger populations. In a 2004 study by Kaiser Family Foundation, American youngsters reported that they spend 5.8 hours per week for interacting with a computer (as their main activity). While being on a computer, the most common non-media activities were: doing homework not on the computer (7% of total computer time), eating (6%) and chores (3%) (Roberts et al. 2005).

Jeong and Fishbein (2007) have also investigated prevalence of external media multitasking in American youth. Most common relevant combinations were the Internet-homework (24% of respondents often engage in such combinations), the Internet-interaction with friends (22%) and Internet-eating (17%).

(b) Prevalence of **internal media multitasking** (i.e. multitasking that involves only media activities) has been investigated more thoroughly and warrants a more detailed overview. Here, I will focus on diary studies done by Kaiser Family Foundation (fielded in 1999, 2004 and 2010 for 8-18 year old population of USA), Ofcom (fielded in 2010, 2014 and 2016 for 16+ year old population of UK) and Voorveld and van der Goot (fielded in 2010 for 13-65 year old population of the Netherlands).

It is important to note that the studies do not use the same measurement instruments and that their results are not comparable in this regard. Otherwise, all studies are relatively similar in terms of research aspects presented Section 2.2.2. They attempt to cover both sides of the multitasking continuum and are focused on the rational band (i.e. they measure and analyze activities in minutes and hours). Moreover, they account for a wide range of devices and activities, take a more granular approach in defining the basic measured media activities, and are generally not concerned with distinction between primary and secondary activities.

All studies report daily averages for total media use (i.e. how much time participants spend on media activities) and the time spent on internal

media multitasking. Kaiser Family Foundation and Ofcom series of studies also provide valuable information on overall trends throughout the years and information on total media exposure (i.e. on how much of media contents a person is exposed to per day). The latter is larger than total media use due to internal media multitasking.

The most recent report by Kaiser Family Foundation estimates that an average US youngster uses media almost a third of the day (7 hours and 38 minutes) and is exposed to 10 hours and 45 minutes of media contents per day. In total, an average participant is engaged in internal media multitasking 29% of total media use time (i.e. 3 hours and 7 minutes per day) (Rideout et al. 2010). In the 2010 study by Ofcom (2010), the same share (29%) is reported for the youngest age group (16-24 year olds); however, UK participants in general have about an hour less of total media use and total media exposure compared to their US counterparts. Similar results are reported for the youngest age group (13-16 year olds) in Voorveld and van der Goot's study (2013): Average total media use time for these participants is also about 6 hours and 30 minutes per day; 31% of this time is spent on internal media multitasking. On the other hand, figures for 17-19 year olds (4 hours of total media use time; 24% share of internal media multitasking) and 20-24 year olds (5 hours; 25%) are considerably lower than in the UK Ofcom study.

Table 2.1 summarizes results of all studies by Kaiser Family Foundation and Ofcom. In the US studies, total media use has been rising throughout the years. More importantly, the share of media use time spent on internal media multitasking has almost doubled from 1999 to 2009; correspondingly, daily averages for total media exposure have increased for more than three hours in a decade. Similar trend can be observed for

the UK studies from 2010 to 2014; however, it stagnates in the latest report from 2016[6].

Table 2.1: Average daily media behavior for young populations in USA and UK

	USA, 8-18 year olds			UK, 16-24 year olds		
	1999	**2004**	**2009**	**2010**	**2014**	**2016**
Total media exposure (hh:mm)	7:29	8:33	10:45	9:32	14:07	13:11
Total media use (hh:mm)	6:19	6:21	7:38	6:35	9:08	8:56
% of total media use time spent on internal media multitasking	16	26	29	29	35	33

Sources: Kaiser Family Foundation studies for the US data **(Roberts et al. 2005; Foehr 2006; Rideout et al. 2010)**, Ofcom studies for UK data **(Ofcom 2010, 2015, 2016)**

Ofcom studies and Voorveld and van der Goot (2013) also investigate internal media multitasking for other age groups. As observable from Table 2.2, estimated daily averages for general population in 2010 are about 7 hours of total media use with 20% of this time spent on multiple media activities. Again, we can observe a substantial increase in the UK data between the years 2010 and 2014 and relative stagnation between 2014 and 2016.

Relevant results into prevalence of internal media multitasking can be also found in the report by IAB Europe (2012). This study is based on a cross-sectional survey and is focused only on the combinations of activities worded as "watching TV" and "using the internet" (IAB Europe 2012, 7). Results from surveys fielded in 2010 and 2012 reveal that the share of respondents who are regularly engaged in such behavior has risen from 32% to 53%. The 2012 study included 19,000 respondents from 30 different European countries and there are notable differences between

[6] As authors note themselves, this stagnation could influenced by the slightly changed methodology, see Ofcom (2016) for more details.

countries – from 71% (France) to 31% (Romania). Slovenia is ranked sixth with 60%.

Table 2.2: Average daily media exposure and media use for general populations in the Netherlands and UK

	The Netherlands, 13-65 year olds	UK, 16+ year olds		
	2010	2010	2014	2016
Total media exposure (hh:mm)	N/A	8:48	11:06	10:52
Total media use (hh:mm)	6:46	7:05	8:41	8:45
% of total media use time spent on internal media multitasking	22	20	21	20

Sources: Voorveld and van der Goot **(2013)** for the Dutch data, Ofcom studies for UK data **(Ofcom 2010, 2015, 2016)**

In addition to the presented prevalence rates, another key insight from literature on internal media multitasking is the importance of PC and the Internet in such behavior. In the 2014 iteration of their study, Ofcom reports that 54% of time spent on laptops involves internal media multitasking (higher shares are noted only for smartphones and landline phones), while this is true also for 40% of time spent on desktop computers. Furthermore, Ofcom also reports that Internet-related activities appear in five out top six most common multitasking combinations (Ofcom 2015). In the Voorveld and van der Goot (2013) study, Internet-related activities appear in the top three most common multitasking combination for each age groups. In both, Ofcom and Voorveld and van der Goot studies, we can also observe that computer/Internet activities have an especially central role in multitasking for younger age groups. Authors of the Kaiser Family Foundation studies have also observed that computer has a well-established "role in fostering" media multitasking among their young participants (Foehr 2006, 13). They performed computer activities in solitude only about a third of the time (Foehr 2006).

(c) As noted in Section 2.2.2, studies on **single-device media multitasking** (i.e. multitasking that involves only media activities on a single device) on PC are typically interested in sequential multitasking. More specifically, primary focus of these studies is frequency of task switches and duration spent on one activity before switching to another. In this context, activity is typically defined by the (content of the) active window or a tab on a computer.

Judd and Kennedy analyzed computer logs from an open-access computer laboratory at an Australian university from 2007 and 2009 respectively (Judd and Kennedy 2011; Judd 2013). The 2009 data contained over 6,000 logs from 536 different students. Based on this data, authors report that the mean time between task switches is 2 minutes, with 77% of all durations less than 5 minutes. Over 70% of computer sessions from 2009 contained some level of multitasking, a 20% increase from 2 years before.

Yeykelis, Cummings, and Reeves (2014) conducted an interesting research on twelve students from an US university. They analyzed logs from participants' PC. Mean time between task switches was about 1 minute, with 75% of all durations shorter than 1 minute. Only 8% of activities were longer than 4 minutes and majority of them were video related (e.g. watching YouTube or Netflix).

Zhang et al. (2015) analyzed a month's worth of data from PC logs by 3,000 volunteers from China with diverse socio-demographic characteristics. Mean time between task switches for 10-30 year olds is about 2.5 minutes and becomes longer for older age groups (it is about 5 minutes for participants aged over 50 years). Authors also develop a more detailed measure of Average Multitasking Switching Rate. This measure refers to task switching sequences in form of A -> B -> A; i.e. if a person switches from one software process (e.g. text editor) to another process (e.g. web browser) and back to the original process (text editor again). Such sequences are accounted for only if they occur within a 5-minute

span. Here, differences between age groups are not so notable, but are still present: participants aged between 10 and 20 years averaged 2.8 of such combinations per 5 minutes, while the figure drops to about 2.3 for participants aged over 50 years.

Factors

I present factors by following a (slightly adapted) model by Jeong and Fishbein (2007) which includes: (a) socio-demographic factors, (b) individual factors, and (c) structural/technological factors. Following elaboration by Kononova and Chiang (2015), I also include (d) motives for media multitasking behavior. Majority of research summarized here is based on cross-sectional surveys that takes a similar wide (in terms of included devices/activities) and granular (in terms of how basic activities are defines) approach as the presented studies of prevalence of everyday media multitasking.

(a) Socio-demographic factors. Age is such an important factor for media multitasking research that I already had to elaborate on it during the overview of literature on prevalence. To summarize, age is a notable factor in two ways. Younger people are more likely to multitask with media in general. Therefore, many studies are focused solely on college or high school students. Moreover, different age groups are usually engaged in different patterns and combinations of media multitasking. Here, younger generations again stand out due to increased prevalence of activities related to the use of PC and/or the Internet. (e.g. Voorveld and van der Goot 2013; Hwang et al. 2014; Ofcom 2015; cf. Carrier et al. 2009)

With regards to gender, there are several empirical evidences that females are more likely to be engaged in media multitasking (e.g. Foehr 2006; Hwang et al. 2014; Gil de Zúñiga et al. 2015). Some researchers go in greater details and report that females are especially more involved in some specific forms of media multitasking, such as combinations involving

audio-based activities (Jeong and Fishbein 2007) or TV activities (Hwang et al. 2014); however, research on this is relatively scarce and rather inconclusive (cf. Cotten et al. 2014; Hwang et al. 2014; Gil de Zúñiga et al. 2015).

Some limited research has also been done for other demographic factors, such as ethnicity, education and income (e.g. Foehr 2006; Jeong and Fishbein 2007; Kononova and Alhabash 2012). However, wider and deeper research on demographic factors is warranted in order to have a solid understanding on their roles. E.g. Kononova (2013) alerts that relationship between gender and media multitasking varies between countries, while Cotten, Shank, and Anderson (2014) argue that differences between genders can be at least partly attributed to other factors such as technology use and ownership. Similarly could be claimed for other socio-demographic factors as well.

(b) Individual factors. *Ownership, access and use of technology* have been confirmed as important factors of media multitasking also in several other empirical studies (e.g. Foehr 2006; Jeong and Fishbein 2007; Kononova and Chiang 2015). In addition to PC, other modern media and communication devices such as smartphones and tablets are particularly important drivers of internal media multitasking, as are combinations of different screen-based devices (e.g. Smith and Boyles 2012; Google 2012; Nielsen 2012).

Another group of individual factors refers to *personality traits*. Sensation seeking refers to individual's preference for strong stimulation experiences. Consequently, people with a strong preference for sensation seeking tend to seek "varied, novel and complex sensations" (Hwang et al. 2014, 543) and could thus engage more frequently in media multitasking. Positive relationship between media multitasking and sensation seeking has been confirmed in several empirical studies (e.g. Jeong and Fishbein 2007; Sanbonmatsu et al. 2013; Hwang et al. 2014).

Polychronicity relates to individual's preference to engage in multitasking behavior (König and Waller 2010). The broader concept originates from anthropology where it has been used to categorize cultures into monochronic (i.e. where people like to one thing at a time) and polychronic (i.e. where people like to do multiple things at a time). A typical example of the later is Northern Europe, while a typical example of the former is Latin America. However, polychronicity has also been explored on an individual level. Several empirical studies have confirmed that polychronicity is an important driver of increased media multitasking behavior on an individual and country level (e.g. König et al. 2010; Voorveld et al. 2014; Kononova and Chiang 2015).

The Big Five personality traits[7] have been less often studied in media multitasking research. Existing scarce empirical findings on whether these traits relate to increased media multitasking behavior are inconclusive. E.g. Wang and Tchernev (2012) report positive relationship only with neuroticism; Loh and Kanai (2014) only with extraversion; while Ophir et al. (2009) did not find any relationship between the Big Five personality traits and media multitasking. These findings are rather surprising since sensation seeking is positively correlated with the Big Five traits agreeableness and openness, as it has been confirmed by several studies included in the literature review by Roberti (2004).

(c) Structural/technological factors. As Foehr (2006) notes, overall technological development and penetration of modern media devices and services are also important drivers of media multitasking. In contemporary societies, individuals are embedded into complex media environment (Petrič et al. 2011). It is safe to assume that such

[7] The Big Five personality traits is a widely adapted model used to describe human personality with regards to five dimensions: extraversion, agreeableness, conscientiousness, neuroticism, and openness to experience (Goldberg 1990).

environment not only enables but also stimulates media multitasking[8]. While technological factors are often debated on a theoretical level, they are rarely a subject of empirical research. Nevertheless, two interesting studies provide some insight into these factors.

Cross-sectional survey conducted by Kononova (2013) in USA, Russia, and Taiwan shows that country-level ICT development is an important factor of media multitasking even after accounting for socio-demographic, individual (including technology ownership and use) and other structural factors. More cross-cultural studies would be a welcome addition to the field.

An interesting experiment by Lottridge et al. (2012) has highlighted that differences in graphical user interface design can also importantly influence extent of media multitasking. Authors electronically logged web browser usage on PC of 14 participants for a week. In the second week of the experiment, participants used the same browser, but with a modified interface. Visited web sites were automatically categorized into two groups: "work" and "non-work" related contents (Lottridge et al. 2012, 1957). The modified interface visually expressed this categorization in two ways. Firstly, work-related tabs at the top of the browser had a more prominent look and placement. Secondly, a timer at the bottom-right corner of the browser showed how much time has a user spent on a specific tab, as well as the overall share of time spent on work and non-work web sites. Electronic tracks revealed significant differences between the first and the second week of usage. Average numbers of visited web

---

[8] A simple example of this can be seen in technological development and advertising strategies for smartphones and tablets. Compared to the first generations of these devices, newer models are increasingly capable of handling multiple applications at once. Manufacturers of course advertise these increased capabilities, sometimes even in a very competitive and aggressive tone such as the Samsung campaign against Apple (Welch 2014).

sites and non-work web sites, opened tabs, browser sessions and especially tab switches were all considerably lower in the second week.

(d) *Motives*. I structure this brief overview of motivations for media multitasking after Kononova and Chiang (2015). Authors introduce five main motivations and include an empirical study that confirms all five as significant predictors of media multitasking behavior:

- Control. This motivation refers to "being in charge of media use" (Kononova and Chiang 2015, 34). In other words, people are motivated to engage in media multitasking because they believe they can control and handle different media activities for various purposes.

- Efficiency. People commonly report that they engage in media multitasking because it is an effective and productive approach for performing activities. This motivation can be amplified in certain situations such as work overload and transportation (see also Foehr 2006; Baron 2008; Tokan et al. 2011).

- Entertainment. In some cases, the act of media multitasking brings feelings of enjoyment. In related findings, Baron (2008) notes that media multitasking is often used to fight boredom.

- Connection. People engage in media multitasking because it helps them to connect better with other people. These social connections in multitasking combinations can occur in physical (Meng and McDonald 2009) or digital world (Kononova and Chiang 2015).

- Addiction. Kononova and Chiang (2015, 34) define it as: "habitual, routine media multitasking, which goes beyond user's control". This has been exposed as an important reason also by Baron (2008) and Bardhi (2010).

Relationship with performance

Plentiful studies have investigated performance in the context of media multitasking. As different authors note (e.g. Fried 2008; Ophir et al. 2009; Adler and Benbunan-Fich 2012), it is challenging to provide solid empirical evidences on the presumed causality between media multitasking and performance, especially in the real-world setting as defined by the application continuum. Therefore, researchers usually cautiously write about correlations and relationships in this stream of research.

For the purposes of this discussion, I categorize relevant literature into two groups. The first group is focused on direct relationship of media multitasking with performance. In other words, such research typically investigates how secondary activities or disturbances influence performance of the primary task in a specific setting. The second group is concerned with indirect relationship. Such research is aiming to assess whether heavy engagement in everyday media multitasking is related to lower general cognitive abilities that are associated with performance in cognitive activities.

The remaining discussion is organized as follows. First, I present key findings on (a) direct relationship. In line with purposes of this dissertation, this part is mainly based on the overview of human-computer interaction literature in Salvucci and Taatgen (2011). Second, I summarize literature on (b) indirect relationship. This part is focused on literature based on the Media Multitasking Index. In the end, I provide a (c) brief conclusion that includes relevant findings from two other prominent research streams.

(a) *Direct relationship*. Following upon their Threaded Cognition Theory, Salvucci and Taatgen (2011) hypothesis include studies over the whole application continuum. Overall, the literature shows that secondary activities in human-computer interaction have negative consequences for

the main task performance, as measured with time and other assessment criteria.

This is especially true in the case of sequential media multitasking combinations involving secondary activities that are not related to primary activity and require a separate construction of the problem state (i.e. temporary information needed execution of the an activity) and/or need to retrieve information from declarative memory (i.e. the main resource for storage of information; also known as explicit memory) (Salvucci and Taatgen 2011).

Experiments have shown that only one problem state can be active at a time. Therefore, in case of cognitively demanding secondary activities, the problem state of the primary activity is suspended. The Cognitive system tries to maintain the primary activity's problem state by rehearsing it in the declarative memory. If the necessary information is maintained when a person switches back to the primary activity, its corresponding problem state can be retrieved from the declarative memory. However, the possibility that this information is not maintained increases due to the interplay of several factors: length of the suspension period, complexity of the primary activity, and whether the secondary activity requires the limited resources of the declarative memory as well. When this occurs, the problem state needs to be reconstructed and this increases the task switch cost (Salvucci and Taatgen 2011).

Another important aspect of sequential multitasking is the timing of the task switch with reference to mental workload required by the primary task. In other words, performance can be especially negatively affected if the primary task is interrupted at the moment when it is utilizing a high amount of mental resources. Experiments show that task switching in the moments of lower mental workload is much less critical. Related, but separate are also empirical findings that performance can be particularly

hindered in the case of unexpected interruptions (Salvucci and Taatgen 2011).

Several other factors that can determine a direct relationship between media multitasking and performance have already been mentioned in different parts of this dissertation. To briefly summarize, (concurrent or sequential) combinations of the primary and secondary activities can be particularly critical if they require the same perpetual, motor or cognition resources (Salvucci and Taatgen 2011). There are also numerous individual factors (e.g. general cognitive abilities, prior knowledge of activities, factors specific to the task) that can influence the relationship between media multitasking and performance (Salvucci and Taatgen 2011). Research shows that situational factors such as tiredness and mindfulness can also have an important role in this relationship (e.g. Ie et al. 2012; McCarthy 2013).

(b) *Indirect relationship*. Media Multitasking Index (MMI) has a prominent role in research on media multitasking relationship with performance. It was introduced by Ophir et al. (2009) and has been used in numerous studies since. MMI is based on self-reported data collected via the Media Multitasking Questionnaire – the higher the score, the more have participants reported on being engaged in everyday life media multitasking. It can be used to identify heavy media multitaskers (HMM) and light media multitaskers (LMM) who score one standard deviation above or under the mean sample MMI respectively.

In the study by Ophir et al. (2009), participants solved several cognitive tests related to attention and distraction filtering. Results have shown that HMM differ from LMM in approaches to fundamental information-processing activities – they have more difficulties in "filtering out irrelevant stimuli from their environment", "are less likely to ignore irrelevant representations in memory" and "are less effective in suppressing the activation of irrelevant task sets" (Ophir et al. 2009,

15585). The authors conclude that the data suggests that LMM more easily focus on a single task and filter our distractive stimuli compared to the HMM greater tendency for exploratory information processing.

Several subsequent studies have generally confirmed these findings with different sets of cognitive experiments related to attention and focus in single goal-oriented activities (Cain and Mitroff 2011; Lui and Wong 2012; Ralph et al. 2015; Moisala et al. 2016; Uncapher et al. 2016; cf. Minear et al. 2013). In contrast, research by Ralph et al. (2014) has taken MMI outside the laboratory setting[9] by conducting a survey among university students that included the Media Multitasking Questionnaire along with standardized instruments for self-reported measures on different aspects of everyday attention. Results show that higher MMI scores are related to higher self-reported frequencies of attentional failures, as well as spontaneous and deliberate mind wandering.

On the other hand, several studies have investigated the relationship of MMI with other types of cognitive tests and provide more optimistic news for HMM. E.g. HMM performed better in multisensory tests (Lui and Wong 2012). With regards to task switching, Alzahabi and Becker (2013) report that HMM exhibit a better performance than LMM, while Minear (2013) found no effects for either group. Similar conclusion echoes in the Ralph et al. (2014) study, where MMI was not correlated with self-reported measures of difficulties in attention switching. Moreover, Ophir et al. (2009) note that HMM are more prone to the exploratory processing of information, which can be beneficial in certain settings (see also experiment by Lottridge et al. (2015)).

Overall, the literature on MMI gives an important insight into the potential advantages and disadvantages regarding the performance of people who

[9] MMI has been used also in other studies outside laboratory settings (e.g. Pea et al. 2012; Kononova and Chiang 2015; Nooner and Schaefer 2015); however, none of them investigated the relationship with performance.

are heavily engaged in everyday media multitasking. It is important to note that researchers do not claim that intensive everyday media multitasking behavior causes positive or negative consequences for individual's cognitive capabilities such as attention. As Ophir et al. (2009) state themselves, it is quite possible that the opposite true. Another important thing to note is that majority of discussed studies is positioned heavily on the less-applied side of the application continuum. More research investigating the relationship between MMI and real-life performance is warranted.

(c) Another relevant and prominent stream of literature focuses on applied research on the relationship between media multitasking and academic performance. Researchers in this stream use varied conceptualizations of media multitasking and research approaches. Data collection methods include cross-sectional surveys (e.g. Fried 2008; Junco and Cotten 2012), diaries (e.g. Jacobsen and Forste 2011; Mokhtari et al. 2015), and electronic tracks (e.g. Hembrooke and Gay 2003; Kraushaar and Novak 2010). Nevertheless, these studies provide relatively consisting findings. Increased engagement in secondary media activities during schoolwork (whether in-class or at home) is predominately correlated with worse academic performance.

Lastly, I would like to underscore again that none of the authors summarized here claim that all media multitasking behavior is related to worse performance or other aspects of life. Accordingly, different authors from organizational and occupational health psychology recommend taking controlled breaks during cognitively demanding activities. Controlled here refers to length, timing, and activities performed during such break. Experiments have shown that such task switches rarely hinder performance or can even improve it. More importantly, they have beneficial effects on well-being and other health-related aspects (e.g. Henning et al. 1997; Zacher et al. 2014; Mijović et al. 2015).

2.3 Respondent multitasking in web surveys

In this Section, I synthesize the findings from the literature on respondent multitasking (RM) in surveys and particularly in web surveys. Moreover, this Section serves as a link between the more general (media) multitasking literature, specifics of respondents multitasking and my empirical research. Unless explicitly stated otherwise, this Section deals with web surveys and not with the other survey modes.

Findings on the prevalence of media multitasking and its relationship with performance presented in Section 2.2.4 should be quite remarkable for survey methodologists. Overall, people spend increasingly more time on media multitasking. This is especially true for younger generations and for the use of modern media devices such as PC, tablets and smartphones (i.e. devices that are used for responding to web surveys).

Moreover, one should keep in mind that most of the presented research on the prevalence of everyday media multitasking is based on self-reports which tend to undervalue the amount of media use and multitasking (Section 2.2.3). With this in mind, reports on the frequency of task switching from research on internal media multitasking on PC are particularly confounding and relevant for web surveys.

Such conclusions can be worrying since media multitasking is typically related to worse performance, in particular for those activities where keeping attention on a single task is crucial (Section 2.1). More often than not, the key research questions in surveys assume and require a considerable amount of respondents' attention.

Considering all this, it is not hard to presume that web surveys are just another stream in respondents' personal modern media environment. In this environment, web surveys fight for respondent's attention with other

activities, even (or especially) after the participant has started responding to the web questionnaire. Media multitasking literature indicates that the results of this battle can have consequences for the quality of respondent's performance.

In conclusion of this introduction, I restate the structure in the light of Sections 2.1 and 2.2: In Section 2.3.1, I provide a definition and conceptualization of respondent multitasking with regards to the research aspects of media multitasking. These aspects are also used to expand on a taxonomy of secondary activities, based on Zwarun and Hall (2014). In Section 2.3.2, I elaborate on the relationship between respondent multitasking and survey performance. This includes relevant theoretical insights from survey methodology and Threaded Cognition Theory. Sections 2.3.3 and 2.3.4 are focused on an overview of the literature on RM. In the former Section, I discuss the published studies in the context of measurement approaches for respondent multitasking. In the latter Section, I provide an overview of key empirical findings and discuss them from relevant multitasking and media multitasking perspectives that also help to highlight the most considerable research gaps in the existing literature on respondent multitasking.

2.3.1 Conceptualization of respondent multitasking and taxonomy of secondary activities

After establishing multitasking and media multitasking, defining respondent multitasking (RM) in web surveys is relatively straightforward. Respondent multitasking in web surveys is any type of multitasking where one of the activities is responding to a web survey.

However, it is useful to position RM within the research context presented in Section 2.2.2:

(a) Generally, a well-rounded research on RM should account for a wide range of activities and devices. At the current state of research, not much

is known about which activities are commonly combined with responding to a survey and/or have a notable relationship with the response quality.

(b) Similarly, RM research should take a granular approach for defining basic media activities. A device-based activity definition is especially inadequate for research on RM in web surveys due to a large volume of single-device media multitasking on modern media devices.

(c) In the context of RM research, responding to a survey is considered as the primary activity, while all other activities are regarded as secondary activities. To clarify, the adjective "primary" here does not mean to imply that this is the main activity for the respondent in terms of dedicated cognitive resources, time or perceived importance. Instead, "primary" implies that this activity is the main research interest. In other words, it is quite possible to imagine a situation where a respondent is more invested in other activities during the observation period (i.e. the time between the start and end of responding to the questionnaire) than in the survey itself. While responding to a survey might be "secondary" from the respondent's perspective, it is still "primary" from the researcher's perspective.

(d) Again, at the current state of research, a study of RM should be interested in combinations of primary activity with secondary activities across the whole multitasking continuum. Criteria for determining whether a certain secondary activity is in a concurrent or sequential relationship with the primary activity depends on research goals and positioning on the abstraction continuum.

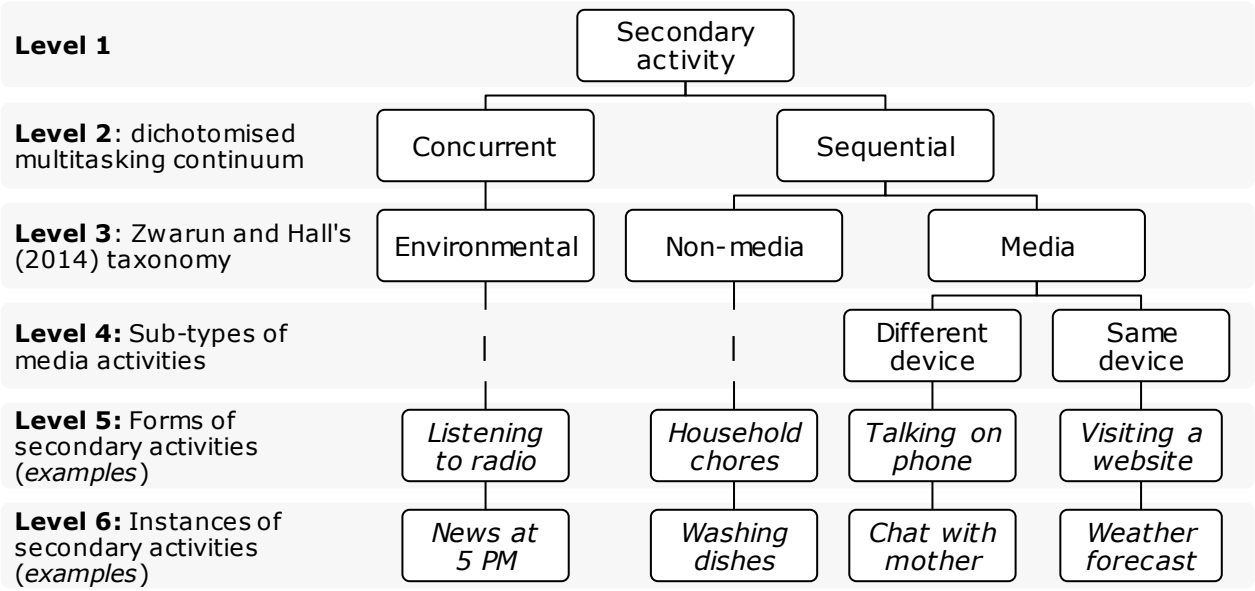
(e) Considering survey lengths, research on RM can generally occur on all time scales of the abstraction continuum except for the social band. Positioning on this continuum also depends on the data collection method. For example, it is quite unrealistic to prompt respondents to report on the length of their secondary activities in milliseconds or seconds.

(f) Research on RM can occur on both sides of the application continuum. For example, experiments in cognitive laboratories can investigate the

effects of secondary activities in details, while applied studies can assess the prevalence of RM and its relationship with survey performance in real-world situations.

In the second part of this Section, I will present the taxonomy of secondary activities in web surveys. The taxonomy is based on a paper by Zwarun and Hall (2014). The authors differentiate between three types of secondary activities: *environmental activities* (which do not require respondents to "task-switch by shifting their primary attention from the survey", but have "the potential to take up cognitive processing"), *non-media activities* (which require task-switching to non-media activities), and *media activities* (which require task-switching to other media activities) (2014, 239).

Figure 2.2: Taxonomy of secondary activities in web surveys



Drawing on the overview of multitasking and media multitasking literature, I further detail this taxonomy into two directions, thereby providing a more exhaustive classification of RM. Following the multitasking continuum, environmental activities can be associated with concurrent multitasking, while non-media and media activities belong to the sequential multitasking dimension. Following conceptualization of

media multitasking in Section 2.2.1, secondary media activities can be divided into two subtypes: *same-device media activities* refer to media activities using the same device that is used for responding to a questionnaire (i.e. *survey device*), while *different device media activities* refer to media activities on other devices.

Taxonomy can also be extended into two lower levels. Level 5 refers to *forms* of secondary activities. For example, same-device secondary activities can occur in different forms such as writing a school assignment or visiting a website.

Level 6 refers to specific *instances* of secondary activities. Each form of secondary activities can usually describe specific complete activities that are not related to each other. For example, in cases of complex cognitive activities, they require a separate problem state. To continue with the example, imagine a scenario where respondent visits two different websites while responding to a web survey, one to check the weather forecast, the other to check the stock market. Each website is considered as a specific instance of this form of secondary activity.

In practice, Levels 5 and 6 cannot be so rigorously defined or observed as the above levels. They partly depend on several research aspects, such as how granularly the activities are defined and the abstraction continuum. Still, these concepts need to be in consideration when discussing RM research.

The described taxonomy is presented in Figure 2.2 and will be used in the remainder of this dissertation. It is notable that concurrent activities are not further divided into media/non-media subcategories as it is done for sequential activities. While such distinctions could useful and relevant for RM research, it is beyond the scope of this dissertation to investigate these types of activities in details.

Lastly, it is important to define *survey suspension* or *interruption*. This relates to the concept of task switch and suspension of the primary

activity due to secondary activities, as introduced in Section 2.1. In the context of RM, survey suspension happens when the majority of cognitive resources are allocated to secondary activities. The frequency of survey interruptions is not necessarily related to the frequency of instances of sequential activities. For example, a single interruption could be related to multiple instances of secondary activities (e.g. respondent takes a break from a survey, writes an e-mail, goes for a walk, and then continues with a survey). However, a single instance of secondary activity could be related to multiple interruptions (e.g. respondent switches to a web chat with a friend multiple times during the survey).

In this sense, survey suspension cannot be included in the taxonomy as presented in Figure 2.2. It is better to understand it as an additional dimension of RM.

2.3.2 Proposed relationship between multitasking and performance in surveys

In survey methodology, respondents' performances in a survey are typically assessed with *response quality* (RQ) indicators. RQ indicators refer to those aspects of data quality that are influenced by respondents' individual differences in perception, understanding, and interaction with a questionnaire (Ganassali 2008).

Response process has a central role in explaining the relationship between RM and RQ (e.g. Holbrook et al. 2003; Lavrakas et al. 2010; Lynn and Kaminska 2012). Response process theories attempt to explain how people respond to survey questions. The most widely adopted information processing model of the response process by Tourangeau, Rips, and Rasinski (2000) includes four stages: *comprehension* (i.e. understanding the question and what information is requested), *retrieval* (i.e. retrieving specific information related to the question), *judgment* (i.e. evaluating and integrating the retrieved information) and *response* (i.e. mapping the judgment into available response categories). Each stage has several

demanding cognitive processes[10]. If a respondent does not dedicate sufficient cognitive resources for all stages (or does not even complete them), this can result in a lower RQ.

Based on general multitasking and survey methodology literature, as well as on RM studies, I suggest that RM could be related to RQ on the following three levels:

1. **Causal relationship between RM and response process.** Involvement in secondary activities could delay or disrupt the execution of these processes. Kennedy (2010) suggests that secondary activities can distract all four stages of the response process: 1) comprehension might suffer due to the inability to completely process the entire question; 2) respondents might employ less demanding retrieval strategies; 3) at the judgment stage, respondents might not fully contemplate all possible considerations; and 4) in the final stage, respondents might fail to optimally map their judgment onto the required response format. Lynn and Kaminska (2012) provide a similar elaboration.

2. **Interdependent relationship between RM and survey engagement.** In the context of web surveys, engagement refers to how respondents are focused and motivated to respond to a survey; It is suggested that less engaged respondents are less likely to dedicate necessary cognitive resources for the optimal completion of a response process (Callegaro et al. 2015). It is possible to assume that RM and engagement are in an interdependent relationship: less engaged respondents are more likely to start multitasking; multitasking leads to a lower engagement (e.g. the respondent might rush through the survey so that they can give full attention to the secondary activity).

---

[10] For more details, including an informative overview of other models, see Berzelak (2014).

Therefore, even if secondary activities do not directly influence the response process, they could still be related to it due to its relationship with survey engagement.

3. **Chronic or high media multitaskers.** Based on discussed literature on media multitasking index (MMI) in Section 2.2.3, it is reasonable to assume that chronic or high HMM could have a particular approach to response process due their typical attentional problems. These issues can occur whether HMM are engaged in secondary activities or not.

These three levels are not mutually exclusive and there could also be an important interplay between them. As an example, studies show that HMM are less successful in maintaining focus on the primary activity (e.g. Ophir et al. 2009; Stone et al. 1999; Lui and Wong 2012). Therefore, the independent relationship between RM and survey engagement could be particularly strong for such respondents, meaning that HMM are more likely to start multitasking if they are not very engaged in the survey since this is a part of their habitual behavior.

Moreover, when discussing relationship on the first two levels (i.e. causal relationship between RM and response process; interdependent relationship between RM and survey engagement), it is important to keep in mind factors that generally influence the relationship between media multitasking and quality of performance (see Section 2.2.3). Here, I specifically discuss how characteristics of respondent multitasking can influence its relationship with response process and/or survey engagement:

- Whether the problem state of response process needs to be reconstructed. Secondary activities that require the same required resources as the response process (i.e. cognition resources), are more complex and/or results in longer interruptions increase the possibility that a respondent will need to problem state related to a survey. This increases the task switch cost, meaning that respondent needs to

increase the effort to optimally complete the survey. To give a specific example, on one hand listening to music in the background typically requires perpetual and not cognition resources, is not a complex activity or does it result in longer interruptions. On the other hand, the opposite is true for following TV news or having a conversation about a serious topic.

- Timing and expectedness of survey interruptions. Secondary activities have a particularly negative influence if they interrupt the primary activity during a high mental workload. In this regard, I broadly separate survey interruptions into two categories: (a) interruptions that occur during a response process and (b) interruptions that do not occur during a response process (e.g. respondent interrupts a survey after completing a question and before starting to read the next question). Moreover, it is necessary to differentiate between (c) response process interruptions that occur during cognitively demanding survey questions (e.g. questions about personality traits, everyday behaviors, etc.) and (d) less cognitively demanding questions (e.g. demographic questions). For the former, the mental workload during a response process is higher and interrupting it could be especially disruptive.

- Prior knowledge of primary and secondary activities. The importance of prior knowledge the primary activity is emphasized in both, survey methodology literature (e.g. Olson and Parkhurst 2013; Ansolabehere and Schaffner 2015; Callegaro et al. 2015) and general multitasking literature (e.g. Srivastava 2010; Salvucci and Taatgen 2011; Alexopoulou et al. 2015). The latter stream also elaborates on the importance of prior knowledge of secondary activities. Based on this, it is possible to deduce that multitasking of respondents who are more familiar with survey topics and/or have more experiences with responding to web surveys is less likely to be in a strong relationship with response quality. Moreover, routine secondary activities are less

likely to be disruptive than activities for which respondent does not have many prior experiences.

Existing research on RM typically investigates the relationship between RM and two specific respondent behaviors that can be measured with RQ indicators:

Non-response behavior. Bosnjak and Tuten (2001) provide a detailed taxonomy of response and non-response behavior in web surveys, based on the number of answered questions and number of displayed questions[11]. First, there are respondents who view all questions by successfully navigating to the last page of the questionnaire that contains questions. Such respondent can be either: a *complete responder* (i.e. respondent has responded to all questions), an *item non-responder* (responded to some questions) or a *lurker* (did not respond to any question). Next, there are respondents who dropped out of the survey before reaching the last page containing questions. They can be divided into *answering drop-outs* (they answered to all displayed questions), *item non-responding dropouts* (they answered to some displayed questions), and *lurking drop-outs* (they did not respond to any displayed questions). Finally, *unit non-respondents* are those who do not view or answer any questions since they do not reach the first page containing questions.

Callegaro et al. (2015) discuss two relevant factors for non-response behavior that can be associated with all three levels of relationship between multitasking and RQ: *time resources* and *perception of burden*. As an example, if engagement in secondary activities decreases related cognitive resources required for response process, respondents might

[11] As Bosnjak and Tuten (2001) note, not all non-response behavior is voluntary. It can be a consequence of technical problems or other situational factors, as well as poorly constructed questionnaires (e.g. none of the available options in a closed question is appropriate for a respondent so they do not provide a response).

need longer time to complete a question or a whole survey and/or experience survey as a larger burden. This could result in a non-responding behavior. In a more extreme scenario, respondents might habitually engage in secondary activities losing their interest or even forgetting about the survey they were responding to.

Satisficing. Satisficing occurs when respondents do not "expend the mental effort necessary to generate optimal answers" and instead "settle for generating merely satisfactory answers" (Krosnick 1991, 215).

Satisficing in web surveys can be observed in the several forms such as: selecting the first response alternative that seems reasonable (*primacy effect*), *agreeing with assertions* (i.e. tendency to agree with or accept any assertion, regardless of content), *endorsing the status quo* (e.g. choosing the "keep things as they are" type of response to questions about social or political changes), non-differentiation in using rating scales (*straight-lining*), saying "don't know" and mental coin-flipping (i.e. randomly choosing an answer to close-ended question), etc. (Krosnick 1991).

Krosnick (1991) identifies three general factors that foster satisficing: *difficulty of the task* (i.e. required cognitive resources to respond to a question), *respondent's ability and respondent's motivation*. The latter two can be related to the three levels of relationship between multitasking and RQ. Respondent's ability can be affected by secondary activities, lower engagement or everyday chronic media multitasking behavior, while motivation is a major component of respondent's engagement.

2.3.3 Overview of approaches for measuring respondent multitasking

To the best of my knowledge, there are currently four publications that investigate RM in web surveys to a considerable extent: three papers where this was the a main research focus (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Sendelbah et al. 2016) and a small,

but important part of one PhD dissertation (Antoun 2015). These four studies will be the focus of Section 2.2.3 and Section 2.2.4. The topic of RM is also briefly touched upon several studies that report on share of respondents who have taken long breaks during a survey (e.g. Heerwegh 2005; Stieger and Reips 2010; Beckers et al. 2011).

In Section 2.2.3, I discuss measurement approaches used in the published literature, while empirical findings are presented in Section 2.2.4 along with other key methodological aspects. I separate discussion in three parts: (a) reactive approaches (self-reports), (b) non-reactive approaches (paradata) and a (c) concluding discussion.

(a) Three studies have used **self-reports** to measure RM (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Antoun 2015). In all three cases, researchers included closed-ended questions on RM at the end of the questionnaire. These questions were developed by authors specifically for each study and differ considerably.

Question wordings all three studies gravitated more towards descriptions of forms and not in instances of secondary activities. For the purposes of his research, Antoun (2015) was only interested in how many different secondary activity forms respondents were engaged in, while the other two studies prompted respondents to report whether they were engaged in specific secondary activity forms. However, the selection and description of forms that were included as response options also differed considerably (cf. Zwarun and Hall 2014; Ansolabehere and Schaffner 2015). For example, only Zwarun and Hall (2014) include a question, particularly targeted at concurrent or environmental secondary activities (it consists of three separate forms: background conversation, background music and background video). They are also the only ones to ask respondents how many survey suspensions occurred due to sequential secondary activities.

In addition, Zwarun and Hall (2014) and Antoun (2015) included questions that can be used as a supplement to understanding RM behavior, such as whether respondents have felt distracted during the survey, their location at the time of responding (e.g. home, at work, in transit) and the survey device.

It is worth mentioning that the overall reactive approach described above in studies on web surveys (i.e. closed-ended questions at the end of the questionnaire) is considerably different from reactive approaches found in published studies on other survey modes. Research on telephone surveys (CATI mode) has used open-ended questions (Kennedy 2010; Lavrakas et al. 2010; Lynn and Kaminska 2011) and interviewer observation (Pew Research Center 2006), while Schober et al. (2015) investigated RM in short telephone and text messaging surveys via subsequent debriefing telephone interviews.

While I could discuss questions on RM in published studies in greater details, I believe that I have already made the case on their diversity. Overall, these pioneering studies on RM developed questions based on specifics of their research settings, goals, and other factors.

It is safe to assume that many of general issues of reactive measurements in media multitasking research (Section 2.2.3) also apply here. While some issues might not be so accentuated in the case of RM (e.g. reference periods are short compared to research on everyday media multitasking), others could be more prevalent (e.g. some respondents might not want to report that they have not been fully focused on a survey due to social desirability bias).

(b) As already elaborated in Section 1.1, there are two general approaches for investigating RM in web surveys with **paradata**.

- First approach relates to *response times*. Influence of secondary activities on longer completion times has been well established in general (media) multitasking literature (see Section 2.1 and

particularly Section 2.2.4). Similarly, there is a long line of research on response times in survey methodology[12]. However, the relationship between (longer) response times and RM have been scarcely investigated.

On the other hand, the relationship between response times and RQ has been more often a subject of empirical research. The focus of such studies is predominantly on *speeders*, i.e. respondents who complete the survey in such a short time that it is reasonable to assume that they did not optimally or fully complete their response process to all questions (Matjašič 2015). While some studies report that speeding is associated with worse RQ in some aspects (e.g. Malhotra 2008; Callegaro et al. 2009; C. Zhang and Conrad 2014), other authors stress that removing speeders from the final sample does not change the substantive findings (e.g. Greszki et al. 2015; Thomas and Barlas 2014; Barlas et al. 2016).

Outside the context of RM research, I am not aware of any study that would directly investigate the relationship between lower RQ and longer response times. As Callegaro et al. (2009, 7) state, longer response times in survey methodology "have a positive connotation and are linked to optimizing strategies and higher levels of engagement with the questionnaire; exactly the kind of response behaviors that every survey researcher hopes their respondents will use". However, some authors do note that certain respondent behaviors that substantially

[12] Web survey researchers have used paradata to explore how much time respondents need to complete a question (e.g. Heerwegh 2005; Couper and Kreuter 2013), a page (e.g. Kaczmirek 2008; Gutierrez et al. 2011) or a whole questionnaire (e.g. Heerwegh 2004; Malhotra 2008) for a variety of other reasons. Based on a literature review by Olson and Parkhurst (2013), factors of long response times have been associated with different respondent-dependent factors (e.g. cognitive abilities, engagement in the survey) or questionnaire-dependent factors (e.g. poor question wording, usability problems).

prolong response times could also be negatively associated with RQ (e.g. Heerwegh 2003; Stieger and Reips 2010; Campbell et al. 2017).

The analysis of response times with regards to RM faces two major challenges. First, it is important to account properly for a variety of factors that can also contribute to longer response times (see footnote 12). A promising response for this can be found in several recent studies on response times in web surveys (Yan and Tourangeau 2008; Weyandt 2014; Gummer and Roßmann 2015). While these studies did not investigate RM, their approach could be extended to counter this particular challenge.

These studies have used linear mixed models with response times as the independent variable. Observed explanatory variables of response times are included in the model as fixed effects predictors. These typically include respondent characteristics (e.g. age, education, and experience with the web)[13] and questionnaire characteristics (i.e. variables expressing the structure of the questionnaire or its elements; e.g. frequency of all questions in the survey, frequency of words on a questionnaire page, question format type)[14]. Unobserved factors of response times for individual respondents or elements of the

---

[13] Older respondents and less educated respondents had slower response times in all three studies (Yan and Tourangeau 2008; Weyandt 2014; Gummer and Roßmann 2015). Greater level of experience with the Internet was associated with shorter response times in both studies that included this indicator in their models (Yan and Tourangeau 2008; Weyandt 2014). Findings on effects of gender are inconclusive (cf. Weyandt 2014; Gummer and Roßmann 2015).

[14] As one could expect, response times of surveys (Gummer and Roßmann 2015), pages (Weyandt 2014) and questions (Yan and Tourangeau 2008) increase with a higher word count and with a greater number of questions or response options. Moreover, response times of pages and questions become longer towards the end of the survey (Yan and Tourangeau 2008; Weyandt 2014). All three studies report that different question designs and contents also influence response times.

questionnaire can be accounted by random effects. Residual analysis of such models could reveal which response times are particularly long even after accounting for several key factors.

The second challenge relates to determining when a response time is long enough to be categorized as an instance of RM, i.e. determining threshold or a *cut-off point* that defines a *time-out event*. As Campbell et al. (2017, 18) state, the question "how fast is too fast and how slow is too slow" is one of the fundamental questions in response time analysis, but also non-trivial and rarely investigated in depth. This challenge is related, but also notably different from the detection of extreme outliers of response times that appears in several studies (e.g. see an overview in Matjašič (2015)). The main difference is that detection of extreme outliers is concerned only with particularly long (or short) response times. Here, the issue is more complex since it is reasonable to assume that RM is a more common form of respondent behavior.

Several authors have tried to identify which respondents have taken longer breaks while responding to a web questionnaire and used very different approaches for determining the cut-off point (e.g. Heerwegh 2004; Beckers et al. 2011; Sendelbah et al. 2016). Reported shares of respondents who recorded at least one time-out event are considerably different (Section 2.2.4 for more details).

Overall, even with an extensive additional research, determination of cut-off points will probably be always at least partly arbitrary and prone to false negative and false positive detections of RM. Another limitation is a general lack of qualitative information on secondary activities. Furthermore, the frequency of time-out events also does not equate to the frequency of survey suspensions or instances of secondary activities. It is also important to note that the maximum number of time-out events is bounded by the level at which response times are

measured (e.g. if time-out events are based on page response times, a respondent cannot have more time-out events than the number of pages he has visited).

Despite these considerable limitations and particular characteristics, I believe it is useful to investigate further the relationship between RM and response times. In addition to all benefits of non-reactive measures outlined in Section 2.2.3, response times have two specific advantages. Firstly, they are by far the most researched paradata type. Knowledge from existing literature could be useful to stimulate and inform further research in the context of RM. Secondly, compared to some other types of paradata, response times are widely available across different survey platforms. This could be another important facilitator in the future research.

- The second approach looks at respondents' **temporary inactivity** within the web questionnaire. Compared to response times, this is a much less commonly investigated topic in RM and in general web survey methodology literature (Callegaro 2013).

  I am aware of two papers that have used this approach in relation to RM (Stieger and Reips 2010; Sendelbah et al. 2016). Both papers utilized different paradata types and subsequent criteria for determining inactivities. Stieger and Reips (2010) analyzed activities by input devices. The authors defined long inactivity if no activity was recorded for at least 5 minutes.

  In Sendelbah et al. (2016), *focus-out* events were determined with analysis of paradata files containing information on when respondents lost and regained focus on the browser window or tab that contained the web questionnaire. This change of focus happens when respondent switches to another window or tab. Thus, such events can generally reflect survey suspensions that start due to same device activities.

Compared to time-out events, inactivity-based measures are not bounded by an upper limit of possible observations and can thus present a more realistic and detailed picture of RM, especially on survey suspensions and sequential (same-device) secondary activities. Moreover, it is reasonable to assume that the possibility of false negatives and false positives is smaller compared to time-out events (but still notable)[15].

A specific limitation of focus-out events is that they prioritize measurement of sequential activities, while time-out events can theoretically account for a wider variety of secondary activities. Other disadvantages of inactivity-based measurements compared to time-out events have already been explained (i.e. a much smaller adoption in survey methodology literature and across survey platforms).

(c) In **conclusion**, while I have already elaborated on the general differences of reactive and non-reactive approaches in Section 2.2.3, it is useful to discuss how some of these key differences are actualized in the context of RM research.

Overall, self-reports allow for a wider and deeper investigation of RM. This is not only holds for measuring prevalence and characteristics of secondary activities, but also for exploring the motives and potential effects of RM on RQ (e.g. questions on whether respondents have felt distracted as used in Zwarun and Hall (2014) and Antoun (2015)).

On the other hand, paradata events offer more accurate and precise measures of the observed forms of RM without increasing the respondent

[15] For example, an example of a false negative in a case of focus-out is a respondent who occasionally suspends responding to a survey due to an audio conversation over Skype, but never changes focus outside the web questionnaire. An example of a false positive is the already mentioned case of a respondent switching focus to another browser tab to search for information related to the survey.

burden. Besides the already elaborated problems with over- or underreporting, it is also important to note that questions on RM usually need to be placed at the end of the web questionnaire. In addition to missing out on respondents who have dropped out before reaching questions on RM, scholars have observed that item non-responding and satisficing tends to increase towards the end of the questionnaire (e.g. Galesic and Bosnjak 2009). Moreover, lowering the respondent burden is also important in the light of the general issue of over-surveying (Callegaro et al. 2015).

Technological development has different consequences for reactive and non-reactive approaches. The former need to be especially careful to include new prevalent forms of secondary activities in response options of close-ended questions (Wallis 2010). The main issue here for paradata-based approaches is that innovation and evolution of technology used for accessing web surveys need to be carefully followed by assuring that paradata events are accurately measured for a variety of software and hardware combinations.

Lastly, it is important to note that both approaches for measuring RM are still in the nascent stages. In fact, I am aware of only one study that would triangulate or validate measurements of RM. Ansolabehere and Schaffner (2015) have compared the frequency of self-reported distractions with the number of time-out events, determined by whether page response times were longer 5 minutes. The authors report that there is a "clear positive relationship between the two measures" (Ansolabehere and Schaffner 2015, 239).

2.3.4 Overview of empirical findings on respondent multitasking

Before presenting empirical findings, it is important to overview, key methodological aspects of the four studies in focus (their measurement approaches have already been explained in Section 2.2.3).

The Paper by Zwarun and Hall (2014) investigated RM on a survey conducted on a sample of members from a commercial non-probability online panel. Their analysis included 5,853 panelists from seven different countries older than 18 years. The survey was about 30 minutes long and questions were about "the policies, economy, facilities, culture and society" (Zwarun and Hall 2014, 239). Survey devices, i.e. devices used for responding to the survey in this study included PCs, tablets and smartphones.

Paper by Ansolabehere and Schaffner (2015) included four different political surveys also conducted among members of a commercial non-probability online panel in the USA. Surveys varied considerably in terms of the sample, length and other factors. Similarly to Zwarun and Hall (2014), respondents could use all prevalent web survey devices.

Data in Sendelbah et al. (2016) came from a survey on incoming and outgoing members of a student exchange program at the University of Ljubljana. The survey was about 13 minutes long. The main topics were students' satisfaction with exchange program and their experiences of living and studying abroad. The analysis included 267 respondents that used PC to respond to the web survey.

Research by Antoun (2015) was done on an academic probability online panel in the Netherlands. However, due to the design of the experiment, the final sample included in his research was not representative for the Dutch population. Author conducted a crossover design experiment among respondents older than 15 years. Specifically, 1,390 panelists were split into two equally sized groups. The first group responded to a web questionnaire via smartphone. A month later, the group responded to the same questionnaire, but this time by using a PC. The order was opposite

for the second group. Median completion time for web surveys on PC was about 10 minutes[16].

Prevalence

Firstly, I attempt to present the results on overall prevalence rates of RM, i.e. share of respondents who were engaged in at least one instance of RM. Due to differences in methodology and reporting, comparisons between studies are severely limited even for such simple statistic.

The most comparable are Ansolabehere and Schaffner (2015) and Antoun (2015) who both report on the overall prevalence of all observed forms mainly related to sequential activities. Zwarun and Hall (2014) have used most extensively reactive approach with regards to accounting for concurrent and sequential secondary activities. Unfortunately, they do not report on the overall prevalence for all observed or for all sequential secondary activities. They do provide prevalence rates separately for all media activities[17] and for each measured non-media activity. As already mentioned, in the paradata approach by Sendelbah et al. (2016) focus-out events relate mainly to same-device media activities, while time-outs generally account for all types of RM.

Table 2.3 summarizes the most related prevalence rates from these four studies. Considering the elaborated differences, it is not surprising that

---

[16] Unless explicitly noted otherwise, all further discussion on this study is limited to findings related to web surveys on PC.

[17] One of the sequential media activities in this study is worded in the questionnaire as: "I heard a noise notifying me that I had received an instant message, voice mail, text or email". I argue that this is hardly a case of sequential activity since it does not necessarily imply that respondent has switched his attention and key cognitive resources from the survey. Still, authors include observations of this activity in their reported prevalence rate of sequential activities that are presented here in Table 2.3. In further discussion, I will not consider this activity as a part of sequential activities, unless explicitly referring to this footnote.

prevalence rates range from 22% for the shortest survey included in Ansolabehere and Schaffner (2015) to 62% in Sendelbah et al. (2016).

Table 2.3. Reported overall prevalence rates of respondent multitasking in web surveys

	Survey length (minutes)	**% of respondents engaged in at least one instance of (sequential) secondary activity**	
Zwarun and Hall (2014)	30	Media	30
		Temporarily leaving computer	16
		Direct conversation	6
Ansolabehere and Schaffner (2015)	10, 12 (2 surveys)	Media & non-media	22-37
Ansolabehere and Schaffner (2015)	27, 32 (2 surveys)	Media & non-media	~50
Antoun (2015)	10	Media & non-media	44
Sendelbah et al. (2016)	13	Focus-outs (Media)	40
		Time-outs (Media & non-media)	53
		Both	62

Next, I look at the most commonly self-reported forms of secondary activities. This information is available in Zwarun and Hall (2014) and Ansolabehere and Schaffner (2015). To broadly summarize, the most common forms of both studies are: watching TV, taking a break and/or physically leaving the survey device, having a face-to-face conversation, and various media activities. There are considerable differences between the results of these two studies. This can at least partly be contributed to question design since the selection of secondary activity forms that are included as response options and overall question wording notably differ.

All three studies based on reactive measures compute the number of unique secondary activity forms observed per respondent, i.e. *range of*

activities. Antoun (2015) reports that the majority of multitasking respondents were engaged in one activity (25% of all respondents), followed by two (14%), three (4%), and four or more activities (1%). A similar distribution can be observed in Ansolabehere and Schaffner (2015). Zwarun and Hall (2014) do not report descriptive statistics on this measure, but do use it in their explanatory models that will be discussed further bellow.

Related, but conceptually different findings are reported in Sendelbah et al. (2016) with regards to survey suspensions as measured by focus-out events. Looking at the distribution of focus-out event frequencies and total duration on a respondent level, we encounter a Poisson-like shape (similar is also observed for time-out events). About 40% of respondents recorded at least one such event.

A specific insight gained by focus-out events in Sendelbah et al. (2016) is the duration of such survey suspensions. An average duration of all focus-out events per respondents was 90 seconds, but this distribution is again strongly skewed to the right. In this 13-minute survey, about 20% of respondents spent more than a minute on focus-out events and 11% more than two minutes.

As already mentioned, several studies investigated the prevalence of time-out events with very differently defined cut-off points. Heerwegh (2005) looked for respondents who spent more than half of their total survey time on a single question (30% of respondents exceeded this cut-off point). In another example, Beckers, Siegers, and Kuntz (2011) searched for respondents who spent more than 30 seconds per item, four minutes per page, or more than seven minutes to complete a five-page questionnaire (in total, 23% of respondents exceeded at least one of these cut-off points). In Sendelbah et al. (2016), cut-off points are determined separately for each questionnaire page based on median and median absolute deviation (MAD) of all response times by complete responders.

These two values were used to standardize respondents' page response times by firstly subtracting it with median and then dividing by MAD. If the standardized value was larger than 2, this was considered to be a time-out event (53% recorded at least one such event). The duration of these time-out events (i.e. for how much respondents' page times exceeded the cut-off point) was also highly skewed to the right.

Distribution of frequencies and durations of focus-out and time-out events in Sendelbah et al. (2016) thus resonates findings from studies on self-reports that while a considerable share of respondents is involved in at least one secondary activity form, a much smaller share are engaged extensively in RM. This is also reflected in the paper by Stieger and Reips (2010) where less than 4% of respondents were inactive for more than 5 minutes.

An interesting finding about RM and duration is reported by Ansolabehere and Schaffner (2015). As already mentioned in the Section 2.3.3, this paper found a positive correlation between the range of activities and page response times longer than 5 minutes. Moreover, the authors also investigated the effects of specific secondary activities forms on survey response timings. The largest time consuming effects were found for "taking a break", "doing a chore", and "answering a phone call", followed by "talking to an adult", "watching television", and "surfing the web" (Ansolabehere and Schaffner 2015, 230). For some other forms, most notably those referring to sending an e-mail or a text, no effects were found.

Factors

Zwarun and Hall (2014), Antoun (2015), and especially Ansolabehere and Schaffner (2015) have investigated factors that influence RM. I structure this part in the same way as the overview of factors in general media multitasking literature (Section 2.2.4). In order, I discuss (a) socio-demographic, (b) individual, and (c) structural/technical factors. Since

motives for RM were never a subject of analysis in published studies, I do not discuss them here.

(a) Zwarun and Hall (2014)[18] and Ansolabehere and Schaffner (2015)[19] both report that age is in a strong negative relationship with occurrence of all observed secondary activity forms. In Zwarun and Hall (2014), the share of respondents engaged in at least one form of media activities (see footnote 17) is 52% for the youngest age group (18-24 year olds) and drops steadily across the age groups to 17% for the oldest age group (65+ year olds). The effects are not so large for non-media and concurrent activities; however, they are still statistically significant in both papers. Ansolabehere and Schaffner (2015) also report that respondents who are married and/or have children are also more likely be engaged in related forms of RM.

(b) With regards to device ownership, Ansolabehere and Schaffner (2015, 226) report that respondents "whose only phone service is a cell phone are much more likely to text, e-mail, and talk to an adult, though less
likely to watch TV". The authors also investigate the influence of respondent's experience in web surveys, measured by the number of surveys respondent had taken at that panel provider. The effect of this factor is small, although statistically significant.

(c) The discussion here is mainly focused on the influence of survey device, i.e. differences between PC, smartphones and tablets. One could argue that this should be considered as an individual factor. However,

---

[18] Zwarun and Hall (2014) used chi-square analysis to compare the prevalence of RM between age groups.

[19] Ansolabehere and Schaffner (2015) used logit models with factors as the predictors. Therefore, they are able to observe the specific effect of each factor controlled for the influence of other factors. The model includes socio-demographic and individual factors, as well as survey device which I categorize as a structural/technological factor.

respondents are not always free to select the device for web surveys. For example, researchers can attempt to restrict certain devices for reasons related to methodological concerns, technical issues, research goals, etc. Moreover, reasons for device selections (by either respondents or researchers) are inherently tied to general socio-technological development.

In Ansolabehere and Schaffner (2015) study respondents were free to select the survey device. Respondents on desktop PC were the least likely to engage in secondary activities, followed by tablet and laptop users. However, smartphone users were by far the most likely to be multitasking during the survey. The authors also noted that patterns of RM differ from device to device. For example, tablet users were less likely to look or write e-mails during the survey than respondents on a desktop PC.

Antoun (2015) used the already mentioned crossover design that eliminates the self-selection bias and is focused on the comparison of PC and smartphones. Overall, 54% of respondents reported on being engaged in at least one secondary activity while being on a smartphone. This is significantly higher than the share of the PC (44%). Use of smartphone is also connected with a larger range of activities, the share of respondents using more than one session to complete the survey (40% for smartphones vs. 10% for PC) and share of respondents that were surrounded by people during the survey (37% vs. 30%).

Discussion on structural factors of RM can be also understood in a narrower sense. I am talking about the characteristics and structure of the web survey and its sub-elements (e.g. individual pages and questions). One obvious factor here is survey length – the longer the survey, the more opportunities for secondary activities to arise. This can be observed in Ansolabehere and Schaffner (2015) where generally the same question on RM was included in four surveys that were of different lengths, but relatively similar in some other crucial characteristics (e.g. all on political

topics and done on a commercial panel). As the authors note themselves, their data shows a clear positive link between the prevalence of RM and survey length.

Other characteristics such as complexity, question design, incentives, topic salience etc. could also be important factors for RM research. While more research is needed to investigate these potential factors, existing general survey methodology literature already highlights the importance of these factors for RQ and engagement. Some insight into this can be found in Ansolabehere and Schaffner (2015) who analyzed page response times. The share of respondents, who needed more than 5 minutes to complete a page was considerably larger for pages that required more cognitive effort. E.g. while shares of such respondents are lower than 7% for the majority of analyzed 47 pages, this figure jumps to 13% for a page that contains factual questions (even though the median time for this page is about 20 seconds). Ansolabehere and Schaffner (2015) also investigated whether a share of over 5-minute page times was related with page location, i.e. whether such time-out events were more common towards the end of the questionnaire. They did not find any notable patterns related to this. However, as the authors note themselves, a more detailed analysis is needed in order to draw broader conclusions for both of these characteristics.

Relationship with response quality and distraction

Here I summarize findings on (a) a relationship between RM and RQ indicators and (b) a relationship between RM and self-reported distraction.

(a) Ansolabehere and Schaffner (2015), Sendelbah et al. (2016), and Antoun (2015) have all investigated the relationship between RM and different sets of RQ indicators. Only in two cases, the authors confirm the connection between RM and lower RQ. Ansolabehere and Schaffner (2015) report that respondents who were engaged in secondary activity forms related to e-mail and talking to a child had a lower number of correctly

answered knowledge questions about US politics (after controlled for other secondary activities, education, age, and interest in news and public affairs). In Sendelbah et al. (2016), higher frequency of focus-out events was associated with a higher number of item non-response. However, for both studies, the effect sizes were considerably small.

Ansolabehere and Schaffner (2015) also found two positive relationships between RM and RQ: respondents who talked to an adult or surfed the web answered correctly to more knowledge questions. This is not so surprising if we consider that websites and adults could be used to get information on these knowledge questions. Effect sizes for these two relationships are small as well.

In most cases, however, there is no significant relationship in either direction. This was true for all RQ indicators included in Antoun (2015): rounding, short open-ended response, the number of incorrect answers to cognitive reflection test, avoiding half-open other response, and non-differentiation. The latter was investigated in Sendelbah et al. (2016) as well, and no relationship was found there either. Ansolabehere and Schaffner (2015) also report that there is no relationship for intra-item correlation with any of the observed secondary activity forms.

(b) Zwarun and Hall (2014) and Antoun (2015) both asked respondents how distracted they felt while responding to the web questionnaire. This can be an important insight into the relationship between RM and RQ. If respondents themselves report that they felt distracted, it is important to see how this was reflected on RQ indicators.

Zwarun and Hall (2014) report that self-reported measures of RM were positively associated with the self-reported measure of distraction. More specifically, all concurrent and non-media sequential activities, except for background music, were significantly positively associated with the level of self-reported distraction. By far the strongest relationship was found for

the range of media activities (see footnote 17). Unfortunately, the analysis in this paper did not include RQ indicators.

This implicitly resonates also in Antoun (2015) where participants were more commonly distracted when using a smartphone than a PC for responding to a web survey (this is parallel to the fact that more participants also reported on being engaged in RM while responding from smartphones). However, the author reports that there is no relationship between distraction and RQ indicators.

For Zwarun and Hall (2014) study, it is interesting to note that the frequencies of specific forms of media activity were mostly significantly negatively associated with self-reported distractions. In other words, self-reports reveal that respondents who more commonly switched to specific forms of media activities were less likely to feel distracted. The only exception is activity worded as "[leaving] the browser screen the survey is on to do another task on my computer" (Zwarun and Hall 2014, 243) where increased self-reported frequencies were associated with higher levels of self-reported distraction.

Zwarun and Hall (2014) also investigated other factors in relation to distraction. The level of self-reported distraction significantly varied in terms of location of respondents. The least distraction was observed for respondents who were at home, followed by those at work. Most distracted were respondents who replied to the survey from public spaces. Moreover, authors also investigated whether age is a factor in the relationship between RM and distraction. While the relationship was significant for all age groups, there was a curvilinear trend where the correlation was the strongest for 35- to 44-year-olds.

Discussion

In this discussion, I firstly summarize main empirical findings from the literature on RM in web surveys that were observed by at least two studies. I also place and contrast them with relevant insight from RM

studies on other survey modes and media multitasking literature. I conclude the discussion by listing some crucial limitations and research gaps.

For the majority of included surveys, it is possible to infer that over 40% of respondents have been engaged in at least one instance of sequential secondary activities. While this might seem a lot, existing literature also indicates that the share of respondents who have been frequently engaged in secondary activities is much smaller and follows the Poisson-like distribution. For example, only 5% of respondents reported on being engaged in more than two secondary activities in Antoun (2015) and only 8% of respondents recorded more than two focus-out events in Sendelbah et al. (2016).

Some similar patterns can be observed in studies on CATI surveys on mobile phones. E.g. in Lavrakas et al. (2010) about a half of respondents were multitasking (based on self-reports), and 16% were engaged in "distractive" secondary activities (as assessed by the authors of this study). In a Pew Research Center study (2006), interviewers observed that 20% of respondents were multitasking and 8% of respondents were marked as somewhat or very distracted by secondary activities.

An interesting question to ask here is: are people more likely to multitask during web surveys than during some other comparable media activities? Considering certain figures on everyday media multitasking, such as that people are multitasking almost half of the time while visiting websites (Voorveld and van der Goot 2013), and more than half of the time while using laptops and smartphones (Ofcom 2015), it seems that web surveys get about the same amount of (or a little bit more) attention as typical visits of websites on modern media devices.

Age is an important factor also for multitasking in web surveys. Younger respondents are much more likely to engage in secondary media activities. This is in line with observations that younger people are more

likely to be engaged in different forms of internal media multitasking. However, it is interesting to note that younger respondents in Zwarun and Hall (2014) and Ansolabehere and Schaffner (2015) were also more likely to engage in non-media secondary activities.

Another prominent factor is the survey device. Everyday media multitasking is generally more common on those modern media devices that are easier to carry around (e.g. Ofcom 2015). This resonates also in RM research. Smartphone stimulates more RM whether respondents are free to choose the survey device (Ansolabehere and Schaffner 2015) or not (Antoun 2015).

Overall, current empirical findings do not confirm the proposed relationship between RM and RQ. While there are some small indications, effect sizes are small (Ansolabehere and Schaffner 2015; Sendelbah et al. 2016). Similar outcomes have also been reported in studies on other survey modes (Kennedy 2010; Lavrakas et al. 2010; Schober et al. 2015). This is somewhat surprising, especially since respondents who are (more) engaged in secondary activities also tend to report that they are feeling more distracted (Zwarun and Hall 2014; see also Antoun 2015).

In relation to this, it is important to note that none of these studies directly investigated a causal relationship between RM and response process (i.e. the first level of relationship between multitasking and RQ, as elaborated in Section 2.3.2). Such investigation can be carried out only in an experimental study design that was not used in any of these studies. Furthermore, none of these studies investigated the relationship between everyday media multitasking habits and RQ (i.e. the third level of relationship between multitasking and RQ). Thus, all findings related to the relationship between RM and RQ indicators can be attributed to the investigation of the second level of this relationship, i.e. how the interdependence between RM and survey engagement reflects on RQ.

Moreover, as elaborated in Section 2.3.2, a complex set of factors could influence the proposed relationship between RM and RQ. It is possible that different statistical approaches are needed for this type of analysis. It is of course also possible that while a substantial share of respondents engages in at least one secondary activity and/or feels distracted, this does not generally reflect on RQ indicators.

One conclusion that can be (very carefully) drawn on the relationship between RM and RQ based on existing studies is that sequential media activities, and especially same-device media activities tend to be in a bit stronger relationship with RQ indicators and self-reported disturbance (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Sendelbah et al. 2016). This is not surprising since these activities are also among the most commonly observed activities in these studies. Moreover, it is safe to assume that such secondary activities are likely to be cognitively demanding and are requiring the same cognitive and perceptual resources as the primary activity of responding to a web survey. Furthermore, a considerable share of sequential activities has probably suspended the problem state related to responding to a survey. If this problem state needs to be reconstructed, this results in a specifically high task switch cost. Following the Threaded Cognition Theory (Section 2.1) and related empirical evidences in experiments on media multitasking (Section 2.2.4), such multitasking combinations could thus be especially disturbing for the response process (as elaborated in Section 2.3.2).

When interpreting these results on RM prevalence, factors and relationship with RQ, it is also important to keep in mind that little to no research was done on respondents that are not panel members. Research has shown that panel members, especially those from non-probability panels, can exhibit different patterns in behavior, and some of this differences could influence results on RM and/or RQ (e.g. see an overview in Callegaro et al. (2015)). Ansolabehere and Schaffner (2015) report that respondents who have taken more surveys at that specific panel provider were more likely

to be multitasking (although the effect is small). However, other important aspects still need to be researched. E.g. panel members typically have more experiences with web surveys than non-members. General (media) multitasking literature shows that increased experiences and knowledge of the primary activity mitigates negative effects of secondary activities on performance. It is reasonable to expect that this happens also in the context of web surveys.

It is also important to note that none of the included studies attempted to separate secondary activities (i.e. those that are not related to the primary activity) from complementary activities (i.e. those that can be considered as a part or support of the primary activity). Instead, both are reported together as secondary activities. As Ansolabehere and Schaffner (2015) have indicated, such complementary activities (e.g. looking for additional information in order to provide a response to survey question) do sometimes happen and have considerably different effects on RQ. Considering this, figures on the prevalence of RM might be overstated while effects on RQ might be understated.

Then again, it is also reasonable to assume that other factors contribute to understating prevalence rates of RM. In the case of the three studies based on reactive measures, this could be true because people tend to under-report on the multitasking behavior (Section 2.2.3). In the case of the study based on reactive measures, it is also safe to assume that focus-out and time-out events did not account for all forms of secondary activities.

There are of course numerous other methodological concerns to consider (as already elaborated throughout this Chapter, especially in Section 2.3.3). These concerns can be summarized as one the most crucial research gaps in current research on RM.

- **Research gap 1**: **Further development and evaluation of data collection procedures in RM research in web surveys.** It is

important to continue to evolve reactive and non-reactive approaches for measuring RM. This includes:

- Developing more elaborate questions that prompt respondents to report on different aspects of their RM behavior. In particular, questions need to account for a variety of concurrent and sequential secondary activity forms that can occur during the survey. In this context, Closed-ended questions can quickly become problematic, especially if they do not include some prevalent everyday media multitasking activities in a quickly changing technological society. Researchers here can take cues from studies done on telephone surveys (Kennedy 2010, e.g.; Lavrakas et al. 2010) and use (semi-)open questions; relevant insights on selection and wording of media activities can also be found in media multitasking literature.

- Moreover, the approach based on self-reports also need to account for the second dimension of RM behavior, i.e. interruptions during the survey. General literature on multitasking and media multitasking indicates that greater intensity of secondary activities can particularly degrade the quality of performance of the primary activity. Information on how often respondents interrupted the survey can thus be very valuable in research on RM that is concerned with its relationship with RQ.

- Developing and utilizing reactive approach. Reactive, i.e. paradata approach has not been fully utilized in research on respondent multitasking, despite the several benefits elaborated throughout this dissertation. To summarize, paradata could provide a cost-effective measurement of respondent multitasking, without increasing their respondents' burden and without being under the influence of bias when it comes to self-reporting on such behaviors.

On the other hand, researchers also need to be aware of limitations of such approach. Compared to self-reports, paradata cannot account for as a wide array of secondary activities and generally lacks qualitative information.

- Triangulation and/or validation of different data collection approaches. Currently, no published study on RM in web surveys (or any other survey mode) has systematically compared findings from different data collection approaches. Approaches can be compared on several levels. It would be useful to compare how different designs of questions on RM influence results. Moreover, the comparison between self-reports and paradata would provide an important insight on the differences and likeness between these two approaches. Lastly, a thorough validation of observations (obtained by either self-reports and/or paradata) is warranted, but also challenging to conduct without putting respondents in an unnatural environment.

- Overall, the research community should strive to develop as uniform as possible approach for measuring RM. Currently, it is hard to compare findings from different studies since each used a distinctive approach. Establishing some common ground on how to conceptualize and measure RM would be beneficial for this emerging stream of research in survey methodology.

Considering that only a handful of published studies, one could say that another crucial gap of RM literature on web surveys is a lack of empirical research. While it might be true that at this stage of research any additional study is a welcome addition to the field, researchers should be especially focused on exploring how RM is related to other respondents' characteristics and behaviors. Such insights can help to establish whether RM is an important respondent behavior or not. In relation to this, two crucial research gaps are:

- **Research gap 2**: Gathering more insight on factors of RM. Thus far, only Ansolabehere and Schaffner (2015) explored a wide variety of factors of RM. More research is needed to assess the importance of socio-demographic, individual and structural factors in relation to prevalence of RM as well as its relationship with RQ. Here, researchers could follow the research on factors of media multitasking research (Section 2.2.4), as well as the relevant findings from web survey methodology (Section 2.3.2).

- **Research gap 3**: Gathering more insight on the relationship between RM and RQ. All four studies included in this overview have explored this relationship to some degree. However, more refined methodological and analytical approaches could reveal some additional insight into the relationship between RM and RQ, especially if they manage to account for all key factors. It is also notable that quantitative measures of RM (e.g. in how many cognitively demanding secondary activities has respondent been involved in) tend to reveal a stronger relationship with RQ and distraction than simple dichotomous measures (Lavrakas et al. 2010; Zwarun and Hall 2014; Sendelbah et al. 2016; cf. Kennedy 2010; Ansolabehere and Schaffner 2015; Antoun 2015).

3 Research questions and hypothesis

In the introduction of this hypothesis, I have discussed on two general challenges in research on RM in web surveys (see Section 1.1.2). The first general challenge is "How to measure respondent multitasking?". After an overview of literature in the theoretical framework (Chapter 2), I have further elaborated on this challenge under Research gap 1 (Section 2.3.4): "Further development and evaluation of data collection procedures in RM research in web surveys".

The second general challenge is "What is the relationship between respondent multitasking (RM) and response quality (RQ)?". This is specifically further elaborated under Research gap 3: "Gathering more insight on the relationship between RM and RQ". However, Research gap 2 ("Gathering more insight on factors of RM") is also importantly related to this issue. As explained in the discussion on the proposed relationship between RM and RQ (Section 2.3.2), numerous factors can affect RM, RQ and their relationship.

This dissertation attempts to address the crucial research gaps by developing two data collection approaches for measuring RM (one based on paradata, another on self-reports) and using them in an empirical study.

A particular emphasis is given to the paradata approach, which has been underutilized in RM research so far. In this dissertation, paradata approach includes focus-out and time-out events. As described in an overview of approaches for measuring RM (Section 2.3.3), one focus-out event equals to a single same-device interruption. Further, time-out events are based on page-level response times. Whenever a certain respondent's time on a certain page is longer from the cut-off point, this

results in a time-out event. Detailed information on the development of paradata approach can be found in Section 2.3.2 which is then implemented into a specific empirical survey in Section 4.1.2.

Self-reports on RM are collected with questions that are developed with regards to existing RM and media multitasking literature, as well as specific research goals of this dissertation. Development of this approach is detailed in Sections 2.3.2 and 4.1.3.

Based on initial overview of research challenges (Section 1.1.2), I already outlined the four preliminary research questions, as well as the preliminary hypothesis (Section 1.2). Later when I fully elaborated the conceptual background, I further reformulated the challenges and use them to identify the three research gaps (Section 2.3.4).

Here, based on the conceptual work in Chapter 1 and 2 , I can reestablish – and elaborate them with much more precision – the research questions and the hypothesis in relation to the key theoretical and methodological concepts, which were discussed within conceptual discussion of the first two chapters. Considering the complexity of these specific topics and lack of research on them, research objectives are written in a relatively straightforward, descriptive manner. However, research objectives are further detailed with research sub-objectives and corresponding elaboration based on the theoretical framework.

Research question 1: How are the specific paradata observations (focus-out and time-out events) of RM associated with self-reports on multitasking?

This research question specifically addresses the lack of studies that would compare different approaches for measuring RM (as discussed under Research gap 1 in Section 2.3.4). Based on the theoretical framework, I expect that the association between these two approaches will be influenced by two general dynamics. Firstly, comparative studies of non-reactive and reactive approaches (Section 2.2.3) have shown that people

tend to underreport on their media and/or multitasking behavior. If this holds true also in the specific case of RM in web surveys, self-reports will underestimate the prevalence of secondary activities. Secondly, as elaborated in an overview of approaches for measuring RM (Section 2.3.3), focus-out and time-out events do not conceptually observe all forms of RM behavior (i.e. concurrent multitasking or short interruptions due to other-device and non-media secondary activities). In this aspect, paradata approach is likely to underestimate prevalence of certain secondary activities. An important part of analysis related to this research question is an approximation on how much these two dynamics influence the association between paradata and self-reports.

It is also important to account for conceptual differences between focus-out and time-out events. Therefore, this research question includes two subquestions:

> **Research subquestion 1a:** How are same-device interruptions observed using paradata (i.e. focus-out events) associated with corresponding self-reports?

Focus-out events allow for a granular analysis of this RM behavior, e.g., it is possible to observe an exact time and length of such interruptions. However, it is not rational to prompt respondents to self-report on their interruptions in such details for several reasons: comparative studies on media and/or multitasking behavior (Section 2.2.3) indicate that people have problems with recalling interruptions even on a less granular level; moreover, some respondents might perceive such questions as too demanding and/or intruding. A compromise I follow is prompting respondents to report on frequency of all same-device interruptions they had during a survey. Their responses will be compared with frequency of all focus-out events recorded for a respondent.

Research subquestion 1b: How are time-out events associated with self-reports?

While focus-out events are mainly limited to observing same-device interruptions, time-out events can account for a wider array of RM behavior. However, a major limitation of time-out events is a lack of insight on what specific secondary activity forms and/or types of interruptions of RM are behind them. In contrast to focus-out events, the concept of time-out events cannot be directly associated with questions (i.e. self-reports on RM) for respondents. Therefore, I will investigate the relationship of time-out events with all indicators based on self-reports; however, I do not expect that time-out events will as closely related with any indicator based on self-reports as in the case of the focus-out events described at the previous subquestion.

Research question 2: What is the prevalence of different types of respondent multitasking behavior in the web survey, according to specific RM indicators based on paradata and self-reports?

While this research question does not directly address any of the three specific research gaps identified in Section 2.3.4, it is a fundamental part of research on RM and inherently tied to the research goals of this dissertation. While all empirical studies on RM in web surveys have reported on prevalence of RM, the manner of conceptualizing and measuring RM, as well as reporting of obtained results varies considerably. This is probably one of the reasons on why reported findings on prevalence of RM in studies on web surveys range from 22% (Ansolabehere and Schaffner 2015) to 62% (Sendelbah et al. 2016).

I will attempt to provide a detailed insight into different types of RM behavior and deliver a holistic reporting on different aspects of prevalence of RM in the web survey. In order to do that, I develop three related research subquestions:

Research subquestion 2a: How many respondents have multitasked during the survey, according to paradata and self-reports?

This subquestion relates to the most basic inquiry on prevalence of RM: what is the prevalence rate of RM, i.e. share of respondents who have been engaged in any form of secondary activities? At the highest level, this subquestion can be answered by looking at the share of respondents who have a non-zero value of any RM indicator based on paradata or self-reports. More detailed, I will also look at prevalence rates of RM according to measurements via only paradata or only self-reports.

Research subquestion 2b: What are the most prevalent secondary activity forms based on self-reports?

As elaborated throughout the Chapter 2 (particularly in Section 2.2.4), qualitative information on what forms of secondary activities has a person been engaged in can be crucial for research on multitasking. This is especially true because some secondary activities are potentially more intruding for the primary activity. Since such information cannot be obtained with paradata, this subquestion relies on self-reports obtained via detailed questions that prompt respondents to report on their engagement in secondary activity forms.

Research subquestion 2c: How often do respondents interrupt the survey due to same-device (according to paradata and self-reports), other-device (according to self-reports) or non-media (according to self-reports) activities?

This subquestion is also tied to the relationship of RM and RQ. Research on media multitasking (Section 2.2.4) and RM (Section 2.3.4) indicates that certain types of sequential multitasking are more strongly related to the quality of primary activity's performance than concurrent multitasking. Following the taxonomy of RM presented in Section 2.3.1, interruptions of sequential multitasking can happen due to secondary activities on the

same device, other device or due to non-media secondary activities. All three types of interruptions can be measured with self-reports. In addition, same-device interruptions can be also observed with paradata (i.e. focus-out events).

Research question 3: What are socio-demographic and individual factors related to RM?

This research question directly addresses the second research gap identified in Section 2.3.4 – lack of insight on factors of RM which has so far been comprehensively investigated only in one study (Ansolabehere and Schaffner 2015). Based on the media multitasking literature, the general theoretical framework identified four groups of factors of general media multitasking behavior: socio-demographic, other individual factors and structural/technological factors, as well as motives. Full investigation of all these factors is outside the scope of this dissertation, specifically for structural/technological factors and motives. A thorough research on structural/technological factors would require an experimental design (e.g. see Antoun (2015)). Since no research on motives for RM has been published to date, a qualitative approach would be more appropriate for initial investigation on this group of RM factors (e.g. see research on motives for everyday media multitasking behavior done by Tokan et al. (2011)).

Considering the general research plan concerning other goals of this dissertation, it is viable to measure socio-demographic and individual factors with self-reports (by including relevant questions that measure these socio-demographic and individual factors of respondents) and then investigate their relationship with RQ indicators based on paradata and self-reports. Corresponding research subquestions are:

Research subquestion 3a: What socio-demographic factors influence RM?

Age is perhaps the most thoroughly investigated factor of multitasking, media multitasking and RM behavior. As shown throughout the Chapter 2 , younger people exhibit some specific tendencies of multitasking behavior. Not only do they multitask more commonly, they are also engaged in some specific forms of secondary activities that are not that common for older age groups (e.g. an intensive use of modern media devices and services). Age will be a prominently exposed factor in my analysis. In addition, I will also look at two other typical socio-demographic factors – gender and education – which have been investigated so far only in Ansolabehere and Schaffner (2015).

Research subquestion 3b: What other individual factors influence RM?

Other individual factors were included according to the elaboration in Section 2.3.2, where I expose two individual factors, which can importantly influence not only prevalence of RM but can also potentially have an important role in relation to RQ. This dissertation is thus particularly focused on these two factors:

- *Everyday media multitasking habits*. In the theoretical framework I have highlighted the importance of Media Multitasking Index (MMI) in media multitasking literature (Section 2.2.4). Moreover, I have also elaborated on how everyday media multitasking habits could be related to RM, as well as RQ (Section 2.3.2). Based on this, I include MMI in my empirical analysis as an individual factor reflecting everyday media multitasking of respondents.

- *Survey engagement*. Best to my knowledge, there is no standardized approach for measuring respondents' engagement in the survey. Still, various approaches do exist (e.g. trap questions, final evaluation, etc.).

This factor will be approximately measured by prompting respondents to report on how much effort and attention they have given to the survey participation. Due to lack of theoretical support, I restrain myself from referring to this directly as survey engagement. Instead, I simply talk about *self-reported effort* and *self-reported attention*.

Other important factors included in this research are general experiences with web surveys, motivations for being a member of the askGfK panel (on which the survey for this study was conducted), evaluation of this particular survey etc.

Research question 4: What is the relationship between RM indicators (based on paradata and self-reports) and RQ indicators?

This research question addresses the third research gap presented in Section 2.3.4, which is probably one of the main motivations for research on RM in general. As overviewed in Section 2.3.4, empirical studies on RM in web surveys did not find strong evidences that RM is related to worse RQ (this generally holds true also for studies on other survey modes). New methodological approaches used in this dissertation could reveal new insights into this relationship or confirm findings from previous empirical studies.

It is important to note that this dissertation does not attempt to investigate a causal relationship between RM and response process. Such relationship can be properly investigated only with experimental research design. Instead, analysis here is focused on the second (i.e. non-causal relationship between RM and RQ), and also the third level (i.e. non-causal relationship between everyday media multitasking and RQ) (see Section 2.3.2 for elaboration on these two levels of the relationship).

As also elaborated in Section 2.3.2, it is important to account for other factors that can influence RQ. This study specifically controls for socio-demographic factors (i.e. age, gender and education), self-reported effort,

self-reported attention and overall survey response time. The role of the latter factor is to control for respondents who are generally moving through the survey either very fast (i.e. speeders) or very slow (even after accounting for their RM behavior).

Following the overview of two types of respondent behavior associated with worse RQ in Section 2.3.2, this research question contains two subquestions:

> **Research subquestion 4a:** What is the relationship of RM indicators (based on paradata and self-reports) and unit-level item non-response?

Following the taxonomy of response behavior by Bosnjak and Tuten (2001) in Section 2.3.2, I am investigating respondents who view all questions and answer at least some of them (i.e. complete responders and item non-responders). More specifically, I am interested if unit-level item non-response of these respondents (i.e. the frequency of all non-responses per respondent) is related to RM indicators. Other types of response behavior (e.g. lurkers and drop-outs) are not included in this study because such behavior is not common enough among respondents from a commercial online panel (which served for the empirical study in this dissertation) for inclusion in the analysis.

> **Research subquestion 4b:** What is the relationship between RM indicators (based on paradata and self-reports) and satisficing indicators?

Based on the existing studies on RM, as well as on the literature from more general streams of survey methodology, I am focusing on the following specific indicators of satisficing: non-differentiation, length of responses to open-ended questions, inconsistent responding, trap item and estimation as the response strategy for behavioral frequency questions. More information on these indicators can be found in Section 4.1.5.

The elaboration of the above research questions and subquestions will provide a rich empirical insight into numerous aspects of RM, as well as into the relation between RM and RQ. Some of these findings will also serve when addressing the central hypothesis of the dissertation, which deals with the general potential of paradata approaches for measuring RM. Namely, the main hypothesis of this dissertation is related to the question, whether - and to what extent - the paradata approaches could be used for the detection and evaluation of RM. This is a very important conceptual and practical question. Namely, on one side, the issue of potential multitasking of respondents is increasingly important, as researchers want to know about respondents who were involved in (excessive) multitasking while answering the web survey. On the other side, the paradata approaches can provide an answer to this question by elegant technical procedures, which can run unobtrusively, i.e. without additional response burden for the respondents.

> **Hypothesis:** Paradata-based procedures (focus-out and time-out events) can measure respondent multitasking in web surveys and they provide similar general findings on prevalence of multitasking behavior, its factors and relationship with response quality, as the approaches based on self-reports.

This hypothesis is not directional – in essence, a null hypothesis is being evaluated here since the hypothesis states that there are no key differences between the two approaches. Based on the overview of studies that compared reactive and non-reactive approaches for measuring media multitasking (Section 2.2.3), one might think that a directional hypothesis would be appropriate within the context of this dissertation. However, here it is important to note that comparison of paradata and self-reports for measuring respondent multitasking differs from comparisons in the media multitasking studies in one important aspect. The media multitasking studies were focused on the single-device media multitasking. In other words, these studies were not comparing how the

two approaches measure other types of multitasking behavior where non-media and different-device media activities are involved. This is not the case in this dissertation where both approaches attempt to measure as wide a range of media and non-media secondary activities as possible. Best to my knowledge, there no research has been published that would compare reactive and electronic non-reactive approaches outside the context on single-device media multitasking.

Under discussion on Research question 1, I have mentioned that specific limitations of each approach (e.g. underreporting for self-reports; false positive and false negative for paradata) will importantly determine the outcome for this research question. Similarly can be assumed also for the hypothesis. However, with no prior empirical research, it is not possible to assume the direction of hypothesis.

More generally, it is important to stress that paradata has been much less utilized than self-reports in research on RM (and similar can be said for non-reactive approaches in research on media multitasking when compared with reactive approaches). Following the calls for a greater utilization of reactive approaches in research on media multitasking (e.g. Greenberg et al. 2005; Wallis 2010; Möller et al. 2013) and different applications of paradata in web survey methodology (e.g. Lynn and Nicolaas 2010; West 2011; Callegaro et al. 2015), the main purpose of this hypothesis is to establish whether a novel paradata-based approach can provide comparable substantial insights as the more traditional approach based on self-reports. If paradata approach does provide comparable insights (i.e. hypothesis is fully confirmed) while also being a more effective approach than self-reports (e.g. in terms of respondent burden), this would be a considerable stimulation for a greater utilization of this approach.

I will seek support for the hypothesis in a close link to the answers on four research questions, particularly with respect to Research question 1,

which deals directly with the relation between paradata and self-report approaches. In addition and more specifically, I will seek the support for the hypothesis with three streams of arguments, which will support the following sub-theses:

- **Sub-hypothesis 1:** Paradata-based indicators can identify majority of respondents that are multitasking according to self-reports. This claim is in direct relation to Research question 2.

- **Sub-hypothesis 2**: Respondents who are multitasking according to paradata approaches have similar relations with factors of RM as respondents who are multitasking according to self-reports. This issue is addressed already within Research question 3.

- **Sub-hypothesis 3:** Respondents who are multitasking according to paradata approaches have similar relations with RQ as respondents who are multitasking according to self-reports. This sub-hypothesis will be closely elaborated in relation to Research question 4, which investigates the relation between RM and RQ.

The above research questions and hypothesis will be elaborated with the findings from the empirical study, which is central to this dissertation. The Chapter 4 will thus introduce the corresponding research design of the study, systematically present the results and provide discussion of the findings, structured according to the above research question and hypothesis.

4 Case study on media multitasking in a web survey

This Chapter describes a case study that addresses the research questions and the hypothesis (Chapter 3). I firstly describe the study and its methodology (Section 4.1), then I present results (Section 4.2) and finally discuss the findings (Section 4.3).

4.1 Research design

The empirical study was carried out in a special survey, designed specifically for the needs of this dissertation. I first describe the research design, together with the questionnaire and instruments for measuring multitasking.

4.1.1 Study description

The empirical study was conducted in a 20-minute web survey that ran on the askGfK online panel from 2nd of June to 7th of July 2016. This commercial panel is not probability based; however, interlocking quotas involving gender and 10-year age groups were employed to ensure that the sample structure matches the Slovenian residents between 15 and 55 years old. Another non-interlocking quota was set on all 12 statistical regions in Slovenia. It should be noted that the panel only provided the respondents that were then redirected to open source web survey service (1KA), where the web questionnaire was hosted. This service 1KA was thus used for the actual data collection, including paradata recording.

An online preview of the questionnaire is available ("1KA Survey on Media Multitasking in Web Surveys" 2017). The first part of the web questionnaire contained questions about media behaviors, including an adapted version of the Media use questionnaire (Ophir et al. 2009). The

second part contained questions about different social and psychological aspects of respondents' lives, including the 10-item TIPI measure of the Big Five dimensions (Gosling et al. 2003) and a 16-item version of the Zimbardo time perspective inventory (ZPTI) (Zimbardo and Boyd 2015). The third part contained questions about respondent's attitudes towards participation in the askGFK panel survey, as well as questions about respondent multitasking and socio-demographic questions, which I describe in more detail later. In total, the questionnaire contained 61 questions with 202 items across 37 pages.

In total, 2,059 panelists have reached the first page of the questionnaire. The gross number of invitations sent to the panel members was not reported by GfK, but is also not important for the research aims of this study. We then excluded respondents who did not fulfill all of the following criteria: 1) they reached the last page of the questionnaire; 2) they responded to at least half of the items; 3) they used a PC (i.e. not a smartphone or a tablet) to participate in the survey. In total, 1,428 respondents fulfilled these criteria.

We further excluded 2% of the respondents with the fastest and 2% of the respondents with the slowest survey completion times, because we assume these are very specific users. This is a primitive, but practical approach for the elimination of extremes that has been previously used in survey methodology literature (e.g. see Matjašič (2015) for an overview). Four additional respondents needed to be excluded because of various technical problems. This left us with 1,366 respondents in our final sample. With this we get the bulk of typical respondents, which could be said to represent the "mainstream" behavior in a typical web survey. We believe that with elimination of extremes we did not lose the generality of our study.

The basic demographic structure of our sample is presented in Table 4.1. As expected with quota sampling, gender and age structure closely

correspond to the official data provided by the Statistical Office of the Republic of Slovenia (SURS). Age is equivalently distributed between both genders.

Table 4.1: Demographic structure of the study sample and comparisons with the official structure from Statistical Office of Republic Slovenia (SURS).

	f	%	SURS %
Females	679	49.3	48.3
Males	697	50.6	51.7
15-24 year olds	209	15.1	17.8
25-34 year olds	343	24.9	24.3
35-44 year olds	391	28.4	27.9
45-55 year olds	433	31.5	30.2
Low education level	139	11.2	37.8
Middle education level	421	39.8	35.3
High education level	603	49.0	26.8

I categorized respondents into three groups based on their finished education level. Respondents who have finished 4- or 5-year high school programme are labeled as “middle education”. Respondents with lower education level than the “middle education” group are labeled “low education", while respondents with higher level are labeled as “high education”. While education was not controlled for with the quota sampling, I illustratively contrast it with the official SURS data for the population within the same age range. As observable in Table 4.1, the share of respondents with middle education in this sample is relatively comparable with the population’s share (39.8% vs. 35.3%). There are however considerable differences between for low (11.2% vs. 37.8%) and high (49.0% vs. 26.8%) education groups. Overall, panel members are

more highly education than the general population, but this is an effect that can be found in almost any survey.

4.1.2 Data collection with paradata

I extensively elaborated the corresponding definitions, context and usage of paradata already in Chapter 2. Here I only focus on the specific implementation of paradata collection in this empirical study.

Figure 4.1 presents a simple excerpt from a paradata file produced by 1KA. Overall, the output contains all observed paradata events. Each specific event is logged as a new row in the output and mapped to specific respondents (the first column in Figure 4.1) and timestamp (the second column). In total, the paradata output for this survey contained 592,686 rows of data.

Figure 4.1: The paradata captured in the survey

ID USER	TIMESTAMP	EVENT	PAGE	QUESTION
13135984	31.5.2016 16:21:39	entered page	11835985	
13135984	31.5.2016 16:21:43	lost focus	11835985	
13135984	31.5.2016 16:21:49	got focus	11835985	
13135984	31.5.2016 16:21:53	data change	11835985	vrednost_7148707
13135984	31.5.2016 16:21:56	data change	11835985	vrednost_7148706
13135984	31.5.2016 16:22:10	left page	11835985	

For **focus-out** events one needs to look at the events names as "lost focus" (i.e. when has a specific respondent changed the focus away from the window or a tab containing the web questionnaire) and "got focus" (i.e. when a respondent has changed the focus back to the window or a tab containing the web questionnaire). These two events are highlighted in blue in the example in Figure 4.1. A pair of "lost focus" and "got focus" events represent a single focus-out event and the differences between the timestamps of the two events in a pair are the length of a focus-out event (6 seconds in the above example).

The only technical caveat in the detection of focus-out events is that for some web browsers "lost focus" and "got focus" are logged every time a respondent proceeds from one questionnaire page to the next

questionnaire page. Due to this, some respondents had many short focus-out events. In order to counter this, our analysis only includes focus-out events longer than 5 seconds.

To summarize the discussion of focus-out events in Section 2.3.3, a single focus-out event denotes a single interruption due to same-device secondary media activity. We should also be careful not to equate a single focus-out event with a single instance of secondary activity. As an example, one focus-out event could indicate two or more same-device secondary activities (e.g., a respondent switches from the web questionnaire window to write an e-mail and check the weather forecast and then returns back to the questionnaire) or a combination of same-device activities with other types of activities (e.g., a respondent switches from the web questionnaire window to write an e-mail, goes for a walk, and then returns to the questionnaire). On the other hand, multiple focus-out events could be associated with a single activity (e.g. having a conversation via an instant messaging client with a friend throughout the survey process).

Moreover, we should note that false positives and false negatives can occur. An example of a false negative is having an audio conversation over Skype while Skype's program window is minimized. An example of a false positive would be an automatic switch to screensaver mode while the respondent leaves the PC to do some non-media activities (while this is also RM, it is not an interruption due to media activity on the same device).

Time-out events are registered when respondent's time spent on a specific page exceeds a defined cut-off point (threshold). Firstly, total page response times are obtained by calculating the differences between timestamps of "entered page" and "left page" events for each respondent on each page. If a respondent also has one or more focus-out events recorded on a specific page, the length of these focus-out events is

deducted from the page response times. This is done in order to not account for same-device interruptions twice (firstly with a focus-out and secondly with a time-out event). In Figure 4.1, the two events denoting entering a specific questionnaire page and proceeding to the next page are highlighted in green. In this example, respondent proceeded to the next questionnaire page after 31 seconds (i.e. the difference between time stamps of the lines highlighted in green). However, since he also recorded a 6-second long focus-out event, his "pure" page response time is 25 seconds.

As elaborated in Section 2.3.3, one of the key challenges in determining time-out events is to define the cut-off points. For the purposes of this research, I take cues from the studies using linear mixed models for response times analysis (Yan and Tourangeau 2008; Weyandt 2014; Gummer and Roßmann 2015). Data containing "pure" page response times for all respondents on all pages was analyzed with linear mixed model with crossed random effects for respondent and page factors (see Appendix B for more details on the model). Time-out event for a certain respondent on a certain page occurred if its standardized model residual was larger than one. Such approach should generally account for differences in page response times due to differences in the content of the pages (e.g. word count, frequency and complexity of questions, etc.) and due to specific respondents' factors that influence the overall speed of responding (e.g. general cognitive abilities, familiarity with participation in web surveys etc.).

Nevertheless, one should still be wary of false positives. For example, it is possible that a respondent generally answers questions relatively quickly, but has some specific problems with understanding questions on a particular page. This could result in a relatively longer response times on this page. Consequently, linear mixed model residual of this particular observation is larger than one, even though the respondent was not engaged in any secondary activity while being on this page.

Leaving the concerns of false positives aside, time-out events also have other important limitations. Compared to focus-out events, time-out events offer even less qualitative information on the nature of RM behavior. In the case of focus-out events, it is at least probable to conclude that interruption has happened at the device used for survey participation. On the other hand, time-out events can be a consequence of concurrent RM or interruptions due to non-media or other-device activities, or even same-device activities. Similar to focus-out events, one time-out event can relate to multiple secondary activities and one secondary activity can result in multiple time-out events.

Another important difference between focus-out events and time-out events is that the former have no theoretical upper limit (there is no theoretical boundary on how many focus-out events a single respondent can record during a survey session), while the number of possible time-out events per respondent is bounded by the number of questionnaire pages they have visited. In the context of this study, one respondent could not record more than 37 time-out events.

4.1.3 Data collection with self-reports

Following the literature overview of existing reactive approaches for measuring respondent multitasking in web surveys (Section 2.3.3), I have determined that new questions on RM need to be developed for the purposes of this study. Development was guided by specific goals: to obtain as much qualitative and quantitative information on RM via self-reports as possible (without overburdening respondents) and to achieve some degree of comparability with the paradata approach, as well as with the prevailing reactive approaches used in respondent multitasking and broader media multitasking literature.

Figure 4.2: English translation of Q1

What, if anything else, have you been doing **on any electronic device** while responding to this survey?
Multiple answers are possible. Please select all that apply.

- [ ] Browsing the web, reading online news and documents, reading e-books
- [ ] Texting, instant messaging or e-mailing
- [ ] Listening to music, radio, podcast or other audio content (e.g. TV in background)
- [ ] Talking on the telephone or other devices (including video chatting, e.g. Skype)
- [ ] Playing games (computer, video, web)
- [ ] Using social networks (e.g. Facebook, Twitter, etc.)
- [ ] Watching TV or video content (e.g. movies, series, news, YouTube clips)
- [ ] Working on text documents, presentations, spreadsheets, or similiar activities
- [ ] Other:
- [ ] *I was not engaged in any other activities on any device.*

What, if anything else, have you been doing while responding to this survey?
Multiple answers are possible. Please select all that apply.

- [ ] Eating, drinking, or preparing a meal
- [ ] Doing household chores (cleaning, washing dishes, doing laundry)
- [ ] Taking care of other people (e.g. children)
- [ ] Talking to a person face-to-face
- [ ] Listening to a person talking (e.g. attending a lecture)
- [ ] Shopping or running errands (e.g. bank, post office)
- [ ] Walking around (e.g. taking a walk)
- [ ] Using means of transport (e.g. car, bus, train)
- [ ] Other:
- [ ] *I was not engaged in any other activities.*

The reactive approach used in this study consists of two survey questions. Question 1 (or *Q1*) aims to gather qualitative information on secondary activity forms respondent has been engaged in. Media and non-media activities are separated as two sub-questions and listed eight common activity forms for each of the sub-question. The selection of listed activities was partially based on Ophir et al. (2009). Moreover, I have added an open-ended item and a non-substantive item ("I have not been engaged in any activities […]") at the end of each sub-question. If respondents selected the latter item, the selection of all other items in the

sub-question was automatically disabled. An English translation of Q1 is shown in Figure 4.2.

Figure 4.3: English version of Q2

While you were responding to this survey, how many times have you been interrupted for **more than 5 seconds**?
Count only interruptons longer than 5 seconds.

	0 times	1 time	2 times	3 times	4 times	5-9 times	10+ times
Interruptions for activities on **this device**							
Interrputions for activities on **other devices (e.g. TV, PC, tablet, phone, etc.)**							
Other interruptions (not related to any electronic device)							

Question 2 (*Q2*) prompts respondents to report on quantity of interruptions during the survey. Respondents were asked to count all instances of interruptions that occurred due to secondary activities on the device used for responding to the web survey, secondary activities on other media devices and non-media secondary activities. An English translation of Q2 is shown in Figure 4.3.

Of course, as elaborated in Chapter 2, the self-reporting on multitasking could be inclined towards underestimation. This is also linked to the danger of inconsistency if some respondents report on some multitasking activities in Q2, but they do not report on them in Q1. The opposite situation does not necessarily indicate inconsistent responding, because respondent can, for example, listen to music without interrupting the questionnaire.

4.1.4 Respondent multitasking indicators

Three indicators of RM are based on paradata approach as explained in Section 4.1.2. Specifically:

- ***P_FOCUS*** is the frequency of all focus-out events per respondent
- ***P_TIME*** frequency of all time-out events per respondent
- ***P_ALL*** is the sum of P_FOCUS and P_TIME

All three paradata-based indicators are ratio variables. Unless noted otherwise, they are used as such in the analysis reported in Section 4.2. In certain sections, I will refer to all respondents who have recorded at least one focus-out and/or time-out event. Abbreviations used for such respondents are:

- ***P_FOCUS Rs*** are respondents who have at least one focus-out event
- ***P_TIME Rs*** are respondents who have at least one time-out event
- ***P_ALL Rs*** are respondents who have at least one focus-out or time-out event.

With regards to self-reports, numerous indicators can be constructed. Firstly, there are several indicators based on responses given to Q1. Firstly, there are ratio variables that express the count of how many different secondary activities has respondent reported on in Q1:

- ***S_RM_MM*** indicates the range of media activities (***S_RM_MM Rs*** are respondents who reported on at least one media activity)
- ***S_RM_NM*** indicates the range of non-media activities (***S_RM_NM Rs*** are respondents who reported on at least one non-media activity)
- ***S_RM*** indicates the range of all (media and non-media) activities (***S_RM Rs*** are respondents who reported on at least one activity)

Secondly, there are ordinal indicators based on Q2:

- ***S_SEQ_SD*** is the number of self-reported same-device interruptions (***S_SEQ_SD Rs*** are respondents with at least one self-reported same-device interruption)
- ***S_SEQ_OD*** is the number of self-reported other-device interruptions

(***S_SEQ_OD*** Rs are respondents with at least one self-reported other-device interruption)

- ***S_SEQ_NM*** is the number of self-reported non-media interruptions (***S_SEQ_NM Rs*** are respondents with at least one self-reported other-device interruption)

- ***S_SEQ_ALL*** is a number of all interruptions (***S_SEQ_ALL Rs*** are respondents with at least one self-reported interruption)

Lastly, based on responses to both questions (Q1 and Q2), I derive a dichotomous indicator:

- ***S_ALL*** reflects whether respondent has provided at least one self-report on respondent multitasking in either Q1 or Q2 (***S_ALL Rs*** are respondents with at least one self-report on respondent multitasking in either Q1 or Q2)

Figure 4.4 below summarizes the definitions and labeling. It also shows the relation of key RM indicators to various types of multitasking.

Figure 4.4: Comparison of key RM indicators based on paradata and self-reports

Secondary activities		Self-reports		Paradata	
		Indicator	**Definition**	**Indicator**	**Definition**
Concurrent	**Environmental**	N/A	N/A	P_TIME	Frequency of all time-out events per respondent
Sequential	**Non-media**	S_SEQ_NM	First Q2 item		
	Different device	S_SEQ_OD	Second Q2 item		
	Same device	S_SEQ_SD	Third Q2 item	P_FOCUS	Frequency of all focus-out events per respondent
All		S_ALL	Any substantive response in Q1 or Q2	P_ALL	P_FOCUS + P_TIME

4.1.5 Response quality indicators

As elaborated in many details within the Section 2.3.2, the respondents' performance in surveys is evaluated with response quality (RQ) indicators. Various components of RQ exist and in Section 2.3.2 I have already selected the two aspects on which I will focus: nonresponse behavior and satisficing. The nature of the empirical data enables and allows us to include RQ indicators as listed below, where unit-level item nonresponse serves for the evaluation of non-response behavior, while the remaining RQ indicators serve for evaluation of satisficing. It is important to note that all indicators are computed in such way that a higher value denotes lower response quality, except for the indicator of length of open-ended responses (listed here at the 6th position).

1. *Unit-level item non-response*. This ratio-scale RQ indicator is computed by counting all unanswered items by a respondent. Here it is important to note that only a portion of the questionnaire was used for computation of this indicator. Some questions were excluded due to their design where a non-response is not necessary caused by behavior related to a lower response quality, but because respondents

did not have anything to report. In total, 63 items (29.7% of all survey items) were excluded from the computation of this indicator[20]. All respondents received the same number of questions in the included set of questions, so the index is fully comparable across all respondents.

2. *Estimation as the response strategy for behavioral frequency question.* This dichotomous indicator reflects whether a respondent has used a less cognitive demanding strategy to respond to a particular behavioral question, i.e. by providing a response based on a general impression instead of attempting to enumerate all relevant episodes related to the questions. The questionnaire contained a question prompting respondents: "How many of the last 14 days were you physically active continuously for 20 minutes or longer?". After providing a response, respondents were presented with a follow-up question: "Which of these best describes how you came up with your answer?" with four possible answers: "I knew the answer off the top of my head"; "I thought about my physical activities for each of the past days and add them up"; "I gave my estimate based on a general impression"; "Other". In case respondent selected the third answer, his value on this RQ indicator was one, otherwise zero. This indicator is used similarly as in a study of RM in telephone surveys by Kennedy (2010) which provides a further theoretical elaboration of the approach.

3. *Trap item*. At the end of a matrix question on the time perspective (see Q29 in Appendix A), I have added another item, which explicitly

---

[20] The excluded questions are related to IT ownership and use (questions Q5 and Q7 through Q13 in the Appendix A), media multitasking index (Q14-Q21), experiences with surveys (Q34-Q36). The two open-ended questions (Q25 and Q26) are also excluded since they are investigated in details with the satisficing indicator of the length of open-ended responses.

instructed respondents to select the answer "strongly agree". This indicator is expressed as a dichotomous variable, where respondents who have not selected "strongly agree" will have a value of 0, while all other respondents will have a value of 1 (including those who did not respond to this item). This dichotomous indicator thus reveals which respondents were not attentive enough to follow the specific instruction for this item. See Downes-Le Guin et al. (2012) for more discussion on this indicator.

4. *Composite non-differentiation index* and *straightlining*. Differentiation refers to variability of responses in matrix questions. A low variability of responses can be a sign of careless responding (Berzelak 2014). This composite index is computed in two steps. First, standardized differentiation index (Berzelak 2014) is calculated for each respondent for each of the two matrix questions in the survey: Big-5 and time perspective (without the trap item). Next, an average of these two indexes is computed. This composite index is a continuous variable with values between 0 (denoting a great variability of responses) and 1 (denoting no variability of responses in the two matrix questions, also referred to as straightlining).

5. *Inconsistent responses*. This dichotomous indicator reflects whether a respondent has provided at least one substantively inconsistent response to Big-5 or time perspective questions. As an example, if a respondent strongly agreed that he sees himself as "extraverted, enthusiastic" and also strongly agreed that he is "reserved, quiet", this means that he provided an inconsistent response.

6. *Length of responses to open-ended questions*. Our questionnaire contained two open-ended questions: the first one prompted respondents to describe what is "a good friend" and the second one to describe "the best friend". This ratio indicator is computed by summarizing the number of words provided to both questions.

7. *Short responses to open-ended questions*. This dichotomous indicator reflects whether respondent provided shorter than three-word responses to both open-ended questions.

4.2 Results

In Section 4.2 the results of the empirical study are systematically presented. I first overview the relation between paradata (Section 4.2.1) and self-reported evidence on prevalence of the multitasking. Next, I include the socio-demographics and other factors into the analysis of the prevalence of the respondent multitasking. (Section 4.2.2). Finally, I observe both approaches within the context of the relation between multitasking and response quality (Section 4.2.3).

Introduction of the each of following three Sections briefly describes the analytical approach used in the corresponding Section. Overall, each Section is a mix of descriptive and inferential statistics. In principle, statistical inference is not possible for non-probability samples; however, in practice researchers often use inferential approach also outside the context of probability samples. Following discussion by Callegaro et al. (2013), my analysis thus includes inferential statistics but that does not mean that findings are intended to be generalized to a general population. Instead, the main purpose of inferential approach here is to illustrate the most noteworthy findings in the context of online panels. All statistical procedures were done in R (R Core Team 2017) by using the functions from core packages, unless referenced otherwise.

4.2.1 Prevalence of respondent multitasking

I systematically elaborate in full details the paradata (based on focus-out and time-out events) and self-reported evidence (based on Q1 and Q2 as seen in Figure 4.2 and Figure 4.3) on RM.

The first part is dedicated to description and comparison of paradata indicators on RM. The second part presents on results from Q1 (i.e. self-

reports on secondary activity forms) and is particularly focused on determining which combinations of secondary activity forms are most prevalent. The third part is dedicated to description of responses on Q2 (i.e. self-reports on interruptions). The fourth part compares responses to Q1 and Q2 and the final part compares results from paradata and self-reports.

Depending on characteristics of indicators, comparisons are made by using proportion test, exact McNemar test (Fay 2010) or Mann Whitey U test. Investigation into most prevalent combination of secondary activity forms is assisted with cluster analysis (Maechler et al. 2017).

Focus-out and time-out events (based on paradata)

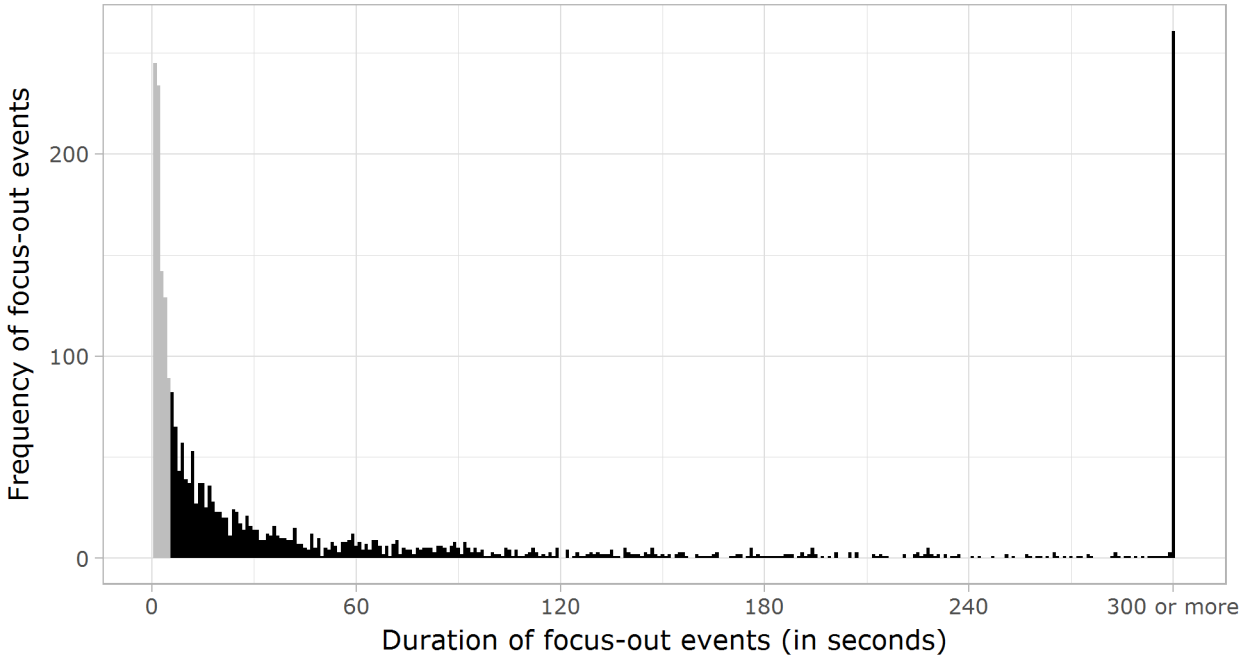
In total, paradata from this survey contained 2,596 focus-out events (for 1,366 respondents). Distribution of all focus-all events with regards to their duration is presented in Figure 4.5. About 32% of all recorded events lasted 5 seconds or less and are colored gray. These events are excluded from the further analysis due to reasons described in Section 4.1.2.

Overall, this distribution is heavily right-tailed. Median focus-out length is 38s (seconds), with 25% of events lasting less than 14s, and 25% of events more than 139s. Out of 1,595 focus-out events included in this analysis, 40 were longer than 1 hour, with the longest one lasting almost 2 days (about 45.5 hours).

I am of course more interested in the focus-out events aggregated on a respondent level. The share of respondents with at least one recorded focus-out event (i.e. the share of P_FOCUS Rs) is 37% (503 respondents).

Among these 503 respondents, the median value of focus-out events per respondent is 2. Frequency of focus-out events per respondent steadily drops from 1 being the most common (i.e. 39% of all P_FOCUS Rs recorded only 1 focus-out event) to the maximum of 43 events recorded for one respondent. Distribution of the P_FOCUS indicator is shown in Figure 4.6.

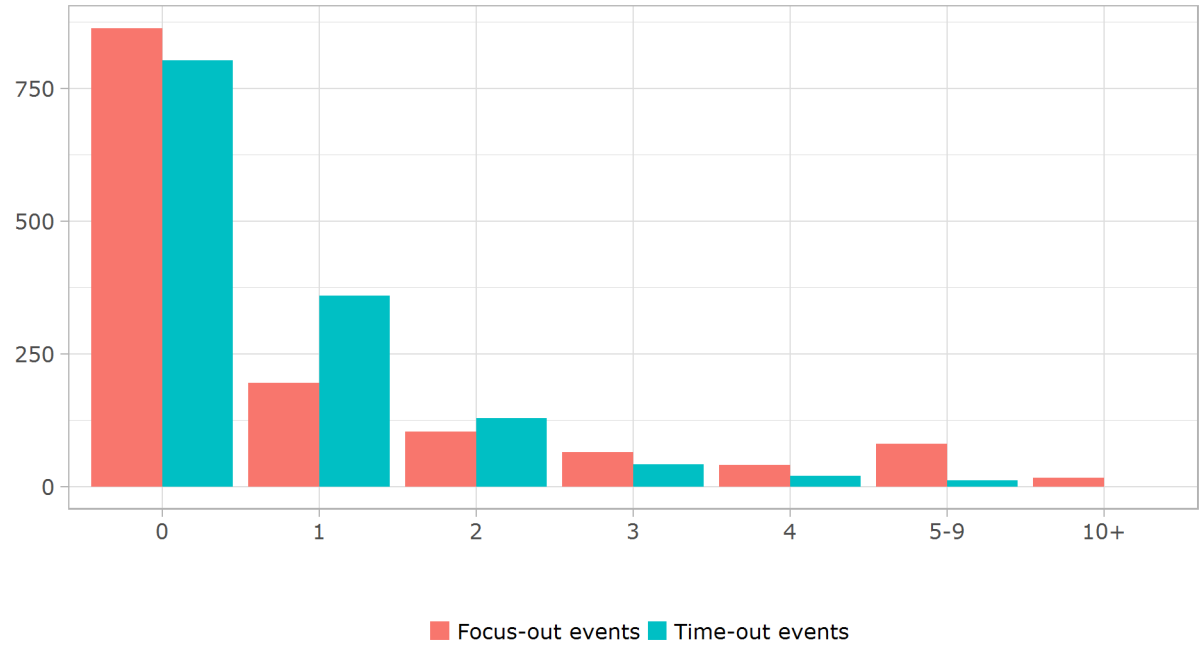
Figure 4.5: Distribution of focus-out events in relation to their length



For P_FOCUS Rs, it is also interesting to look at the total length of time a responded was not focused on the questionnaire, i.e. the sum of all focus-out durations recorded for a respondent. The median value of this indicator is slightly over 2 minutes (129s). Looking at its distribution, we again encounter a Poisson-like shape, with 25% of lengths being under 33s and 25% over 498s.

Now I turn my attention to time-out events. In total, 884 time-out events were determined for this sample. The share of respondents with at least one time-out event (i.e. the share of P_TIME Rs) is 41%. However, only 36% of these respondents have more than one time-out event and only six respondents have more than five time-out events. Distribution of the P_TIME indicator is shown along with the P_FOCUS indicator in Figure 4.6. The share of respondents with at least one focus-out or time-out event (i.e. the share of P_ALL Rs) is 59%. P_ALL indicator, i.e. the sum of P_FOCUS and P_TIME, is similarly distributed as its summands. Looking again at the 25th and 75th percentile, their values are one and four respectively.

Figure 4.6: Histogram of focus-out (P_FOCUS) and time-out (P_TIME) events per respondent



At this point, one might be worried that P_FOCUS and P_TIME significantly overlap with each other, i.e. that the same instance of respondent multitasking might be recorded by both paradata events. This is not the case. Here, I will test the assumption that these two paired observations are significantly different with an exact McNemar's test. For the purposes of this test, P_FOCUS is recoded into a dichotomous variable reflecting whether the respondent has at least one focus-out event. The same is done for P_TIME. As observable from Table 4.2 (left panel), both events were observed only for 19.5% of respondents, while 17% had only focus-out and 22% only time-out events. An exact McNemar's test reveals that the difference between these two variables is indeed statistically significant (p = 0.01).

Table 4.2: Overlap between focus-out and time-out events

	On respondent level		**On respondents' visit level**	
	P_TIME = 0	**P_TIME > 0**	**P_TIME = 0**	**P_TIME > 0**
P_FOCUS = 0	41.4	21.7	95.0	1.9
P_FOCUS > 0	17.3	19.5	2.7	0.3

To further confirm this, I conduct the same test also on a more detailed level: respondents' visit of pages. In other words, the basic unit of analysis here is each questionnaire page visit by respondent and the two dichotomous variables reflect whether the respondent had any focus-out or time-out events on this specific page. Data in such form contains 40,264 units (pages). The relationship between focus-out and time-out events on this level is shown in the right panel of Table 4.2. In 95% of cases, there was no event of any type, while both events appear only on 0.3% of respondents' visits. Not surprisingly, the exact McNemar's test reconfirms significant differences ($p < 0.01$).

Secondary activity forms (self-reports from Q1)

Firstly, I look at self-reports from Q1 (see Section 4.1.3). Responses from this survey question reveal what forms of secondary activities respondents have reported on being engaged during the survey participation. Almost a half of respondents reported on being engaged in at least one form of secondary activity. In other words, the share of S_RM Rs is 48%.

In total, respondents reported on being engaged in 0.8 activity form on average. S_RM – indicator on how many different activity forms has respondent selected – follows the Poisson-like distribution. The majority of S_RM Rs (62%) were engaged in a single activity form, followed by two (24%), three (8%) and so on. Seven respondents selected more than five different activity forms and one respondent even selected twelve.

Media activity forms were more frequently selected (on average 0.45 per respondent) than non-media (on average 0.33 per respondent). Share of respondents who selected at least one media activity form (i.e. the share of S_RM_MM Rs) is 35%. The share of respondents with non-media forms (i.e. the share of S_RM_NM Rs) is 26%.

Most frequently, respondents reported on listening to different audio contents (22%). Next are two non-media activity forms: "eating, drinking, or preparing a meal" (14%) and face-to-face conversations (10%).

Following the latter are two activity forms related to mediated communication: "texting, instant messaging or e-mailing" (6%) and audio/video conversations over the phone or other devices (5%). No other activity form was selected by more than 5% of respondents.

Such low prevalence rates of the majority of secondary activities impede analysis and interpretation in explanatory models on RM. One way to mitigate this is by clustering respondents based on similarity on self-reported engagement in activity forms.

Cluster analysis includes all 627 respondents who selected at least one activity form in Q1[21]. Jaccard distances between these respondents are computed, with each of the 16 activity forms from Q1 treated as a separate binary attribute (indicating whether the respondent has selected a specific activity form or not). The resulting set of dissimilarities is then used in a hierarchical cluster analysis with Ward's clustering method.

Different numbers of clusters were tried out. Ginal selection is primarily lead by interpretability of clusters and with this in mind, I have decided for seven clusters. Their characteristics are presented in Table 4.3 which in addition to indicators that were used in the cluster analysis (i.e. self-reports on secondary activity forms) also includes other indicators based on self-reports for illustrative purposes.

Interpretation of the first five clusters is relatively straightforward. This is particularly true for clusters #1 and #3 since they contain respondents who were solely engaged in "listening to music or other audio content"

---

[21] As described in Section 4.1.3, Q1 contained two open-ended items. These options were selected 139-times by 80 different respondents (6% of the whole sample). In five of those instances, respondents did not write in what the secondary activity was. Out of remaining 134 answers, I have recoded 96 into one of the existing categories (i.e. other items in Q1). The remaining 38 uncategorized answers most commonly refered to various job-related tasks (7 times), toilet (7 times), smoking (6 times) and resting (6 times). Due to low frequencies, these activities are not included in the cluster analysis.

and "talking to person face-to-face" respectively. From now on, I will refer to them as *Listeners* and *F2F talkers*.

Table 4.3: Characteristics of respondents in the clusters who reported certain activities based on self-reporting questionnaire.

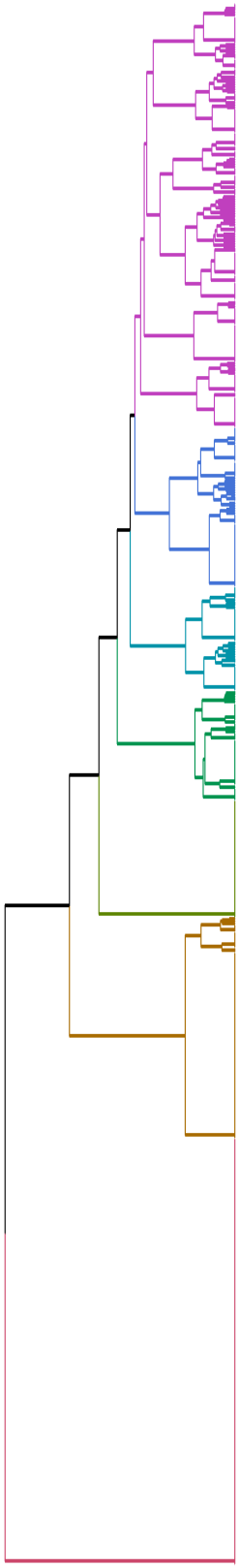
Activity forms	Cluster (% per cluster)						
	#1	**#2**	**#3**	**#4**	**#5**	**#6**	**#7**
Cluster size (n)	***171***	***89***	***47***	***44***	***42***	***64***	***170***
% browsing the web, reading online news and documents, reading e-books	0	2	0	2	7	2	21
% texting, IM or e-mailing	0	11	0	11	10	27	25
% listening to music or other audio content	100	0	0	20	19	81	31
% talking on the telephone or other devices (including video chatting)	0	1	0	0	100	2	10
% playing games (computer, video, web)	0	0	0	0	0	0	4
% using social networks	0	7	0	5	0	8	19
% watching TV or video content	0	0	0	100	2	2	6
% working on text documents etc.	0	0	0	0	5	3	14
% eating, drinking, or preparing a meal	0	100	0	9	21	81	15
% doing household chores	0	0	0	0	2	0	8
% taking care of other people	0	0	0	9	0	2	19
% talking to a person face-to-face	0	2	100	11	40	50	13
% listening to a person talking	0	0	0	0	2	3	24
% shopping or running errands	0	0	0	2	0	0	2
% walking around (e.g. taking a walk)	0	0	0	0	0	0	4
% using means of transport	0	0	0	0	0	2	1
% more than one media activity form	0	4	0	30	36	25	32
% more than one non-media activity form	0	2	0	2	17	36	16
% more than one activity form (any)	0	16	0	59	69	100	61
% more than one same-device interruption	10	15	17	14	24	33	38
% more than one other-device interruption	7	2	4	32	43	22	24
% more than one non-media interruption	11	16	45	14	26	39	29
% more than one interruption (any)	23	38	66	55	71	75	71

Respondents in cluster #2 are also relatively uniform. All of them have reported on "eating, drinking, or preparing a meal" while a small share (16%) of them were also engaged in other activities. Clusters #4 and #5 are similarly fixated on "watching TV or video content" and "talking on the telephone or other devices"; however, respondents in these two clusters are also more commonly engaged in other activities such as listening to music and face-to-face conversations. While it is important to keep variability within these clusters in mind, I will refer to cluster #2 as *Eaters*, #4 as *Watchers*, and #5 as *Phone talkers*.

In contrast to all other clusters, all respondents in cluster #6 reported on being engaged in two or more secondary activities. Most frequently, these respondents are engaged in activities that are among the most common in general: listening to audio contents (81% of all members of this cluster); eating, drinking, or preparing a meal (also 81%); face-to-face conversations (40%); texting, instant messaging or e-mailing (27%). All members of this cluster reported on being engaged in at least one of the following two activity forms: listening to audio contents or having a face-to-face conversation. I will refer to these respondents as *Common mixers*.

Cluster #7 contains the most diverse group of respondents. Not a single activity has been reported by more than a third of members of this cluster. Compared to other clusters, activities such as "browsing the web", "using social networks", "working on text documents", "taking care of other people", and "listening to person talking" are more frequently represented here. Arguably, most of these activities are typically more cognitively demanding than those that are more prevalent in some other clusters (such as listening to audio contents, eating, watching video contents). Moreover, a substantial share (61%) of members of this cluster also reported on being engaged in more than a single activity. Considering these characteristics, I will refer this cluster as *Complex mixers*.

Figure 4.7: Clusters of respondents according to reported secondary activity forms



Complex mixers: majority engaged in more than one activity (esp. media); relatively higher presence of certain cognitively demanding activities (when compared with other clusters); majority has more than one interruption (esp. due to same-device activities)

Common mixers: all engaged in more than one activity; most commonly listening to audio, eating, drinking or preparing a meal and having a F2F conversation; majority has more than one interruption.

Phone talkers: all have mediated audio/video conversations; majority engaged in other activities (esp. F2F conversation) and has more than one interruption (esp. due to other-device activities)

Watchers: all watching video contents; majority also engaged in other activities; about a third interrupts the survey due to other-device activities more than once

F2F talkers: all are only having a F2F conversation; about half of them interrupts the survey due to non-media activities more than once

Eaters: all are eating, drinking or preparing; rarely engaged in other activities or interruptions

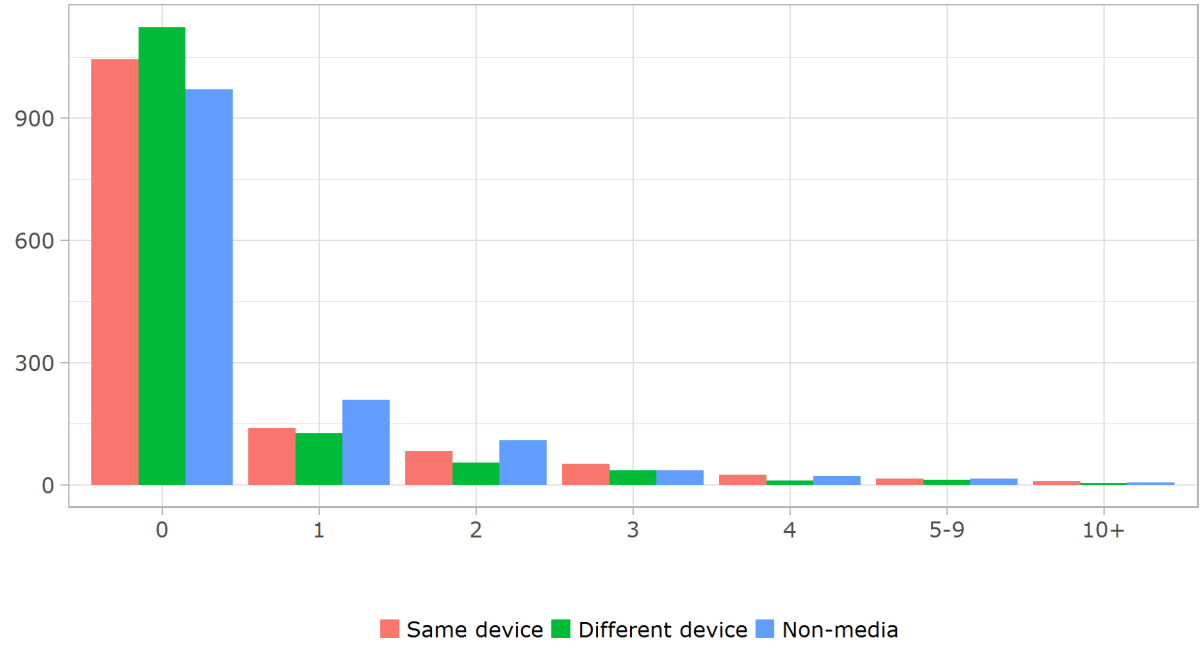
Listeners: all are only listening to audio contents; rarely interrupt the survey

While responses to Q2 on the frequency of interruptions will be separately described in the next part, it is useful to briefly look at how clusters based on responses Q1 differ in terms of responses to Q2. This can be observed in the last four rows of Table 4.3. Almost two fifths of Complex mixers reported on having more than one interruption due to same-device activities, followed by a third of Common mixers and about one quarter of Phone talkers. Phone talkers most commonly reported on having more than one interruptions due to other-device activities (43%), followed by Watchers (32%), Complex mixers (24%) and Common mixers (22%). Respondents with more than one interruption due to non-media activities are most common among F2F talkers (45%), followed by Common mixers (39%), Complex mixers (29%) and Phone talkers (26%). Overall, more than 70% of Common mixers, Complex mixers and Phone talkers interrupted the survey more than once due to an activity of any kind. On the other hand, shares of such respondents are considerably lower among Listeners (23%) and Eaters (38%). Summarized interpretation of clusters is presented in Figure 4.7.

Interruptions due to sequential secondary activities (self-reports from Q2)

Now I look at the responses to the second self-report question on RM – Q2 as presented in Section 4.1.3. A half of respondents reported on suspending the survey for at least 5s at least once (I will refer them to as S_SEQ_ALL Rs). Most commonly, these interruptions happened due to non-media activities. The share of all respondents who have reported on at least one non-media interruption (i.e. the share of S_SEQ_NM Rs) is 29%; in the case of same-device activities, this is true for 24% (S_SEQ_SD Rs) and in the case of other-device activities for 18% (S_SEQ_OD Rs) of all respondents. Distribution of all three indicators can be seen in Figure 4.8.

Figure 4.8: Distribution of self-reported interruptions



Comparison of self-reports based on Q1 vs Q2

To obtain full insight into RM prevalence according to self-reports, it is useful to further compare responses from Q1 and Q2. It is reasonable to assume that a respondent who reported on at least one interruption in Q2 (S_SEQ_ALL) would also report on being engaged in at least one activity form in Q1 (S_RM). In other words, a vast majority of S_SEQ_ALL Rs should also be S_RM Rs. However, 29% of S_SEQ_ALL Rs did not select any answer in Q1. An exact McNemar's test also confirms that these two groups are significantly different ($p < 0.01$).

However, one should not attribute these differences simply to careless responding. A quick check in Table 4.4 reveals that such respondents reported fewer task switches than those S_SEQ_ALL Rs who are also S_RM_Rs. So, it is possible that at least some of respondents regarded their interruptions to be too minor to warrant reporting on them in Q1.

Table 4.4: Mean values of the indicators on sequential activities (frequency of interruptions) for non-SR_RM Rs and SR_RM Rs

	Mean values for	
	non-S_RM Rs	**S_RM Rs**
S_SEQ_SD[a]	0.78	1.07
S_SEQ_OD[a]	0.34	0.81
S_SEQ_NM[a]	0.70	1.18

[a] Ordinal variables. Values 0-4 reflect the actual frequencies of self-reported interruptions, while the value 5 refers to all frequencies above 4. The mean thus relies on truncation (i.e. trimmed mean).

It is also interesting to look at the relationship between responses from Q1 and Q2 from the perspective of Q1. Of respondents who reported on being engaged in at least one secondary activity form in Q1, 25% did not report on having any interruption in Q2. Such respondents could thus be classified as concurrent multitaskers, i.e. respondents who were engaged in secondary activities but did not interrupt the survey process for more than 5s. Table 4.5 compares concurrent and other[22] multitaskers with regards to clusters of secondary activity forms. Prevalence of clusters is notably different between these two groups. Listeners represent 61% of all concurrent multitaskers and only 16% of non-concurrent multitaskers. A higher share among concurrent multitaskers can be also found for Eaters. All other clusters are much more present among non-concurrent multitaskers.

[22] One could refer to respondents who have reported on having at least one interruption as sequential multitaskers. However, I am reluctant to do so, since a notable share of such respondents could be engaged in both, concurrent and sequential secondary activities.

Table 4.5: Comparison of concurrent and other multitaskers across clusters of activity forms (%)[ab]

Q1: Clusters	S_RM Rs	
	Concurrent multitaskers	Other multitaskers
Listeners	61 (57)	16 (43)
Eaters	18 (33)	13 (67)
F2F talkers	1 (4)	10 (96)
Watchers	4 (16)	8 (84)
Phone talkers	1 (2)	9 (98)
Common mixers	4 (9)	12 (91)
Complex mixers	11 (11)	33 (89)

[a] Top of cells: % by columns.

[b] Bottom of cells, in brackets: % by rows.

To conclude the overview of the prevalence of RM, based on self-reports, I present an overview table with shares of key categories of respondents (Table 4.6). For example, share of respondents who have reported on at least one instance of multitasking in either Q1 or Q2 (i.e. S_ALL Rs) is 63%.

Table 4.6: Overview of prevalence of different types of RM according to self-reports.

	f	*% of sample*		
S_ALL: any engagement in secondary activity forms (Q1) or interruption (Q2)	858	63		
	f	*% of sample*	*% of S_ALL*	*% of S_RM*
S_RM: any reported engagement in secondary activity forms (Q1)	658	48	77	
S_RM_MM: any engagement in media secondary activities	475	35	55	72
S_RM_NM: any engagement in non-media secondary activities	358	26	42	54
Listeners	171	13	20	26
Eaters	89	7	10	14
F2F talkers	47	3	5	7
Watchers	44	3	5	7
Phone talkers	42	3	5	6
Common mixers	64	5	7	10
Complex mixers	170	12	20	26
No reported interruptions in Q2 (i.e. concurrent multitaskers)	170	12	20	26
	f	*% of sample*	*% of S_ALL*	*% of S_SEQ*
S_SEQ: any reported interruption (Q2)	688	50	80	
S_SEQ_SD: Any same-device interruption	322	24	38	47
S_SEQ_OD: Any other-device interruption	243	18	28	35
S_SEQ_NM: Any non-media interruption	395	29	46	57
No reported activity forms in Q1	200	15	23	29

Comparison of self-reports and paradata indicators on RM prevalence

As already indicated in Section 4.1.4, it is reasonable to believe that focus-out events (P_FOCUS) and self-reported same-device interruptions

(S_SEQ_SD) should be in a strong positive relationship since they measure relatively the same type of interruptions. However, doubts occur just by observing that the share of P_FOCUS Rs is notably higher than the share of S_SEQ_SD Rs (37% vs. 24%). These differences are significant according to exact McNemar's test ($p < 0.01$) and warrant further investigation.

Firstly, a comparison of simple dichotomous outcome of the two indicators on a respondent level (i.e. has a respondent recorded at least one focus-out event and/or reported on at least one same-device interruption) shows that they are in agreement (i.e. both zero or both are non-zero) only for 66% of respondents (see Table 4.13 at the end of this Section).

Specifically, 64% of P_FOCUS Rs did not report on having any same-device interruption. It is worth noting that this segment of P_FOCUS Rs has statistically significantly lower frequencies of focus-out events than the remaining 36% of P_FOCUS Rs who did report on having at least one interruption, as confirmed by a Mann-Whitey U test ($p<0.01$).

On the other hand, 43% of S_SEQ_SD Rs did not record any focus-out events. Similarly, as in the previous paragraph, this 43% of S_SEQ_SD Rs had lower frequencies of self-reported same-device interruptions than the remaining 57%, as confirmed by Mann-Whitey U test ($p<0.01$).

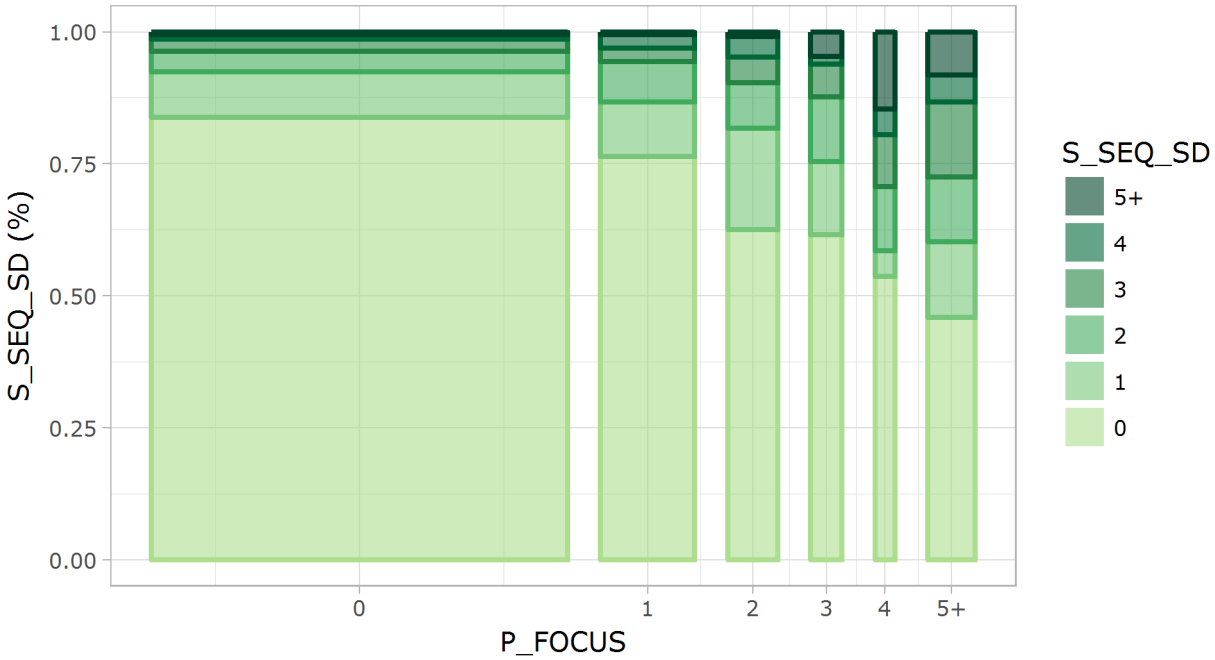
Moreover, I compare the frequencies of recorded focus-out events on the respondent level with how many same-device interruptions were reported by a respondent[23]. Frequencies of P_FOCUS are significantly higher than S_SEQ_SD according to the one-tailed Wilcoxon signed-ranks test ($p<0.01$).

---

[23] For the purposes of this comparison I recode the two indicators into ordinal variables with values 1, 2, 3, and 4 indicating the actual observed frequencies of interruptions, while 5 reflects all frequencies of 5 interruptions per respondent and beyond. Such ordinal recoding is used for all frequency indicators in the remainder of this Section.

This remains true also if I limit the test only to those respondents who have at least one focus-out event and at least one report on same-device interruptions (p<0.01). More specifically, 51% of such respondents have more focus-out events than self-reported same-device interruptions; 25% have more self-reported interruptions than focus-out events, and 24% have the same frequencies of both observations.

Figure 4.9: Spline plot of indicators for focus-out events (P_FOCUS) and self-reported same-device interruptions (S_SEQ_SD)



The overall relationship between these two indicators is visualized with a spine plot in Figure 4.9. P_FOCUS is shown on the x-axis, with width of bars representing the share of respondents who have recorded 0, 1, 2, 3, 4 or 5 and more focus-out events. Heights of differently colored areas inside each bar represent proportions of respondents (as shown on the y-axis) who have reported on having the corresponding amount of survey-device interruptions. We can observe that there is a positive relationship between the two indicators. This is also confirmed by one-sided Kendall tau-b independence test (p<0.01).

Table 4.7 and Table 4.8 provide a more complete comparison of self-reports of P_FOCUS Rs (i.e. respondents who have at least one focus-out

event) and non-P_FOCUS Rs (i.e. respondents who do not have any focus-out events).

Table 4.7 compares these two groups with regards to cluster membership (see Figure 4.7). While 62% of non-P_FOCUS Rs did not report on any secondary activity in Q1, this is true for 40% of P_FOCUS Rs. This difference in proportions is statistically significantly different. If we look at specific clusters, significant differences can be found for F2F talkers, Phone talker, Common mixers and Complex mixers. P_FOCUS Rs are all more likely to be members of these clusters.

If we look at differences in proportions from the perspective of self-reports, 73% (Table 4.7, in brackets at the bottom of each cell) of all respondents who did not report on any activity form in Q1 also did not record any focus-out events, while the remaining 27% did. This proportion is statistically significantly different from 50%. Large majority (67%) of phone talkers are also significantly more likely to have at least one focus-out event. On the other hand, the majority of Listeners (58%) and Watchers (68%) do not have any focus-out events.

Table 4.7: Characteristics of the respondents with no recorded focus-out events (non-P_FOCUS R) and respondents with recorded focus-out events (P_ FOCUS Rs) across the clusters of self-reported activity forms

Q1: Clusters	Non-P_FOCUS Rs	P_FOCUS Rs	Sig. Dif.
No secondary activities	62 (73)	40 (27)	* (*)
Listeners	11 (58)	14 (42)	(*)
Eaters	6 (55)	8 (45)	
F2F talkers	3 (47)	5 (53)	*
Watchers	3 (68)	3 (32)	(*)
Phone talkers	2 (33)	6 (67)	* (*)
Common mixers	4 (48)	7 (52)	*
Complex mixers	10 (48)	17 (52)	*

Top of cells: % by columns.

Bottom of cells, in brackets: % by rows.

* denotes a statistically significantly different proportions according to a two-tailed proportion test ($p<0.05$).

(*) denotes that the proportion of P_FOCUS Rs is statistically significantly different from 50% according to proportion test ($p<0.05$).

Table 4.8 compares non-P_FOCUS Rs and P_FOCUS Rs in terms of self-reported ranges of activity forms and frequencies of interruptions. P_FOCUS Rs reported significantly more on RM according to all these indicators. However, if we look from the perspective of self-reports, only respondents who reported at least one interruption due to same-device activities were more likely to also have at least one focus-out event.

Table 4.8: Characteristics of the respondents with no recorded focus-out events (non-P_FOCUS R) and respondents with recorded focus-out events (P_FOCUS Rs) with regards to ranges of secondary activity forms and interruptions.

		Non-P_FOCUS Rs	P_FOCUS Rs	Sig. Dif.
Q1: Range of activity forms	Media (S_RM_MM)	0.34 (49%)	0.66 (51%)	*
	Non-media (S_RM_NM)	0.26 (52%)	0.45 (48%)	*
	All (S_RM)	0.60 (52%)	1.11 (48%)	*
Q2: Interruptions	Same device (S_SEQ_SD)	0.30 (43%)	0.84 (57%)	* (*)
	Other device (S_SEQ_OD)	0.25 (51%)	0.49 (49%)	*
	Non-media (S_SEQ_NM)	0.37 (51%)	0.79 (49%)	*
	All (S_SEQ_ALL)	0.60 (53%)	1.19 (47%)	*

Top of cells: mean values of indicators for non-P_FOCUS Rs and P_FOCUS Rs.

Bottom of cells, in brackets: % of all respondents who have been engaged in at least one instance of the specific RM behavior (i.e. % by rows).

* denotes a statistically significant different indicator values between the two groups according to a Mann-Whitey U test.

(*) denotes that the proportion of P_FOCUS Rs is statistically significantly different from 50% according to proportion test ($p<0.05$)

Section 4.1.4 indicates that time-out events cannot be directly compared with any indicator based on self-reports. Still, it is interesting to note that 23% of all P_TIME Rs (i.e. respondents who have at least one time-out event) did not report on any multitasking in either Q1 or Q2. On the other hand, only 50% of respondents who did report on any RM in either Q1 or Q2 have at least one time-out event.

Moreover, it is interesting to compare self-reports of P_TIME Rs (i.e. respondents who have at least one time-out event) and non-P_TIME Rs (i.e. respondents who do not have any time-out events). These comparisons are shown in the next two tables. Table 4.9 is similar to

Table 4.7 and Table 4.10 is similar to Table 4.8, but comparisons are here made according to P_TIME indicator rather than P_FOCUS.

Table 4.9: Characteristics of respondents with no recorded time-out events (non-P_TIME Rs) and respondents with recorded time-out events (P_TIME Rs) across the clusters of self-reported activity forms

Q1: Clusters	Non-P_TIME Rs	P_TIME Rs	Sig. Dif.
No secondary activities	63 (68)	42 (32)	* (*)
Listeners	14 (64)	11 (36)	(*)
Eaters	5 (43)	9 (57)	*
F2F talkers	2 (32)	6 (68)	* (*)
Watchers	2 (41)	5 (59)	*
Phone talkers	2 (36)	5 (64)	* (*)
Common mixers	3 (42)	7 (58)	*
Complex mixers	10 (46)	16 (54)	*

Top of cells: % by columns.

Bottom of cells, in brackets: % by rows.

* denotes a statistically significantly different proportions according to a two-tailed proportion test ($p < 0.05$).

(*) denotes that the proportion of P_TIME Rs is statistically significantly different from 50% according to proportion test ($p<0.05$)

We can see that the results from Table 4.9 are similar to the ones in Table 4.7. However, the share of Watchers among P_TIME Rs (5%) is significantly higher from the share of Watchers among non-P-TIME Rs (2%). Moreover, the majority of F2F talkers (68%) are more likely to be P_TIME Rs. None of these differences were significant.

Table 4.10: Characteristics of respondents with no recorded time-out events (non-P_TIME Rs) and respondents with recorded time-out events (P_TIME Rs) with regards to ranges of secondary activity forms and interruptions.

		Non-P_TIME Rs	P_TIME Rs	Sig. Dif.
Q1: Range of activity forms	Media (S_RM_MM)	0.36 (50%)	0.59 (50%)	*
	Non-media (S_RM_NM)	0.22 (41%)	0.48 (59%)	*
	All (S_RM)	0.59 (48%)	1.07 (52%)	*
Q2: Interruptions	Same device (S_SEQ_SD)	0.37 (47%)	0.68 (53%)	*
	Other device (S_SEQ_OD)	0.20 (40%)	0.53 (60%)	* (*)
	Non-media (S_SEQ_NM)	0.34 (42%)	0.79 (58%)	* (*)
	All (S_SEQ_ALL)	0.83 (45%)	1.72 (55%)	* (*)

Top of cells: mean values of indicators for non-P_TIME Rs and P_TIME Rs.

Bottom of cells, in brackets: % of all respondents who have been engaged in at least one instance of the specific RM behavior (i.e. % by rows).

* denotes a statistically significant different indicator values between the two groups according to a Mann-Whitey U test.

(*) denotes that the proportion of P_TIME Rs is statistically significantly different from 50% according to proportion test ($p<0.05$)

Results from Table 4.10 also resonate those from Table 4.8. The key difference between tables is from the perspective of self-reports. While respondents who reported on at least one same-device interruption were more likely to be P_FOCUS Rs, they are not more likely to be P_TIME Rs. However, respondents with the other two types of interruptions (i.e. due to other-device or non-media activities), as well with any kind of interruption, are more likely to be P_TIME Rs, while this is not true in the case of P_FOCUS Rs. In addition to this, P_TIME indicator is significantly correlated with a combined indicator of self-reported other-device and non-media interruptions ($p<0.01$).

Table 4.11: Characteristics of respondents with no recorded time-out events (non-P_ALL Rs) and respondents with recorded time-out events (P_ALL Rs) across the clusters of self-reported activity forms

Q1: Clusters	Non-P_ALL Rs	P_ALL Rs	Sig. Dif.
No secondary activities	69 (53)	43 (47)	* (*)
Listeners	12 (39)	13 (61)	(*)
Eaters	4 (27)	8 (73)	* (*)
F2F talkers	1 (13)	5 (87)	* (*)
Watchers	3 (36)	4 (64)	
Phone talkers	1 (10)	5 (90)	* (*)
Common mixers	2 (19)	7 (81)	* (*)
Complex mixers	8 (26)	16 (74)	* (*)

Top of cells: % by columns.

Bottom of cells, in brackets: % by rows.

* denotes a statistically significantly different proportions according to a two-tailed proportion test ($p<0.05$).

(*) denotes that the proportion of P_ALL Rs is statistically significantly different from 50% according to proportion test ($p<0.05$)

Finally, I conduct similar comparisons for respondents who have recorded any of the two paradata events (i.e. P_ALL Rs) and those who did not (i.e. non P_ALL Rs). From Table 4.11 we can observe that prevalence of all RM clusters is larger among P_ALL Rs than non-P_ALL Rs. Differences are significant for all clusters except for Listeners and Watchers. From the perspective of self-reports, more than a half members from each RM cluster are P_ALL Rs. This is significant for all clusters except for Watchers. The largest shares can be found for Phone talkers (90% of them are P_ALL Rs) and F2F talkers (87%).

Table 4.12: Characteristics of respondents with no recorded time-out events (non-P_ALL Rs) and respondents with recorded time-out events (P_ALL Rs) with regards to ranges of secondary activity forms and interruptions.

		Non-P_ALL Rs	P_ALL Rs	Sig. Dif.
Q1: Range of activity forms	Media (S_RM_MM)	0.34 (28%)	0.66 (72%)	* (*)
	Non-media (S_RM_NM)	0.26 (23%)	0.45 (77%)	* (*)
	All (S_RM)	0.60 (27%)	1.11 (73%)	* (*)
Q2: Interruptions	Same device (S_SEQ_SD)	0.30 (20%)	0.84 (80%)	* (*)
	Other device (S_SEQ_OD)	0.25 (25%)	0.49 (75%)	* (*)
	Non-media (S_SEQ_NM)	0.37 (23%)	0.79 (77%)	* (*)
	All (S_SEQ_ALL)	0.60 (25%)	1.19 (75%)	* (*)

Top of cells: mean values of indicators for non-P_ALL Rs and P_ALL Rs.

Bottom of cells, in brackets: % of all respondents who have been engaged in at least one instance of the specific RM behavior (i.e. % by rows).

* denotes a statistically significant different indicator values between the two groups according to a Mann-Whitey U test.

(*) denotes that the proportion of P_ALL Rs is statistically significantly different from 50% according to proportion test ($p<0.05$)

Results in Table 4.12 reveal that P_ALL Rs have significantly higher values of all indicators related to ranges of activity forms and interruptions. Moreover, about three quarters of all respondents who provided any report on RM behavior in Q1 or Q2 are among P_ALL Rs. These proportions are significantly larger than 50% for all indicators included in this table.

In conclusion, I compare P_ALL Rs with respondents who reported on any instance of RM in either Q1 or Q2 (i.e. S_ALL Rs). The share of P_ALL Rs is 59%, while the share of S_ALL Rs is 63%. This difference is also statistically significant according to the exact McNemar's test ($p < 0.01$).

Table 4.13: Share of respondents (% of the sample) with same-device interruptions and any kind of multitasking according to paradata and self-report indicators

Observation type*	At least one same-device interruption (i.e. P_FOCUS Rs and S_SEQ_SD Rs)	Any kind of multitasking (i.e. P_ALL Rs and S_ALL Rs)
Observed by paradata	37	59
Observed by self-reports	24	63
Observed by paradata, but not by self-reports	23	14
Observed by self-reports, but not by paradata	10	19
Observed by any (paradata *or* self-reports)	47	77
Observed by **both** (paradata *and* self-reports)	13	44
Observed by **neither** (paradata *and* self-reports)	53	23
Agreement rate (sum of the last two rows)	13 + 53 = 66	44 + 23 = 67

To be more specific, 24% of all P_ALL Rs did not report on any RM, as observable in Table 4.13. On the other hand, 30% of S_ALL Rs did not record any focus-out or time-out event. One of the reasons for these differences can be found among concurrent multitaskers. Only 51% of such multitaskers recorded any paradata event. For other SR_RM Rs, this share is 75% and is significantly higher than the overall 70% coverage of all SR_RM Rs by paradata events ($p = 0.02$).

4.2.2 Factors of prevalence of RM

I systematically analyze the relationship between indicators of respondent multitasking and socio-demographic and other respondent characteristics (i.e. individual factors).

The first part is focused solely on the role of age, which is prominently emphasized in the literature (see Chapter 2) as the most important factor

in media multitasking and RM research. The second part includes other factors in the analysis. In addition to descriptive statistics, I follow the example by Ansolabehere and Schaffner (2015) and analyze the effects of all potential factors on prevalence of RM with logit regression models.

Age

Initially, I look at whether age groups differ in terms of clusters of secondary activity forms (as summarized in the previous Section in Figure 4.7). Table 4.14 shows proportions of respondents who are members of a specific cluster for each age group. While some clusters are relatively equally represented across the age groups, respondents from younger age groups more commonly reported on being engaged in at least one secondary activity and are particularly more likely to be members of Complex mixers.

Table 4.15 compares the age distribution for other key RM indicators. It reveals that younger respondents particularly stand out when it comes to media activity forms. The share of SR_RM_MM Rs among the youngest age group is almost twice the size share among the oldest age group. On the other hand, we cannot observe similarly notable trends for non-media activity forms. Still, since media activity forms are generally more common than non-media forms (see Section 4.2.1), they also have greater influence on the overall shares of S_RM Rs per age group. While 59% of respondents aged 15-24 years have reported on being engaged in at least one secondary activity, this is true only for 39% of 45-55 year olds. The youngest respondents also stand out with regards to range of activities (i.e. how many different activity forms has a respondent selected). This is true if we look at the mean range of age groups for media, non-media or all activity forms included in Q1.

Table 4.14: Proportions of members of clusters in the age groups

	15-24	25-34	35-44	45-55	Total
No secondary activities	42	46	58	63	54
Listeners	14	14	13	10	13
Eaters	9	7	6	5	7
F2F talkers	3	4	2	4	3
Watchers	5	3	2	4	3
Phone talkers	2	3	4	3	3
Common mixers	6	4	5	4	5
Complex mixers	19	18	10	7	12
Total	100	100	100	100	100

Table 4.15 also contains aggregate information for indicators based on Q2. Generally, we can observe similar trends, although differences are not so striking. The two youngest age groups are very similar when it comes to interruptions due to secondary activities on the same device (S_SEQ_SD). The youngest age group stands out in prevalence (26%) and mean frequency (0.51) of interruptions due to activities on other devices (S_SEQ_OD). On the other hand, an average 25-34 year old respondent reported on more interruptions (0.62) due to non-media activities (S_SEQ_NM), as well as on more interruptions in total (1.48).

Lastly, Table 4.14 also shows aggregates for paradata indicators. When it comes to P_FOCUS, the two youngest age groups are again exposed more strongly. They stand out in terms of prevalence and mean number of focus-out events per respondent, and both figures notably drop towards the oldest age group. In case of P_TIME and subsequently P_ALL, differences are much less striking.

Table 4.15: RM characteristics of respondents by age groups

		15-24	25-34	35-44	45-55	Total
Q1: Range of activity forms[a]	Media (S_RM_MM)	48 0.75	41 0.52	31 0.39	27 0.32	35 0.46
	Non-media (S_RM_NM)	29 0.43	31 0.38	25 0.32	22 0.25	26 0.33
	All (S_RM)	59 1.18	56 0.90	45 0.71	39 0.57	48 0.78
Q2: Interruptions[bd]	Same device (S_SEQ_SD)	31 0.66	30 0.66	20 0.42	18 0.36	24 0.50
	Other device (S_SEQ_OD)	26 0.51	21 0.43	16 0.29	13 0.23	18 0.34
	Non-media (S_SEQ_NM)	32 0.53	33 0.62	28 0.51	25 0.46	29 0.53
	All (S_SEQ_ALL)	60 1.43	59 1.48	47 1.09	42 0.96	50 1.20
Paradata[cd]	Focus-out (P_FOCUS)	51 1.19	46 1.31	34 0.85	25 0.53	37 0.92
	Time-out (P_TIME)	45 0.75	44 0.65	38 0.61	40 0.63	41 0.65
	All (P_ALL)	69 1.80	63 1.79	56 1.34	52 1.14	59 1.46

[a] Indicators related to how many media, non-media or both secondary activity forms have respondents selected at Q1.
Top of cells: % of respondents who have selected at least one corresponding activity.
Bottom of cells: mean range of corresponding activities.

[b] Indicators related to how many same-device, other-device, non-media or all interruptions have respondents reported at Q2.
Top of cells: % of respondents with at least one interruption of corresponding type.
Bottom of cells: mean of interruptions of corresponding type.

[c] Top of cells: % of respondents who recorded at least one corresponding paradata event.
Bottom of cells: mean of corresponding paradata events.

[d] Ordinal indicators. Values 0-4 reflect the actual frequencies of self-reported interruptions; value 5 refers to all frequencies above 4.

Socio-demographic and individual factors

Models include three socio-demographic factors: age (in years), gender and education (low, middle, and high education; with the latter two entered as dummy variables, with low education as the reference group). The next two variables are related to device ownership: media devices at home (number of different media devices respondent has at home) and smartphone ownership (dichotomous variable reflecting whether respondent owns a smartphone). Everyday media multitasking behavior is included with the MMI measure (see Section 4.1.1).

The rest of factors are related to participants' experiences, as well as their relationship with panel surveys in general and with this particular survey. General respondents' experiences with surveys is expressed with two ratio variables: the number of online panels respondent participating in and the number of surveys the respondents have participated in the last 12 months. Dichotomous variable reflects whether respondent has reported that money, awards or lottery is very important for his participation in commercial panels. Five ordinal variables are from respondents' survey evaluation questions: overall grade of this survey, as well as separate grades on how respondents found the survey to be easy, interesting, too long, and different from other surveys (the higher value the more they agreed with these statements). Lastly, the models include self-reported measures of effort and attention dedicated to the survey (higher value means that respondents reported on giving higher amounts of effort and attention).

Table 4.16: Coefficients of multinomial logistic regressions related to the factors that determine the probability of belonging to specific activity forms clusters

	Lis-teners	Eaters	F2F talkers	Wa-tchers	Phone talkers	Comm. mixers	Comp. mixers
Intercept	-2.35*	-2.89*	-4.87*	-3.47*	-5.26*	-2.86*	-3.28*
Age[a]	-0.15	-0.26*	-0.14	-0.24	0.03	-0.34*	-0.52*
Gender: female[b]	-0.15	0.18	0.69*	-0.43	-0.30	-0.32	0.23
Education: middle[b]	0.60	1.11*	0.69	0.65	2.14*	0.37	0.39
Education: high[b]	0.91*	0.81	0.77	0.17	2.09*	0.21	0.63
Media devices at home[a]	-0.16	-0.06	-0.11	-0.39*	-0.39*	-0.05	-0.17
Smartphone ownership[b]	0.29	-0.14	1.29	-0.36	0.32	0.25	1.08*
MMI[a]	0.36*	0.31*	0.08	0.24	0.61*	0.36*	0.32*
Participation in other panels[a]	-0.09	0.16	0.04	-0.37*	-0.72*	0.00	-0.03
Surveys in last year[a]	0.06	0.12	-0.03	0.39*	0.34*	0.11	0.16
Incentive as motivation[b]	0.13	-0.28	-0.51	1.09*	-0.37	-0.26	0.05
Survey eval.: overall[a]	-0.35*	0.04	0.16	-0.10	-0.12	0.07	-0.14
Survey eval.: different[a]	0.00	-0.30*	-0.09	0.51*	0.16	0.01	0.05
Survey eval.: easy[a]	0.05	-0.06	-0.14	-0.12	-0.25	-0.30	0.00
Survey eval.: interesting[a]	0.20	0.29	0.48	-0.15	0.54	0.61*	0.04
Survey eval.: too long[a]	-0.08	0.29*	0.06	-0.38*	0.18	0.14	0.15
Effort for this survey[a]	0.04	0.01	0.02	-0.12	-0.03	0.17	-0.09
Attention for this survey[a]	-0.65*	-0.37*	-0.82*	-1.04*	-0.87*	-0.99*	-1.00*

* $p < 0.05$

[a] Centered (by subtracting of the mean) and scaled (by dividing by standard deviation)

[b] Dichotomous and/or dummy variable

Table 4.16 presents the results from a multinomial logistic regression where cluster membership is a dependent variable. Respondents who do

not belong to any RM cluster (i.e. non-SR_RM Rs – respondents who did not report on being engaged in secondary activity form in Q1) are the reference group.

Three factors are particularly influential according to this regression model: self-reported attention given to the survey, everyday media multitasking habits and age:

- Lower levels of self-reported attention (i.e. "attention for this survey") increase the probability of being a member of any cluster. This means that respondents with low self-reported attention are more likely engaged in any secondary activity form, with coefficients particularly strong for more intensive secondary activities that denote clusters of Watchers, Phone talkers and both types of Mixers.

- More intensive media multitasking in everyday life (as assessed with MMI) increases the probability of a respondent to be a Listener, Eater, Phone Talker, Common Mixer and Complex Mixers.

- Age is a strong factor for three clusters – younger respondents are more likely to be Eaters or both types of Mixers than to not report on being engaged in any secondary activity form.

All other factors are in a significant relationship with less than three clusters. Overall, we can observe that higher levels of education tend to increase the probability of membership in several clusters. Compared to respondents with primary education, respondents with middle education are especially more likely to belong to Eaters and Phone talkers than to not on report on any form of respondent multitasking (i.e. being a non-SR_RM R). The latter is likewise true for respondents with high level of education, who are also more likely to be Complex mixers and Listeners. Interestingly enough, the remaining socio-demographic factor – gender – is significantly related only to F2F talkers. Compared to males, females are more likely to engage in such activity than to be non-SR_RM R.

Another interesting observation is that respondents who have less different media devices in their household, are more likely to be a Watcher or Phone talker than to be a non-SR_RM R. On the other hand, respondents who personally own a smartphone are more likely to be a Complex mixer.

Factors related to experiences and attitudes towards panel surveys in general and to this particular survey exhibit quite a sporadic behavior in the model. It is worth noting that more frequent participation in surveys in the last 12 months tends to increase the possibility that a respondent will be Watcher or a Phone talker. On the other hand, respondents who are members of more online panels are less likely to be a Watcher or a Phone talker.

Next, I look at the factors related to presence of media (i.e. S_RM_MM) and non-media (S_RM_NM) secondary activity forms as assessed by Q1. Moreover, I also look what factors influence presence of survey interruptions as observed by indicators based on self-reports via Q2 (S_SEQ_SD, S_SEQ_OD, S_SEQ_NM) and paradata (P_FOCUS, P_TIME). Each of these seven indicators is dichotomized (value zero means that respondent did not engage in this specific RM behavior, while value one means that he did engage in this specific behavior at least once). For each dichotomized indicator, a separate logit regression model is fitted, containing the same independent variables as the multinomial logistic model presented in Table 4.16. Results of the seven logit models are in Table 4.17, where each column presents corresponding coefficients of a separate model.

Table 4.17: Coefficients of logistic regressions related to the factors in corresponding models of RM indicators [y]

Factors	S_RM _MM	S_RM _NM	S_SEQ _SD	S_SEQ _OD	S_SEQ _NM	P_ FOCUS	P_ TIME
Intercept	-1.30*	-2.00*	-2.85*	-2.39*	-1.72*	-1.22*	-1.15*
Age[a]	-0.24*	-0.15*	-0.11	-0.23*	-0.14*	-0.38*	-0.09
Gender: female[b]	-0.25*	0.18	-0.06	-0.26	-0.11	-0.07	0.04
Education: middle[b]	0.56*	0.18	0.20	0.08	0.46*	0.20	0.33
Education: high[b]	0.64*	0.16	0.20	-0.02	0.37	0.12	0.15
Media devices at home[a]	-0.19*	0.00	0.06	0.03	-0.01	-0.02	0.02
Smartphone ownership[b]	0.32	0.27	0.70*	0.61	0.09	0.20	0.10
MMI[a]	0.37*	0.14*	0.32*	0.16*	-0.01	0.07	-0.05
Participation in other panels[a]	-0.11	0.01	-0.02	-0.07	-0.01	0.06	-0.06
Surveys in last year[a]	0.13*	0.02	0.07	0.03	-0.03	0.07	0.15*
Motivation for panel[b]	0.24	-0.11	0.07	0.23	0.08	-0.26*	-0.01
Survey eval.: overall[a]	-0.21*	0.00	-0.07	-0.03	0.08	0.12	0.01
Survey eval.: different[a]	0.11	-0.13	0.01	0.12	0.01	0.06	0.01
Survey eval.: easy[a]	0.00	-0.20*	-0.20*	-0.01	-0.20*	-0.04	-0.10
Survey eval.: interesting[a]	0.05	0.43*	0.32*	-0.06	0.20	-0.16	0.16
Survey eval.: too long[a]	-0.06	0.11*	0.18*	0.05	0.09*	0.11*	0.12*
Effort for this survey[a]	0.01	-0.01	0.12	0.07	0.09	0.25*	0.15*
Attention for this survey[a]	-0.73*	-0.41*	-0.49*	-0.36*	-0.36*	-0.20*	-0.22*

[y] Results from seven logit models, one per column. Indicators in the header are dependent variables, dichotomized to reflect whether a value is zero (0) or non-zero (1).

[a] Centered (by subtracting with mean) and scaled (by dividing with standard deviation)

[b] Dichotomous and/or dummy variable

The most consistent and strongest factor is again the self-reported attention. A one-point increase on this five-point scale notably reduces the

odds of a non-zero value for all seven RM indicators, ranging from over 66% for S_RM_MM to about 25% for both paradata indicators. In other words, respondents that are more attentive are substantially less likely to be involved in any type of RM behavior.

Age and everyday media multitasking are again relatively strong factors. Age is a significant predictor in all seven models except for same-device interruptions and time-out events. MMI is significant in all but three models (non-media interruptions and both paradata indicators).

Two out of five statements related to survey evaluation are also prominent factors for different RM behaviors. A one-point increase in agreement (on a five-point scale) with statement "the survey was too long" would thus increase the odds of a respondent having a non-zero value for majority of these seven RM indicators by about 10-15%. Similarly is true also for disagreement with the statement "survey was easy to respond to".

It can be observed that the effects of all other factors are limited only to maximum of two RM indicators. To highlight a few of these effects:

- Females and/or more highly educated respondents are more likely to be engaged in at least one media secondary activity.
- Respondents who were less satisfied with the survey as whole are also more likely to be involved in at least one media activity.
- Self-reported effort (i.e. how much effort respondents gave to this survey) is associated only with both paradata-based RM indicators, particularly P_FOCUS.
- A one-point greater agreement on a 5-point scale increases the chance of at least one focus-out event by about 25%.

4.2.3 Relationship with response quality

After elaborating indicators for the prevalence of respondent multitasking (Section 4.2.1) and their relation to the corresponding factors (Section 4.2.2), I will now discuss the relation between multitasking and response quality. For this purpose, the set of RQ indicators was already presented in Section 4.1.5).

Generally, the relationship of RM indicators with RQ indicators is analyzed with different regression models – i.e. negative binominal (Venables and Ripley 2002), logistic and beta (Cribari-Neto and Zeileis 2009) models – depending on the measurement scale of the related variables. The specific RQ indicators (paradata and self-report) serve in these models as the dependent variable. Independent variables include RM indicators and other factors that could influence the RQ. These factors are theoretically elaborated in Sections 2.3.2 and 2.3.3, while their specific implementation in this empirical study is described in Chapter 3 .

Two streams of response quality indicators will be addressed separately, one related to nonresponse behavior and the other related to satisficing (as outlined already in Section 2.3.2).

Non-response behavior: unit-level item non-response

As noted in Section 4.1.5, unit-level item non-response in this analysis refers to the number of unanswered items per respondent. Over three quarters of respondents (76%) responded to all questionnaire items. As observed from Figure 4.10, this indicator follows a Poisson-like distribution, and only 17 respondents (1.2%) did not answer five or more items. This shows relatively high cooperation, which can be in part explained by respondents being recruited from an online commercial panel, where answers to survey questions are typically mandatory and respondents are therefore conditioned to respond to all items. Although this survey was an exception to this and respondents were even informed

in the introduction that answers are not mandatory, the habit of responding to every question perhaps remained to a certain extent.

Figure 4.10: Frequency distribution of respondents according to the unit-level item non-response

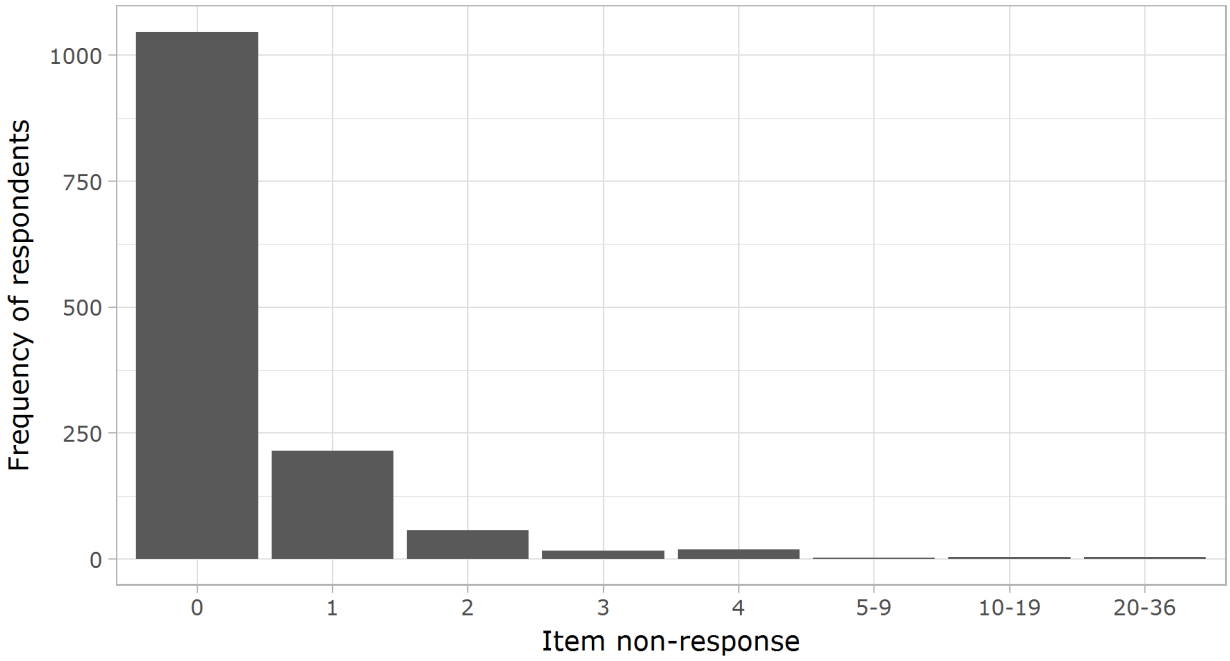


Table 4.18 shows information on negative binomial regression models that contain both paradata indicators as the independent variable and unit-level item non-response as the dependent variable. Second model, containing only P_FOCUS and P_TIME (second column), is significantly better than the model in the first step, i.e. null model (first column) according to the likelihood ratio test. Both paradata indicators are significant; however, effect sizes, pseudo-R^2 and relatively small decrease of the AIC all reveal that the effect of these two indicators is relatively small. It is interesting to note the direction of the effects. P_TIME is positively associated with item non-response, while respondents with more focus-out events are more likely to answer more items.

In the third step, I include the indicator of everyday media multitasking as observed with MMI (Section 2.2.4). More specifically, MMI is entered here in the model as a dichotomous variable indicating whether a respondent is a "high media multitasker" (HMM). In line with Ophir et al. (2009), a

respondent is a HMM if his MMI value is larger than one standard deviation above the average value of MMI of the whole sample. As we can see, HMM respondents are not more likely to have a higher unit-level item non-response. Addition of this independent variable does not improve the model according to the values of log likelihood, pseudo R^2, AIC or the likelihood ratio test.

Next, I added into the model socio-demographic controls, indicators of efforts and attention, as well as indicators expressing overall survey response time. The latter is based on total survey times and is expressed as orthogonal polynomial of second degree. This results in two independent variables (i.e. first-level and second-level coefficient), which are then included into the model.

Overall, these additional controls bring significant improvements according to the likelihood ratio test. However, it should be noted that even the model containing all independent variables included in this analysis (fourth column) still explains a relatively low share of variance of item non-response according to the McFadden's pseudo R^2 (around 4%). Moreover, it is notable that the time-out indicator is not statistically significant at $p=0.05$ level anymore, although it remains below $p=0.1$.

Table 4.18: Negative binomial regression models for unit level item non-response using four hierarchical sets of predictors (including paradata indicators)

	Null model	+ Paradata indicators	+ MMI	+ Controls
Intercept	-0.70***	-0.75***	-0.75***	1.20*
P_FOCUS		-0.12**	-0.11**	-0.14***
P_TIME		0.20**	0.20**	0.13
HMM			-0.04	-0.18
Age				0.00
Gender: female				-0.14
Education: middle				-0.49*
Education: high				-0.32
Effort				-0.03
Attention				-0.35***
Time (first level coef.)				3.07
Time (sec. level coef.				7.36**
2 * log likelihood	-2,369.9	-2,351.9[a]	-2,351.9	-2,267.0[a]
Pseudo R^2		0.0076	0.0076	0.0434
AIC	2,373.9	2,359.9	2,361.9	2,293.0

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

If we look at effects of specific controls, only education is a significant predictor among the socio-demographic factors. By far the strongest effect among all included independent variables can be observed for self-reported attention level, while self-reported effort has no effect. The impact of the time coefficients is positive in both cases, while the second level coefficient is also statistically significant. This indicates that the extremes (i.e. respondents who completed the whole survey either very fast of very slow) are less likely to answer all the items in this survey. We

may recall at this point that the 2% of most extreme units in each directions were already excluded from the very beginning.

I may add that a model with control variables alone (i.e. without the two paradata indicators) has a significantly lower log likelihood than the full model presented in the right-most column in Table 4.18. Pseudo R^2 of the control-only model also notably shrinks, from 0.043 to 0.038.

Next, Table 4.19 presents models that contain RM indicators based on self-reports (instead of paradata indicators as in the previous table). Model in the first column includes only predictors that describe range of media and non-media activity forms. Model reports on a positive relationship between item non-response and the number of different media activity forms respondent is engaged in; no significant relationship is found for range of non-media activity forms. However, as observable from the second column, range of media activity forms becomes non-significant once the model also includes indicators on frequencies due to sequential activities. The only notable effect once all selected reactive RM indicators are included in the model is SR_TS_OD. In other words, respondents who reported on more interruptions due to media activities on other (i.e. non-survey) devices, were more likely to leave more items unanswered. This relationship remains significant with additional controls (third and fourth column). Overall, the effects of these control variables, as well as McFadden's pseudo R-squared and AIC metric are analogous to corresponding paradata-based models.

Table 4.20 presents models with clusters of secondary activity forms as independent variables. As can be seen from the first column, inclusion of these indicators of RM behavior significantly improves the model (compared to the null model with likelihood ratio test). However, only Complex mixers are significantly related with unit-level non-response. Complex mixers responded to fewer items than respondents who did not report on being engaged in any secondary activities (i.e. the reference

group). Nevertheless, even the coefficient related to Complex mixers becomes non-significant once all the controls are introduced into the model.

Table 4.19: Negative binomial regression models for unit-level item non-response using four hierarchical sets of predictors (including self-reports on range of secondary activity forms and interruptions)

	+ Range of activity forms	+ Inter-ruptions	+ MMI	+ Controls
Intercept	-0.82***	-0.87***	-0.85***	0.69
SR_RM_MM	0.28**	0.10	0.12	0.10
SR_RM_NM	-0.09	-0.16	-0.16	-0.19
S_SEQ_SD		0.07	0.07	0.08
S_SEQ_OD		0.29***	0.29***	0.21*
S_SEQ_NM		-0.01	-0.01	0.01
HMM			-0.25	-0.26
Age				0.01
Gender: female				-0.09
Education: middle				-0.39
Education: high				-0.28
Effort				-0.05
Attention				-0.32***
Time (first level coef.)				-1.13
Time (sec. level coef.				4.76
2 * log likelihood	-2,361.3[a]	-2,349.1[a]	-2,345.4	-2,270.8[a]
Pseudo R^2	0.0036	0.0096	0.0105	0.0436
AIC	2,396.3	2,361.1	2,361.4	2,302.8

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

Table 4.20: Negative binomial regression models for unit-level item non-response using three hierarchical sets of predictors (clusters of secondary activity forms)

	+ Clusters	+ MMI	+ Controls
Intercept	-0.79***	-0.78***	1.11
Common mixers	-0.18	-0.17	-0.33
Complex mixers	0.59**	0.62**	0.40
Listeners	0.28	0.29	0.09
F2F talkers	-0.41	-0.41	-0.44
Watchers	0.01	0.01	-0.16
Eaters	-0.19	-0.18	-0.32
Phone talkers	-0.63	-0.61	-1.26*
HMM		-0.16	-0.22
Age			0.01
Gender: female			-0.12
Education: middle			-0.48*
Education: high			-0.40
Effort			-0.06
Attention			-0.34***
Time (first level coef.)			1.00
Time (sec. level coef.)			5.79*
2 * log likelihood	-2,353.7[a]	-2,353.0	-2,268.4[a]
Pseudo R^2	0.0069	0.0072	0.0448
AIC	2,371.7	2,373.0	2,304.4

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

Satisficing

I firstly look at (a) the relation of RM indicators with five dichotomous indicators of satisficing, which include (see Section 4.1.5) estimation as

the response strategy for behavioral frequency question, straight-lining, trap item, inconsistent responses and short open-ended question. Next I investigate the relationship between RM and non-dichotomous indicators of satisficing, namely (b) non-differentiation, and (c) length of responses to open-ended questions.

(a) I firstly report how many respondents have exhibited some form of satisficing according to **five dichotomous indicator of satisficing**:

- 27% of respondents reported that they have used estimation as the response strategy to the behavioral frequency question.
- 23% of respondents had at least one case of inconsistent responding in any of the two investigated matrix questions.
- 16% of respondents failed to properly respond to the trap item.
- 13% of respondents responded to each of the two open-ended questions in less than three words.
- 3% of respondents exhibited straightlining in at least one of the two investigated matrix questions.

In total, 58% of respondents matched at least one of the above criterions of satisficing. One concern here is that the indicators of the behavioral frequency question response strategy and inconsistent responses are too stern, especially because they are also the most commonly violated. If we ignore these two indicators, 26% of respondents have violated at least one of the remaining three dichotomous indicators of satisficing.

Table 4.21: Relationship (coefficients) between RM indicators as dependent and satisficing indicators as independent variables according to logit models

	Estim. behav. freq.	Trap item	Straight -lining	Incon-sistent resp-onses	Short open-ended answer	Any	Any without estim. & incons.
P_FOCUS	-0.09	0.05	0.09	-0.08	-0.15	-0.06	-0.01
P_TIME	-0.11	0.26***	0.36*	-0.03	0.09	0.02	0.23**
SR_RM_MM	-0.07	-0.07	0.01	-0.03	-0.06	-0.05	-0.10
SR_RM_NM	-0.01	-0.02	0.40	-0.03	0.19	0.00	-0.03
S_SEQ_SD	-0.04	0.09	0.27	0.07	-0.04	0.09	0.13
S_SEQ_OD	0.03	0.17	0.23	0.12	0.06	0.09	0.13
S_SEQ_NM	0.01	0.09	-0.11	0.01	-0.44	-0.05	0.07
Common mixers	-0.25	-0.44	0.00	0.21	-1.55	-0.37	-0.58
Complex mixers	0.13	0.22	0.63	-0.05	-0.05	0.10	0.15
Listeners	0.34	-0.60*	-0.21	-0.29	-0.36	-0.37	-0.75**
F2F talkers	0.33	-0.31	0.46	-0.65	-0.74	-0.50	-0.95
Watchers	0.50	-0.20	0.50	0.24	-0.41	0.17	-1.13*
Eaters	-0.04	-0.53	-15.51	-0.11	-15.71	-0.42	-1.18**
Phone talkers	0.34	-1.00	0.38	0.40	-0.93	-0.15	-1.22
HMM	-0.47*	0.87***	1.10**	0.20	0.26	0.22	0.77***

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

In order to investigate relationship between RM and all five dichotomous indicators of satisficing, each of these five indicators is modeled as a dependent variable with a logit regression. More specifically, each of these five indicators is analyzed with four different logit models: one containing the set of paradata RM indicators (the same as model for unit-level item non-response presented in Table 4.18), second containing indicators

based on self-reports on ranges of secondary activity forms (same as in Table 4.19), third containing indicators reflecting the membership of respondents in clusters of secondary activity forms (same as in Table 4.20) and the fourth containing MMI. All control variables are used in these models. Table 4.21 presents the key results from these models – coefficients of RM indicators. Full information on models can be seen in Appendix C.1.

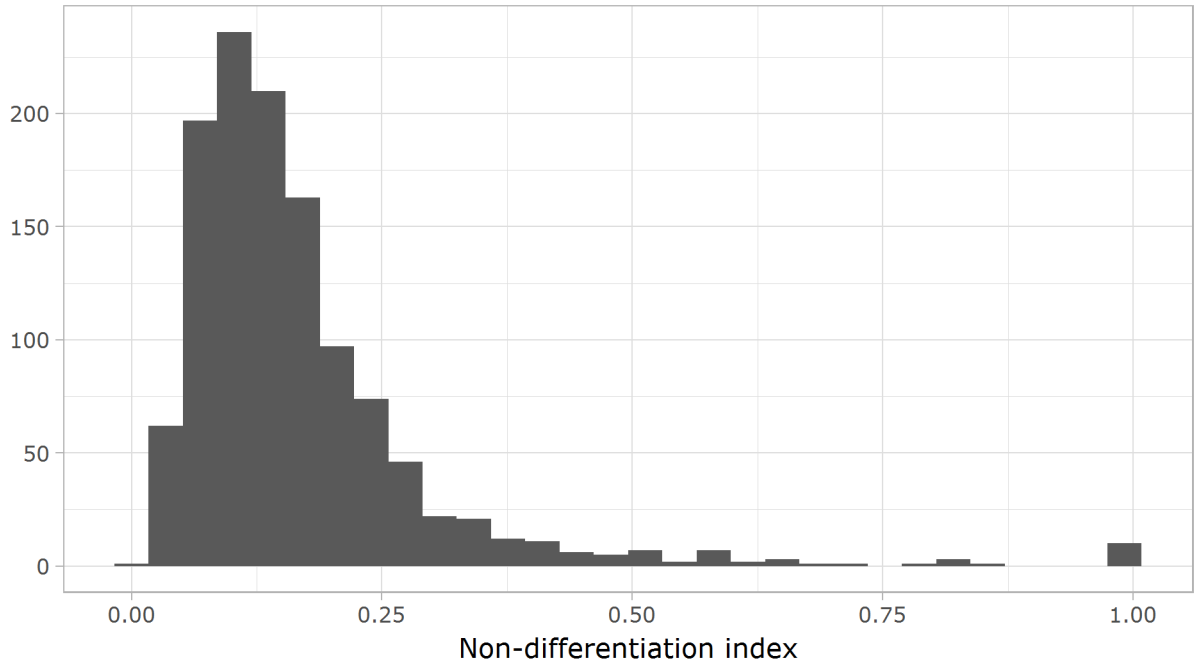
Significant relationship are rare. Among RM indicators, P_TIME stands out. Respondents with more time-out events are more likely to straightline and less likely to provide an adequate response for the trap item. The same can be also said for respondents who are heavy media multitaskers as measured with MMI.

(b) **Composite index of non-differentiation** of responses follows a right-tailed distribution (Figure 4.11). This indicates that majority of respondents did substantially differentiate their responses for both questions.

This indicator is modeled with beta regression in the same hierarchical order as the non-response models. As observable from the second column in Table 4.22, the model containing only paradata indicators of RM are not significantly better from the null model. On the other hand, intensive everyday media multitasking is significantly related to less differentiation, as revealed by the results in the third column.

Interesting patterns emerge once control indicators are introduced to the model. In contrast to previous models on other RQ indicators, all control variables except for gender are strongly related to non-differentiation.

Figure 4.11: Frequency distribution of respondents according to non-differentiation index



Furthermore, once these additional controls are introduced to the model, coefficient of P_TIME becomes significant. Results from the null model thus indicate that respondent with more time-out events were less likely to differentiate their answers in the two matrix questions.

Respondents who are younger, have higher education or the reported on greater levels of self-reported effort and attention tended to differentiate their responses more. Moreover, coefficients related to overall survey time indicate that respondents whose survey completion time was among the fastest or the slowest were less likely to differentiate their responses.

Similar procedures were also done with RM indicators based on self-reports. Tables with full information are in the Appendix C.2. Overall, none of these indicators has a substantial effect on non-differentiation once all the control variables are included in the models. However, there are some indices that respondents who self-reported on more interruptions due to same-device activities were less likely to differentiate their responses.

Table 4.22: Beta regression models for non-differentiation index using three hierarchical sets of predictors (including paradata indicators)

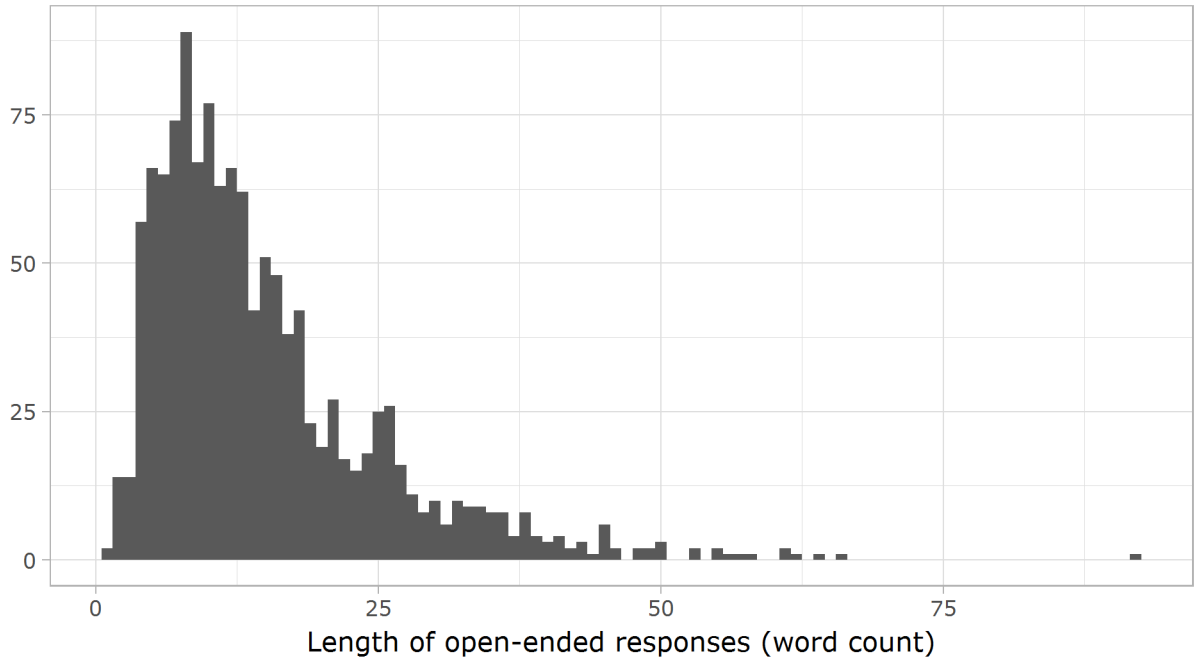
	Null model	+ Paradata indicators	+ MMI	+ Controls
Intercept	-1.36***	-1.36***	-1.40***	-0.43*
P_FOCUS		-0.02	-0.02	0.00
P_TIME		0.03	0.03	0.10***
HMM			0.24***	0.25***
Age				0.01***
Gender: female				0.00
Education: middle				-0.42***
Education: high				-0.43***
Effort				-0.07*
Attention				-0.19***
Time (first level coef.)				-7.07***
Time (sec. level coef.)				5.06***
2 * log likelihood	1529.2	1532.2	1544.0[a]	1737.6[a]
Pseudo R^2		0.0030	0.0138	0.1435
AIC	-1,528.1	-1,532.9	-1,542.7	-1,711.6

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

(c) Next analysis is concerned with the combined **length of responses** to the two **open-ended questions** and only includes respondents who replied to at least one of these two questions (92% of respondents). Mean length of both open-ended responses is 15 words while the median is 12. Distribution of this RQ indicator can be seen in Figure 4.12.

Figure 4.12: Frequency distribution of respondents according to length of open-ended responses



Negative binomial regression is used for modeling combined open-ended response length. Information related to models containing paradata indicators is in Appendix C.3.1. Overall, none of the two paradata-based indicators of RM are significantly related to this RQ indicator once all the controls are included. Information on models containing the range of media activity forms and interruptions are in Appendix C.3.2. Here, the results reveal that respondents who reported on being engaged in a greater number of secondary media activities are more likely to provide longer responses to open-ended questions.

Models that include clusters of secondary activity forms are presented in Table 27. According to the results, none of the clusters is associated with shorter responses compared to non-RM respondents. On the contrary, after the controls are included in the model, results show that Common mixers, Listeners and Eaters tend to provide longer responses.

Table 4.23: Negative binomial regression models for length of open-ended responses using three hierarchical sets of predictors (including cluster activity forms)

	Null	+ Clusters	+ MMI	+ Controls
Intercept	2.71***	2.62***	2.62***	2.20***
Common mixers		0.24**	0.23**	0.20*
Complex mixers		0.14**	0.14*	0.09
Listeners		0.12*	0.12*	0.11*
F2F talkers		0.18	0.18	0.09
Watchers		0.13	0.13	0.15
Eaters		0.29***	0.29***	0.21**
Phone talkers		0.23*	0.22*	0.16
HMM			0.02	0.00
Age				-0.01***
Gender: female				0.23***
Education: middle				0.05
Education: high				0.14*
Effort				0.05*
Attention				0.11***
Time (first level coef.)				3.56***
Time (sec. level coef.)				-2.48***
2 * log likelihood	-8,853.1	-8,823.1[a]	-8,822.9	-8,600.0[a]
Pseudo R^2		0.0034	0.0034	0.0294
AIC	8,857.1	8,841.1	8,842.9	8,636.0

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

4.3 Discussion of results for research questions and hypothesis

4.3.1 Discussion of results for research question 1

I firstly discuss the results for each subquestion separately. Then I summarize and discuss findings for the whole of the corresponding research question.

> **Research subquestion 1a:** How are same-device interruptions based on paradata (i.e. focus-out events) associated with corresponding self-reports?

There is a significant positive relationship between these two indicators. However, there are also important differences. Share (37%) of respondents with at least one focus-out event (i.e. P_FOCUS Rs) is significantly higher than the share (24%) of respondents that reported on having at least one such interruption (i.e. S_TS_SD Rs). These two indicators are in agreement on whether a respondent had at least one same-device activity (13%) or not (53%) in two thirds of the cases.

These discrepancies do not happen only because (a) paradata reveals that there are more same-device interruptions than self-reports, but also because in some cases (b) there are more of these interruptions according to self-reports than according to paradata[24]. The remaining discussion on this research subquestion looks into details into both types of discrepancies, looks at the (c) relationship of focus-out events with other

[24] If we look only at the respondents who have reported on at least one same-device interruption and have at least one focus-out events: 51% of them have more focus-out events than self-reported same-device interruptions; 25% have more self-reported interruptions than focus-out events; and 24% have less the same amount of observations.

RM indicators based on self-reports, and (d) provides a summary on this research subquestion.

(a) Almost two thirds (64%) of P_FOCUS Rs did not report on having any same-device interruption. Here, it is important to keep in mind that these two thirds of P_FOCUS Rs had significantly lower frequencies of focus-out events that the remaining third of P_FOCUS Rs who actually reported on having at least one same-device interruption. Moreover, even when P_FOCUS Rs did report on having such interruptions, the reported frequencies were significantly lower than frequencies of focus-out events.

It is safe to assume that one of the main facilitators for these discrepancies is under-reporting of same-device interruptions by respondents. As evidenced by several studies I have summarized on the overview of non-reactive and reactive measurements (Section 2.2.3), people generally tend to under-report on the frequencies of interruptions and/or switches in the context of media use in everyday life.

In the context of this research, there are several probable reasons for under-reporting. For example, some respondents perhaps failed to account for all (or any) same-device interruptions while responding to a survey. Some subtle empirical evidence for this can be potentially found in the fact that P_FOCUS Rs who did not report on any same-device interruptions had a lower frequency of focus-out events that those P_FOCUS Rs that did report on same-device interruptions. In other words, respondents with fewer same-device interruptions (as observed with paradata) were less likely to report on them than respondents with more same-device interruptions (as observed with paradata). One reason for this could be that they failed to account for some minor instances of same-device interruptions.

It is also possible that some respondents misunderstood the question that prompted them to report on same-device interruptions. For example, instead of reporting on the frequency of same-device interruptions, they

reported on the number of different instances of same-device activities. This could be especially critical for respondents who are engaged in a single instance of secondary activity that interrupts the survey multiple times (e.g., having a continuous instant messaging conversation throughout the survey).

There are of course also other, more general reasons for under-reporting. For example, some respondents might under-report on these interruptions due to privacy issues or social desirability bias (e.g., they could believe that multitasking is undesirable respondent behavior for the authors of this study or the panel managers).

Another source for these discrepancies could be measurement errors and or misclassifications of focus-out events. In Section 2.3.3 I have elaborated on some scenarios that could lead to false positive observations of same-device interruptions via focus-out events such as logging out of operating system account, minimizing the browser without engaging in other same-device activities or various technical errors in the paradata collection procedures. While our pre-testing, pilot study nor the study presented in this dissertation could not find evidences that such false positive cases happen in a considerable amount, it still important to be aware this possibility.

(b) While at least one focus-out event was recorded for a significant majority of S_TS_SD Rs, over two fifths (42%) did not have any focus-out events. However, it is important to keep in mind that these two fifths of respondents had significantly lower self-reported frequencies of same-device interruptions than the other three fifths of S_TS_SD Rs. Moreover, a quarter of S_TS_SD Rs who have at least one focus-out event reported on more same-device interruptions than observed with paradata.

False negatives of focus-out events could be one source of this discrepancy. As elaborated in 2.3.3, false negatives for these events can happen in some specific scenarios where respondent is engaged in a

secondary activity on his survey device without actively switching from the browser tab that contains the web questionnaire (i.e. having a Skype conversation while the program window of this software is minimized). Another source of false negative could be various technical failures of the paradata collection procedures. However, as with false positives, we could find any evidences that such errors are happening at a considerable rate in any stage of our work with paradata.

Another source of this discrepancy is over-reporting by respondents. As already summarized at the discussion for the previous subquestion, empirical literature on media use and media multitasking generally alerts that people tend to under-report on various aspects of their media behavior. However, this same literature also alerts also about over-reports (Section 2.2.3).

In this survey, some respondents could have over-reported simply due to misremembering and/or overestimating the exact number of same-device interruptions. Moreover, over-reports could also be a consequence of misunderstanding of the specific question. For example, some respondents perhaps did not distinguish between same-device and other-device activities. Failing to make a distinction between interruptions and specific instances of same-device activities could also be critical for this type of discrepancies. To give a specific example of this, imagine a respondent who interrupts the survey to do two instances of same-device secondary activity (e.g. checking an e-mail and checking a weather forecast) within this same interruption. Even though this counts for only one interruption, some respondents could have reported on having two.

(c) P_FOCUS Rs tended to report on more different media and non-media activity forms than non-P_FOCUS Rs. Common mixers, Complex mixers, Phone talkers and F2F talkers were all significantly more represented among P_FOCUS Rs. This is not a surprising fact for the first three clusters because notable shares of their members reported on being engaged in

some secondary media activity forms that can be done on a PC (e.g. working on text documents, browsing the web). Over a quarter of members of these clusters also reported on having more than one same-device interruption. It is harder to explain F2F talkers since all members of this clusters reported on being solely engaged in F2F conversations and no other activity forms. Still, 17% of them did report on having more than one same-device activity and this reflects on its association with focus-out events.

On the other hand, a significant majority of Listeners and Watchers did not record any focus-out events, while no significant differences were found for Eaters in either direction. Non-association of focus-out events with these three clusters is understandable. Their characteristics are fixated to secondary activities that are not typically attached to same-device activities.

About a half of respondents who reported on having at least one interruption due to other-device or non-media activities were also P_FOCUS Rs. Moreover, P_FOCUS Rs had significantly higher self-reported frequencies for these two other types of interruptions than non-P_FOCUS Rs. This association is influenced by certain groups of multitasking respondents such as Complex mixers, Common mixers and F2F talker who were commonly reported on being engaged in different types of interruptions.

(d) To summarize, focus-out events are broadly associated with corresponding self-reports on same-device interruptions. There is a significant positive relationship between these two indicators and a significant majority of respondents who reported on same-device interruptions has significantly higher frequency of focus-out events than those respondents who did not report on such multitasking behavior. In general, focus-out events are also meaningfully associated (or not associated) with other indicators based on self-reports. Still, relying solely

on paradata or solely on self-reports would not only give significantly different results on the prevalence and frequency of same-device interruptions, but also on which respondents have even been engaged in such behavior.

Research subquestion 1b: How are time-out events associated with self-reports?

By design, time-out events can potentially detect diverse aspects of respondent multitasking behavior that is observed with a different indicators based on self-reports. In contrast to focus-out events, time-out events do not have a conceptual counterpart among self-reports. Therefore, the discussion for this research subquestion cannot be as rigorous as it was for the previous subquestion.

My discussion here is firstly focused on (a) respondents who have at least one time-out event (i.e. P_TIME Rs) but no self-reports on any form of respondent multitasking. Secondly, I discuss (b) how different indicators of self-reports are associated with time-out events. Lastly, I provide a (c) summary on this research question.

(a) Almost one quarter (23%) of respondents with at least one time-out event (i.e. P_TIME Rs) did not report on any engagement in secondary activities in either Q1 or Q2.

One source for this discrepancy can be attributed to under-reporting. Reasons for under-reporting elaborated at the discussion for previous subquestion also apply here. For example, respondents might fail to account for any instance of their multitasking during the survey or choose to report none at all due to social desirability bias.

On the other hand, I have mentioned that there is a considerable possibility of false positive time-out events. In other words, some respondents' page response times could be marked as time-out events, even though their relatively long response time could happen due to

longer response processes without any engagement in secondary activities. Analysis revealed some limited empirical support for this assumption. These evidences are directly tied to discussion on factors of respondent multitasking. I further elaborate on this under the discussion of the hypothesis (Section 4.3.5).

(b) Overall, half of all respondents who reported on any instance of respondent multitasking had also at least one time-out event. However, P_TIME Rs were significantly more likely to report on media and non-media secondary activity forms. These respondents are also associated with all secondary activity clusters, except for Listeners. Particularly strong associations can be found for F2F talkers (68% of them had at least one time-out event) and Phone talkers (64%).

Furthermore, a significant majority of respondents with at least one other-device (60%) or non-media device (58%) interruption also had at least one time-out event. Overall, P_TIME indicator is significantly correlated with the combined indicator of self-reported other-device and non-media interruptions.

Moreover, P_TIME Rs reported on significantly higher frequencies of all types of interruptions when compared with non-P_TIME Rs, including same-device interruptions. As with the similar association of P_FOCUS with other-device and non-media interruptions, this association is influenced by respondents who commonly reported on being engaged in different types of interruptions (e.g. Common mixers, Complex mixers and Phone talkers).

(c) To summarize, results show that time-out events are particularly strongly associated with other-device and non-media interruptions and a wide array of activity forms (especially face-to-face and audio-/video-mediated conversations). At the same time, one must be aware that relying solely on time-out events would mean ignoring a half of respondents who self-reported that they were engaged in multitasking

during the survey. Moreover, there is also a considerable possibility of false positives among time-out events.

Research question 1: How are paradata indicators on respondent multitasking associated with self-reports?

Both data collection approaches give similar estimates on general prevalence of RM. According to paradata (i.e. focus-out and time-out events), 59% of respondents were multitasking during the survey secondary activities; according to self-reports, this share is 63%. Overall, paradata and self-reports are in agreement on whether a respondent has multitasked during the survey or not in 67% of the cases.

Respondents who were multitasking according to both data sources (who represent 44% of the whole sample) had significantly higher values of both paradata indicators than respondents who multitasked only according to paradata (14%). Likewise, multitasking respondent according to both data sources have reported on significantly more different secondary activity forms and all types of interruptions than multitasking respondents only according to self-reports (19%).

Still, about a quarter of respondents (24%) who multitasked according to paradata did not report on such behavior. The opposite is true for an even greater share – 30% of respondents who reported on any instance of multitasking did not have any paradata events. In the discussion under both research subquestions I have highlighted several reasons for under- and over-reporting in case of self-reports, as well as possibilities of false negatives and false positive for paradata events. In addition to this, two overarching reasons for these discrepancies are:

- As mentioned in Section 2.3.3, quality of responses can decrease towards the end of the questionnaire. It is reasonable to expect that some respondents felt considerably fatigued at the end of this 20-minute survey that contained several cognitively demanding questions. Instead of performing an optimal response process for questions on

respondent multitasking, some respondents might take mental shortcuts and use satisficing strategies. This could lead to both, under- and over-reporting. Satisficing (along with misunderstanding of questions) could also explain why some respondents gave inconsistent responses to Q1 and Q2. E.g., 29% of respondents who reported on having at least one interruption in Q2 did not report on being engaged in any secondary activity form in Q1.

- Focus-out and time-out events together do not conceptually account for all aspects of RM behavior, especially not for concurrent multitasking and short interruptions due to other-device and non-media activities.

Empirical evidences support both reasons for discrepancies between data sources. As already mentioned, 70% of all S_ALL Rs have at least one paradata event. However, if we limit the analysis only on respondents who reported on having at least one interruption, this figure increases to a significantly higher share of 75%. This supports the notion that paradata events are better at detecting non-concurrent multitasking. Furthermore, if we additionally limit the analysis only on respondents who reported being engaged in at least one secondary activity for besides having at least one interruption, the figure increases to 80%. This increase is again statistically significant and indicates that some respondents gave inconsistent responses that over-reported on their respondent multitasking behavior.

Associations between clusters of secondary activity forms and paradata events also show that paradata dominantly covers sequential forms of RM. A significant majority of 90% of all Phone talkers have recorded at least one paradata event. The same is true also for F2F talkers (87%), Common mixers (81%), Complex mixers (74%) and Eaters (73%). In total, 78% of members of these clusters had at least one paradata event. While majority of Watchers (64%) and Listeners (61%) did record at least one

paradata event as well, this proportion is not significantly higher from 50%.

Answer to whether paradata indicators are strongly associated with self-reports or not depends on research goals of RM analysis. If the goal is to measure overall prevalence of different types of RM (e.g. concurrent and sequential activities), the two data collection procedures result in considerably different findings on which respondents were engaged in at least one instance of RM activities. A third of respondents in this study were multitasking only according to one of the data sources and 30% of all respondents who reported on multitasking did not have any paradata events.

If the goal is focused on identifying respondents who were engaged in more intensive and/or sequential types of RM, paradata is more strongly associated with self-reports. This is indicated by the fact that the respondents who multitasked according to both data sources were more intensively engaged in secondary activities than the respondents who were multitasking according to only one of the data sources. Moreover, 78% of respondents who were not members of Listeners and Watchers recorded at least one paradata event. This is true also for 80% of respondents who reported on having at least one interruption and on being engaged in at least one secondary activity form.

Analysis also confirmed that the two paradata indicators are complementary when it comes to association with different types of interruptions. Focus-out events are significantly associated with same-device interruptions, while time-out events are in such association with the other two types of interruptions (i.e. due to other-device and/or non-media activities). However, while these associations are significant, the exact values of paradata indicators and indicators based on self-reported frequencies of interruptions rarely match. This is to be expected in the case of time-out events (who do not observe the exact number of

interruptions due to other-device or non-media interruptions, but only assess whether a delay has happened on an individual questionnaire page after accounting for the length of focus-out events). On the other hand, one might expect a greater accordance between focus-out events and same-device interruptions, but the former were significantly higher than the latter. However, this is not very surprising either, considering that people generally tend to under-estimate the extent of their media multitasking behavior.

To summarize, despite conceptual and methodological limitations of both data collection procedures, results show that paradata in strongly associated with self-reports when it comes to observing whether a respondent was engaged in more intensive types of RM, especially those involving sequential activities. Approaches are less comparable if we try to include all types of RM or observe the exact number of survey interruptions. Published literature, as well as this analysis indicates that this association is strongly influenced by under-reporting. However, we need to be aware that over-reporting, as well as false positives and false negatives among paradata indicators are also all present to some degree.

4.3.2 Discussion of results for research question 2

Similarly to the previous research question, discussion on this research question is done for each subquestion separately. Lastly, I summarize and discuss the findings for the whole research question 2.

Research subquestion 2a: How many respondents have multitasked during the survey?

As mentioned, according to paradata, 59% of respondents have multitasked during the survey. However, as elaborated at previous research question, this figure needs to be taken with some reservations. Firstly, paradata indicators do not account for all types of RM behavior, especially for concurrent secondary activities and other-device and non-media sequential activities that did not considerably delay page response

times. Because of this, some respondents who have multitasked during the survey are not observed as such by using this data collection procedure. Secondly, some focus-out and especially time-out events are most probably false positives. Therefore, some respondents who did not multitask during the survey might be labeled as multitaskers by these paradata indicators.

According to self-reports, 63% of respondents have multitasked during the survey. In contrast to paradata approach, self-reports can observe all types of RM behavior. However, as shown at the previous subquestion, over-reporting and especially under-reporting could influence the overall share of RM.

Due to limitations of both data collection approaches, it is challenging to provide an exact answer to this relatively simple research subquestion. It is more appropriate to provide a conservative and a less conservative estimation. The former is based on self-reports. Under a somewhat idealistic assumption that respondents did not generally under-report on their multitasking during the survey, a conservative estimation is 63%. The least possible conservative estimation is based on the share of respondents who have multitasked according to either paradata or self-reports. In total, 77% of respondents had at least one paradata event and/or reported on at least one type of multitasking during the survey.

Research subquestion 2b: What are the most prevalent secondary activity forms?

Answer to this subquestion depends on self-reports given in Q1. Almost a half of all respondents (48%) reported on being engaged in at least one specific secondary activity form. It is also interesting to note that a quarter of these respondents did not report on interrupting the survey at least once for more than 5 seconds. In other words, 12% of all respondents reported that they were engaged only in concurrent

secondary activities (i.e. they reported on being engaged in at least one secondary activity but did not report on having at least one interruption).

Media-related secondary activities were more commonly reported than non-media activities: 35% of all respondents have reported on being engaged in at least one secondary media activity, while this share is 26% for non-media activities. In total, different media activities were selected 624-times (or 0.45 per respondent), while different non-media activities were selected 448-times (or 0.33 per respondent).

Range of secondary activity forms – i.e. the number of different secondary activity forms reported by a respondent – follows a Poisson-like distribution: 52% of respondents did not report on any secondary activity forms, 30% reported on a single activity form; 12% on two; 4% on three and so on. Only 3% of respondents reported on four or more different secondary activity forms.

Moreover, majority of the 16 secondary activity forms were reported on by less than 5% of respondents. The only exceptions are: "listening to different audio contents" (selected by 22% of all respondents), "eating drinking, or preparing a meal" (14%), "face-to-face conversations" (10%), "texting, instant messaging, or e-mailing" (6%), and "audio/video conversations over phone or other devices" (5%). In relation to this, 6% of respondents also selected at least one of the two open-ended items to report on secondary activity forms that they felt were not included among the close-ended response options in Q1. However, I have managed to recode majority of these additional open-ended responses into secondary activity forms pre-given in the close-ended response options.

Cluster analysis was used in order to get a more concise insight into prevalence of various secondary activity forms. The seven obtained clusters are (in order of their prevalence):

- Listeners (represent 26% of all respondents who have reported on being engaged in at least one secondary form; represent 12% of all

respondents in this sample): respondents who were only listening to audio contents.

- Complex mixers (25%; 12%): a diverse group of respondents where a majority of respondents was engaged in multiple secondary activity forms. No specific activity form dominates in this cluster; however, several relatively cognitively demanding secondary activities such as working on text documents, using social networks and taking care of other people are most strongly represented in this cluster (in comparison to other clusters).

- Eaters (14%; 7%): respondents who were eating, drinking or preparing a meal and were rarely engaged in any other secondary activity forms.

- Common mixers (10%; 5%): respondents who were engaged in multiple secondary activity forms, with at least one of them being either listening to audio contents or having a face-to-face conversation. Another prevalent secondary activity form in this cluster is eating, drinking, or preparing a meal.

- F2F talkers (7%; 3%): respondents who were only engaged into face-to-face conversations.

- Watchers (7%; 3%): respondents who were watching video contents, majority of them also engaged in other activity forms such as listening to audio contents, texting or having face-to-face conversations.

- Phone talkers (6%; 3%): respondents who were having audio- or video conversation via electronic devices. Majority is also engaged in other secondary activity forms, especially face-to-face conversations, but also eating, drinking or preparing a meal, listening to audio contents etc.

Such cluster analysis inevitably leads to a certain simplification's and loss of information; its disadvantages are further discussed under discussion of

the hypothesis (Section 4.3.5) and limitations of the research (Section 5.2). On the other hand, this analysis also reveals some additional insight in comparison on reporting only on prevalence of individual activity forms. Specifically, it allows us to see what combinations of secondary activity forms certain segments of respondents commonly reported.

Research subquestion 2c: How often do respondents interrupt the survey due to same-device, other-device or non-media activities?

According to self-reports, 50% of respondents interrupted the survey because of secondary activities at least once for 5 seconds. Most commonly, this was due to non-media activities (29% of all respondents reported on at least one such interruption), followed by same-device interruptions (24%) and other-device interruptions (18%).

Same-device interruptions were also observed with focus-out events. According to the paradata indicator, 37% of respondents had at least one focus-out interruption. As elaborated already elaborated in the discussion for the previous research question, discrepancy between paradata and self-reports on the measurements of same-device interruptions are likely to be strongly influenced by under-reporting. It is reasonable to assume that similar underestimations are influencing self-reports on other-device and non-media interruptions as well.

What is common to all interruption indicators based on either self-repots or paradata is that frequency of interruptions per respondent follows a similar Poisson-like distribution. For example, 20% of respondents reported on interrupting the survey once, followed by 12% of respondents with two interruptions, 6% with three interruptions, 4% with four interruptions, while 8% reported on having five interruptions or more.

In contrast to self-reports, focus-out events also provide an exact number of observations of same-device interruptions. Overall, 1,595 focus-out events were detected in the context of this analysis. In other words, an average respondent had 1.2 same-device interruptions during the survey.

Research question 2: What is the prevalence of respondent multitasking in the web survey?

Based on indicators from both data collection procedures, I estimate that somewhere between 63% and 77% of respondents have multitasked at least once during the 20-minute survey used in this study. Such estimation is considerably higher from the other two published studies on RM among web panelists that were based on self-reports. Ansolabehere and Schaffner (2015) reported that 50% of respondents in about 30-minute long surveys have multitasked at least once, while in Antoun (2015) this was true for 44% of all respondents during a 10-minute survey. Of course, such direct comparisons between studies must be taken with reservations due to numerous methodological differences between them. One particular reason are different approaches for measuring RM. For example, in Ansolabehere and Schaffner (2015), respondents could report on 10 different secondary activity forms in a closed-ended multiple choice question without open-ended items. In contrast, in these study respondents could select among 16 different secondary activity forms, as well as an open-ended item. Moreover, an additional question specifically prompted respondents to report on any interruptions. Questions about RM in this study were thus more extensive compared to Ansolabehere and Schaffner (2015) and this could be one of the reasons on why more respondents reported on RM.

It is also important to note that much less of respondents were engaged in more than one secondary activity form (18%), had more than one self-reported interruption (30%), or more than one paradata event (36%). Overall, all indicators on ranges of secondary activity forms and frequencies of interruptions per respondent follow a Poisson-like distribution which reveals that the share of respondents who were intensively engaged in RM behavior is relatively small. This is also supported by the fact that two most commonly reported secondary activity forms (listening to audio contents and eating, drinking, or

preparing a meal) are not typically particularly disruptive. Over a half of all respondents who reported on being engaged in secondary activity forms were engaged solely only in one of these two activity forms. In other words, only a quarter of all respondents reported on being engaged in at least one of the other secondary activity forms. Among them, most common are activities related to interpersonal communication: face-to-face conversations, text-based conversations or audio/video conversation via electronic devices. At least one of these activity forms was reported on by 21% of all respondents. None of the remaining 11 activity forms that were included as response items in Q1 was reported by more than 5% of all respondents. In fact, only 18% of respondents reported on at least one of these 11 secondary activity forms (including watching video contents, browsing the web, using social networks, doing household chores or taking care of other people).

4.3.3 Discussion of results for research question 3

As with previous two research questions, I first elaborate on specific subquestions and then for the whole research question 3.

> **Research question 3a:** What socio-demographic factors influence RM?

Younger respondents were more likely to be engaged in secondary activities than older respondents. While about 60% of 15-24 year olds reported on being engaged in at least one secondary activity, this holds true only for 40% of 45-55 year olds. Similar patterns can be also observed with paradata indicators – almost 70% of 15-24 year olds recorded at least one paradata event, while this is true for about 50% of 45-55 year olds. More detailed analysis reveals that younger respondents are more likely to report on being engaged in more than one specific form of secondary activities. They are also more likely to be engaged in secondary media activities. This holds true even after we control for other socio-demographic and individual factors.

Gender and education are less strongly related with different forms of RM. Still, females and/or more highly educated respondents were somehow more likely to be engaged in secondary media activities. Females are also more likely to be F2F talkers while highly educated respondents are more likely to be Phone talkers.

Overall, results reaffirm findings from Zwarun and Hall (2014) and especially Ansolabehere and Schaffner (2015) on the importance that age has in explaining prevalence of RM in overall terms as well as for some specific subtypes of RM.

Research question 3b: What other individual factors influence RM?

Self-reported attention was by far the most important factor of RM in this study. Respondents who reported on being less attentive during the survey were more likely to be involved in all types and subtypes of RM behavior as measured with paradata and self-reports. If we understand self-reported attention as an indicator of respondents' engagement, results here confirm that engagement and RM are in a strong relationship. However, it is interesting to note that the other approximate indicator of respondents' engagement – self-reported effort – was not an important factor of RM. This indicates that a further research on relationship between RM and engagement is warranted.

Results also give solid support for the proposition that everyday media multitasking habits are related to RM in web surveys. Respondents with higher index of everyday media multitasking behavior are more likely to be Listeners, Phone Talkers, Common mixers and Complex mixers. They are also more likely to be particularly engaged in media activities and related same-device and other-device interruptions.

While some sporadic effects can be also found for other individual factors (such as survey evaluation), overall none of them proved as important as other factors described in the above two paragraphs.

Research question 3: What are socio-demographic and individual factors related to RM?

Overall, findings highlight that age, self-reported effort and everyday media multitasking habits are importantly related to RM. These factors are also important in relation to research question 4 and are thus further discussed.

4.3.4 Discussion of results for research question 4

Similar to previous research questions, I firstly discuss specific subquestions related to research question 4. Then, I provide a summary for the whole research question 4.

Research question 4a: What is the relationship of RM indicators (based on paradata and self-reports) and unit-level item non-response?

Overall, only one of the RM indicators is a significant predictor of *higher* unit-level non-response once all controls are included in the regression models. Namely, respondents who reported on more other-device interruptions tended to respond to less items. On the other side, two RM indicators are significant predictors of *lower* unit-level non-response: focus-out events and Phone talkers. In other words, respondents who are Phone talkers and/or have more focus-out events also tended to respond to more items.

In addition to RM indicators, other factors may also impact unit-level item nonresponse. However, according to these models, the only consistently important factor is self-reported attention – respondents who reported on giving less attention to the survey also tended to respond to fewer items. On the other hand, high media multitaskers did not have a higher rate of unit-level non-response.

We may add that in models without control variables two other RM indicators were also significant. Namely, in these models, respondents

who are Complex mixers and/or have more time-out events tended to respond to fewer items.

> **Research question 4b:** What is the relationship between RM indicators (based on paradata and self-reports) and satisficing indicators?

Overall, time-out events had the strongest relationship with satisficing among included RM indicators. Respondents with more time-out events were less likely to respond properly to the trap item and they less likely differentiated their responses in the matrix questions. No other RM indicator was consistently related with at least two of the five different indicators of satisficing once all the control variables were included in the models. It is worth noting that several RM indicators based on self-reports (e.g. range of media activities, Common mixers, Listeners and Eaters) were associated with longer responses to the open-ended questions.

It is notable that indicator identifying high media multitaskers is a significant predictor for a wider variety of satisficing indicators than any RM indicator. Respondents who are intensively media multitasking in everyday lives were more likely to fail to properly respond to the trap item and differentiated their responses in the matrix questions less than other respondents. Interesting enough, high media multitaskers also were less likely to use estimation as the response strategy to the behavioral frequency question.

Self-reported attention was one of the most important factors also in relation to satisficing indicators. Lower levels of self-reported attention were related to higher probability for a proper response to the trap item, lower differentiation of responses in the matrix question and shorter responses to the open-ended questions. Similar relation to these three indicators is also found for self-reported effort and overall survey completion time. To expand on the latter, respondents who were

responding to the survey relatively quickly or slowly were more likely to exhibit these three specific forms of satisficing.

Research question 4: What is relationship between RM indicators (based on paradata and self-reports) and RQ indicators?

Findings of this dissertation are in line with findings from existing empirical literature even though this analysis uses several new indicators of RM. Overall, RM indicators are rarely in significant relationship with RQ indicators. Even when they are, their effects are small and/or coefficients become non-significant once the control variables are included in the models. Time-out events are the only indicator that predicts a lower RQ for more than one specific RQ indicator (trap item and non-differentiation).

It is interesting to see that focus-out events were not a significant predictor of lower RQ in any of the models containing all control variables. This is different from the study by Sendelbah et al. (2016) where focus-out events were a significant predictor of higher unit-level non-response. These contrasting findings probably happened because the two studies differ considerably in terms of methodology (e.g. the target population in Sendelbah et al. (2016) were students, while in this dissertation the target population is much more general in terms of age and education).

Overall, all three types of interruptions – i.e. interruptions due to same-device (as observed with focus-out events and self-reports), other-device and non-media (both observed with self-reports) activities were not related to worse RQ. The only exception here is that a greater number of self-reported other-device interruptions was a significant predictor of greater unit-level non-response.

Absence of a stronger relationship between interruptions and RQ is somewhat surprising considering the theoretical elaboration given throughout Chapter 2 . This empirical finding can be attributed to two major reasons. Firstly, results could be importantly influenced by

measurement errors (mainly false positive and false negatives in case of focus-out event and underreporting in case of self-reports). Secondly, while a relatively high share of respondents did interrupt the survey at least once (37% of respondents had at least one focus-out event; 50% of respondents reported on having at least one interruption), it is possible that a considerable proportion of these interruptions were done in a controlled manner (e.g. respondents interrupted the survey in a moment when the mental workload related to response process was low). As shown in Section 2.2.4, such manner of interruptions is less intruding to the quality of primary activity.

It is also interesting to note that respondents who belong to clusters that define more intensive RM behavior (e.g. Phone talkers, Common mixers and Complex mixers) are also not strongly related to worse RQ. This indicates that the used clustering procedure was non-sufficient for this specific research goal. Perhaps a clustering based on secondary activity forms and frequency of interruptions would yield clusters that are more meaningfully related to RQ.

In conclusion to discussion on this research question, it is important to highlight the importance of control variables. Intensive everyday media multitasking and self-reported effort are not only important factors of RM but also have a substantial role in the relationship with RQ. This indicates that RM, engagement, everyday media multitasking and RQ are in a complex interdependent relationship (as theoretically elaborated in Section 2.3.2). Moreover, overall survey response time was also proven as an important element of relationship with RQ. Results confirm findings from more general survey methodology literature that respondents who complete the survey very fast of very slow can exhibit lower RQ.

4.3.5 Discussion of results for hypothesis

As elaborated in Chapter 3 , confirmation of hypothesis depends on three streams of arguments or sub-theses. Therefore, I firstly look at each individual sub-hypothesis.

Sub-hypothesis 1: Paradata-based indicators can identify majority of respondents that are multitasking according to self-reports.

As already elaborated in discussion on research question 1 (Section 4.3.1), 70% of respondents who reported on being engaged in any type of RM also recorded at least one focus out and/or time-out event. More detailed analysis also shows that the remaining 30% of respondents (i.e. those respondents who reported on being engaged in RM but did not have any related paradata events) were significantly less intensively engaged in secondary activities (e.g. they reported on significantly lower ranges of secondary activities and frequency of interruptions). In other words, paradata approach correctly identified the majority of respondents who were intensively engaged in any secondary activities.

This is further confirmed by the fact that members of majority of clusters of secondary activity forms are significantly more likely to have at least one focus-out and/or time-out event. The only two exceptions here are Listeners and Watchers – both of these clusters denote a RM behavior that is less disruptive than RM behavior related to other clusters.

In relation to this, it is important to note that 24% of respondents who were multitasking according to paradata were not identified as such by self-reports. To stress again, this discrepancy cannot be attributed only to underreporting – false positive and false negatives of focus-out and particularly time-out events could also be an important reason for this.

Overall, 44% of respondents were identified as multitaskers by both data collection approaches. Paradata approach also identified additional 14% of respondents as multitaskers who did not report on being engaged in any

type of RM. The two approaches did not agree on whether a respondent has multitasked or not in 33% of cases. This is a notable share if we are interested on which respondent has multitasked on a case-by-case basis. However, looking at a larger picture, I conclude that the paradata approach has identified majority (70%) of multitaskers according to self-reports and was particularly successful at identifying those respondents who reported on more intensive types of RM behavior. Therefore, I confirm this sub-hypothesis.

Sub-hypothesis 2: Respondents who are multitasking according to paradata approach have similar relations with factors of RM as respondents who are multitasking according to self-reports.

Several importance differences can be seen in how indicators from the two data collection approaches are related to factors. For example, education, gender, media multitasking index and indicators related to survey evaluation are not related to focus-out or time-out events, while they are related to some of the RM indicators based on self-reports. On the other hand, self-reported effort is only related to the two paradata indicators.

An illustrative example of these differences can be observed by looking at focus-out events (which has five significant factors) and same-device interruptions (which has six significant factors). Even though these two RM indicators conceptually target the same type of RM behavior, only two factors (self-reported attention and agreeing with statement "This survey was too long") are significant for both of them.

One particularly concerning pattern here is connected to time-out events. All other RM indicators are significantly related to either age or everyday media multitasking or both. However, this is not true for time-out events. This suggests that certain time-out events attributed to older respondents and/or respondents with low index of everyday media multitasking are false positives. This could be particularly true for older respondents who tend to have longer response times in general (as noted in Section 2.3.3).

It should be noted that self-reported attention is a significant factor for all RM indicators. Moreover, age is a significant factor for focus-out events and majority of indicators based on self-reports. However, despite these similarities, this sub-hypothesis cannot be confirmed due to important differences elaborated in the previous three paragraphs.

Sub-hypothesis 3: Respondents who are multitasking according to paradata approaches have similar relations with RQ as respondents who are multitasking according to self-reports.

Broadly speaking, RM indicators from both data collection approaches all have weak to non-existent relationship with RQ. Within this context, the regression models containing sets of RM indicators based on paradata and regression models containing sets of RM indicators based on self-reports also have comparable evaluation metrics. For example, regression model with unit-level non-response as the dependent variable and paradata indicators and control variables as independent variables has a pseudo R^2 of 0.0434 and AIC value of 2,293. Pseudo R^2 for a similar model that contains indicators based on self-reported ranges of activities and interruptions instead of paradata indicators is 0.0436, while AIC is 2,303. For a model containing indicators based on clusters of secondary activity forms, pseudo R^2 is 0.0448 and AIC is 2,304.

Among all RM indicators, time-out events are in the strongest relationship with worse RQ. Respondents with more time-out events were more likely to fail at the trap item, differentiate less with their responses to the matrix questions and to some degree also have a greater unit-level item non-response. No other RM indicators (based on either paradata or self-reports) is in such relationship with more than one RQ indicator. However, to stress again, effects of time-out events on RQ are also considerably small.

Overall, I conclude that RM indicators based on either paradata or self-reports do provide generally comparable findings on the relationship with RQ. Therefore, I confirm this sub-hypothesis.

> **Hypothesis:** Paradata-based procedures (focus-out and time-out events) can measure respondent multitasking in web surveys and they provide similar general findings on prevalence of multitasking behavior, its factors and relationship with response quality, as the approaches based on self-reports.

Let me first recap the discussion on the three sub-theses. The first sub-hypothesis is confirmed. Despite several important discrepancies, paradata indicators do identify majority of respondents who reported on multitasking during the survey. The second sub-hypothesis is not confirmed. While a couple of factors are in relationship with RM indicators from both data collection approaches, this is not true for majority of factors. The third sub-hypothesis is confirmed. Indicators from both data collection approaches are in a similar, very weak relationship with RQ.

The main hypothesis is only partially confirmed with regards to the above discussion on the three sub-theses. In other words, it is not possible to claim that this paradata approach by itself can substitute or replace self-reports, even if the main research goals are as broadly defined as they are in the three sub-theses.

In theoretical framework I have highlighted two limitations of the paradata approach. Firstly, paradata approaches is not equally viable to measure all different aspects of RM behavior. Secondly, there is a possibility of false positives, especially with time-out events. While I was aware of these two limitations, their extent was greater than I expected. Based on results, it possible to argue that both of these limitations were one of the main reasons not only for rejection of the second sub-hypothesis, but also for important discrepancies related to the first sub-hypothesis.

Rejection of the main hypothesis does not mean to imply that paradata is not a viable approach for RM research. As long as researchers account for its limitations, this dissertation has shown that paradata approach can generate a valuable insight on prevalence of RM and its relationship with RQ that is generally comparable to self-reports. Specifically, paradata approach has identified majority of respondents who were engaged in more intensive forms of RM. Identification of such respondents can be considered to be one the main general research challenges in overall research on RM. This is because more intensive RM behavior respondents is particularly more likely to be related with RQ based on the theoretical background given in Section 2.3.2. If we additionally consider the limitations of approach based on self-reports (e.g. issues with under-reporting and increasing the respondent burden), paradata approach represents a worthwhile alternative that warrants further development and research.

5 Conclusions

Media multitasking is an increasingly common everyday behavior that can be related to worse quality of performance in goal-oriented cognitive activities. As such, media multitasking presents two broad challenges for social research methodology. The first challenge is how to conduct scientific research on various forms of multitasking. The second challenge is how multitasking of participants influences the quality of data obtained from them in empirical social science research. This dissertation mainly addressed the latter challenge within the context of web surveys, an increasingly important survey mode in social sciences.

With only four published empirical studies on respondents multitasking in web surveys, this topic is poorly researched. Two general challenges are related to this and this dissertation addressed both of them. Firstly, there is a lack of discussion on how to measure respondent multitasking in web surveys. Secondly, more empirical evidence is needed on the relationship between response quality and respondent multitasking.

Following the theoretical framework and an overview of relevant empirical findings from literature on media multitasking and respondent multitasking (as presented in Chapter 2), I have identified four research questions with corresponding subquestions and the main hypothesis in Chapter 3 . In Chapter 4 I presented the methodology, results and discussion of the empirical study that was conducted on a web survey.

Below I first summarize the main findings (Section 5.1), then I discuss the limitations of this dissertation (Section 5.2) and directions for further research (Section 5.3). Lastly, I highlight the original contributions of the dissertation (Section 5.4).

5.1 Overview of main findings

In this Section, I firstly summarize the key concepts elaborated in the theoretical framework and the research objectives of this dissertation (Section 5.1.1). Secondly, I sum up the development of new approaches for measuring respondent multitasking (Section 5.1.2). Thirdly, I present the key empirical findings on respondent multitasking of this dissertation with regards to the research objectives and the more general context of media multitasking (Section 5.1.3).

5.1.1 Summary of concepts of the theoretical framework and research objectives

In order to have a holistic understanding of respondent multitasking in web surveys, the theoretical framework draws on scientific literature from multiple disciplines and is built upon the understanding that multitasking is a hypernym of media multitasking and that media multitasking is a hypernym of respondent multitasking online. Correspondingly, the theoretical framework integrates three streams of research (i.e. research on multitasking, research on media multitasking and research on respondent multitasking).

As demonstrated in Sections 2.1 and 2.2.1, conceptualizations of multitasking (e.g. König and Waller 2010; Salvucci and Taatgen 2011; Circella et al. 2012) and media multitasking (Foehr 2006; Wallis 2010; Tokan et al. 2011) substantially differ between and even within different scientific disciplines that research these phenomena. The first important step in developing the theoretical framework was thus expounding on the conceptualization of (a) multitasking and (b) media multitasking. Based on this, I have elaborated on (c) the definition and taxonomy of respondent multitasking. Another important component of the theoretical framework is an (d) overview of reactive and non-reactive approaches. After using the theoretical framework to identify the key gaps in research on

respondent multitasking, I have developed (e) four research questions and one hypothesis. The summarized theoretical elaboration was then a basis for development of the specific methodological approach used in this dissertation (briefly presented in the next section).

(a) My broad understanding of multitasking follows the work of Salvucci and Taatgen (2011). Their unified theory of multitasking describes multitasking and related research with three continua.

The multitasking continuum describes multitasking as a spectrum where one extreme is concurrent multitasking (e.g. activities are combined simultaneously where attention is divided to both activities at once) and the other is sequential multitasking (e.g. two activities are combined in a longer sequence where attention shifts from one activity to another). In other words, sequential multitasking is defined by the task or activity switch – an act of shifting attention from one activity to another.

The abstraction continuum describes the time scale of research – multitasking can be conceptualized, measured and analyzed on a biological band (neural and physiological processes, measured at a sub-second level), cognitive band (specific actions and unit tasks, measured in seconds), rational band (tasks, measured in minutes or hours) and/or social band (long-term behavior, measured in days, weeks and beyond).

The application continuum differentiates research on multitasking with regards on how close the empirical study is to an everyday task scenario, with less-applies studies such as laboratory experiments on one side of the continuum and more-applied studies such as observational studies on the other side of the continuum.

(b) For the purposes of this dissertation, my broad understanding of media multitasking extends on Wallis (2010). Shortly put, media multitasking is any multitasking that involves at least one media activity. Media activities refer to consumption of media contents, creation of media contents, and communication via electronic devices.

Such a broad definition allows for an overview of a wide range of diverse literature, whereas studies are typically concerned with narrower, more specific definitions of media multitasking. In relation to this, it is useful to differentiate three basic types of media multitasking: external (media multitasking that involves at least one external activity), internal (media multitasking that involves media activities on two different devices), and single-device (media multitasking that involves only activities on a single device).

(c) Conceptualization of respondent multitasking is built upon conceptualizations of multitasking and media multitasking. Respondent multitasking is defined as any type of multitasking where one of the activities is related to responding to a survey. Since responding to a web survey involves interaction with media devices and is thus a media activity, respondent multitasking in web surveys is a specific manifestation of media multitasking. In the context of respondent multitasking in web surveys, responding to a web survey is regarded as the primary activity, while all other activities are regarded as secondary activities.

Moreover, I expand on the taxonomy of secondary activities in web surveys by Zwarun and Hall (2014) by integrating key concepts from literature on general multitasking and media multitasking. Secondary activities are firstly differentiated regarding whether they are combined with the primary activity concurrently or sequentially. Sequential activities are further separated into same-device activities (i.e. secondary activities occur on the device that is used for responding to a web survey), other-device activities (i.e. secondary activities occur on media devices that are not used for responding to a web survey) and non-media activities (i.e. secondary activities that are not related to the media).

Two important concepts related to the taxonomy of secondary activities are also instances and forms. An instance of secondary activity refers to a single standalone secondary activity. For example, if a respondent visits a

website to check a weather forecast while he is responding to a web survey, this is regarded as one instance. If he visits another website to check the stock market, this is another instance. A form of secondary activity can be described as a group of similar instances, as defined and observed by researchers. The two instances given in the previous example (weather forecast, stock market) belong to one secondary activity form (visiting a website).

Finally, an important dimension of respondent multitasking behavior are interruptions. An interruption happens when a primary activity (i.e. responding to a web survey) is suspended due to a switch to a secondary activity. It is important to keep in mind that one interruption can contain multiple instances and/or forms of secondary activity and that a single instance and/or form of secondary activity can be related to multiple interruptions.

(d) Based on the literature overview in Section 2.2.3, data collection approaches in media multitasking literature can be divided into reactive and non-reactive approaches. Reactive approaches include all methods in which active participation of research subjects is required. Non-reactive approaches include methods in which no active participation of research subjects is required. In the context of this dissertation, reactive approaches are related to self-reports, while in case of non-reactive approaches I mainly talk about electronic tracks that automatically log the usage of media devices.

Both data collection approaches have specific advantages and disadvantages. As shown by the overview of empirical literature, the main issue with self-reports is that people tend to underreport on their media multitasking behavior. On the other hand, the implementation of reactive approaches using electronic tracks can present a considerable technical challenge. On account of this, such reactive approaches in media multitasking literature are currently mostly limited to measurements of

single-device media multitasking in small-case studies. However, one must take into consideration that the evolution of these reactive methods is still in its early stages. Consequently, several authors (Greenberg et al. 2005; Wallis 2010; Möller et al. 2013) have stated that non-reactive approaches should be more prominently used in multitasking and/or media use research.

Similar circumstances can be observed in literature on respondent multitasking in which reactive approaches are also more common than non-reactive. Reactive approaches for measuring respondent multitasking are typically implemented by including additional questions at the end of the questionnaire. These questions prompt respondents to report on their multitasking behavior during the survey. As indicated by an overview of literature, the design of questions on respondent multitasking can notably influence the research outcome. Question design is particularly important in light of findings of media multitasking literature that people tend to underreport on their multitasking behavior.

In web surveys, the non-reactive approach relates to collection of paradata (i.e. a log of electronic tracks about the process of surveying, automatically generated by the respondent's interaction with the web questionnaire). Paradata is a relatively new data source in web survey methodology and several authors have called for more research utilizing paradata to get insights into different aspects of respondent behavior (e.g. Lynn and Nicolaas 2010; West 2011; Callegaro et al. 2015).

An overview of literature on web survey paradata (Section 2.3.3) identified existing and potential uses of paradata in research on respondent multitasking. Generally, there are two categories of paradata approaches. The first category of approaches – time-out events – is based on the analysis of response times, i.e. the time respondents need to navigate through the question, page or the whole survey. Here, researchers identify response times that are considerably longer than the

typical data from observations. Response times that are longer than the predefined threshold (i.e. cut-off point) are considered to be a consequence of respondent multitasking. The second category of approaches is based on paradata events that indicate the respondent has been temporary inactive during the survey. One particular type of a paradata event that belongs to this category is a focus-out event. A focus-out events contain information on when respondents lost and regained focus on the browser window or tab that contained the web questionnaire. This change of focus happens when a respondent switches to another window or tab.

One general limitation of the overviewed paradata approaches for measuring respondent multitasking in web surveys is the lack of qualitative information on what forms of secondary activities respondents have been engaged in. Moreover, these paradata approaches cannot observe all types of secondary activities equally well (e.g., concurrent sequential activities). Related to this, there is a considerable possibility of false positive and false negative detections of respondent multitasking.

(e) The research questions and the hypothesis were fully elaborated in Chapter 3 . For the purposes of this summary, I list the main four research questions and hypothesis bellow. The key findings related to each of these research objectives are presented in Section 5.1.3.

Research question 1: How are the specific paradata observations (focus-out and time-out events) of RM associated with self-reports on multitasking?

Research question 2: What is the prevalence of different types of respondent multitasking behavior in the web survey, according to specific RM indicators based on paradata and self-reports?

Research question 3: What are the socio-demographic and individual factors related to RM?

Research question 4: What is the relationship between RM indicators (based on paradata and self-reports) and RQ indicators?

Hypothesis: Paradata-based procedures (focus-out and time-out events) can measure respondent multitasking in web surveys and they provide similar general findings on the prevalence of multitasking behavior, its factors and its relationship with response quality, as the approaches based on self-reports.

5.1.2 Description of non-reactive and reactive approaches used in this dissertation

The research objectives of this dissertation required the development of reactive and non-reactive data collection approaches for measuring respondent multitasking in web surveys.

The development of data collection approaches was guided by specific research goals as well as more general insights from the theoretical framework. In this Section, I summarize the approaches I developed and used based on (a) paradata (full description can be found in Section 4.1.2) and (b) self-reports (Section 4.1.3).

(a) The paradata approach was based on two types of paradata events that were already discussed in Section 5.1.1: focus-out events and time-out events. Implementation of focus-out events was relatively straightforward – the only hindrance was abundance of very short focus-out events that happened mostly due to technical reasons. Since such events did not occur due to respondents' behavior, I ignored all events that were shorter than five seconds. Conceptually, focus-out events relate to interruptions due to same-device secondary activities.

Time-out events were based on page response times (after deducting lengths of focus-out events). A novel strategy was used for determining the cut-off point. Respondents' page response times were modeled as a dependent variable with a linear mixed model with crossed random effects

for respondent and page factors in order to account for differences due to the structural elements of specific pages (e.g. word count, frequency and complexity of questions, etc.) and individual differences between respondents (e.g. general cognitive abilities, familiarity with participation in web surveys etc.). A time-out for a specific respondent on a specific page happened if the value of the corresponding model residual was greater than one. Time-out events do not relate to a single specific type of respondent multitasking behavior. Instead, time-out events can conceptually observe any instances of respondent multitasking that prolong the page response time past the cut-off point.

(b) Two questions were developed in order to measure respondent multitasking via self-reports and were included at the end of the questionnaire. The first question prompted respondents to report on what forms of secondary activities (if any) they were engaged in during the survey. An extensive list of media and non-media secondary activity forms that was included in the question as available response options was partly based on one of the most prominent questionnaires in literature on media multitasking (Media Multitasking Questionnaire by Ophir et al. (2009)).

The second question prompted respondents to report on the frequency of interruptions. Following the taxonomy of respondent multitasking, respondents specifically reported on frequencies of interruptions due to same-device, other-device and non-media activities.

Both approaches were implemented in 1KA and were focused on respondents who are responding to a web survey on their PC; however, with some minor modifications both approaches could also be used on other survey platforms and/or on respondents using other devices to respond to a survey.

5.1.3 Empirical findings on respondent multitasking within the context of media multitasking

This Section is organized into four thematic Subsections. Each Subsection begins with a brief overview of insights from literature on media multitasking and respondent multitasking that were addressed throughout Chapter 2 . Within this context, each Subsection then presents key empirical findings of this dissertation from Chapter 4 .

Prevalence

In Section 2.2.4 I synthesized key findings on the prevalence of everyday media multitasking. Here it is important to note that the majority of published large-scale research is focused on internal media multitasking (i.e. multitasking that involves only media activities). Literature shows that the prevalence of media multitasking has been increasing in recent years and represents a substantial share (i.e. around 30%) of total time people spend on media activities. Research also shows that activities on modern media devices (especially PCs) and/or the Internet are especially likely to appear in multitasking combinations. In relation to this, several studies report that same-device multitasking on PCs is especially intensive – participants in these studies switched between two different activities on their PC every few minutes.

Section 2.3.4 contains an overview of the four empirical studies done on respondent multitasking in web surveys. The four studies vary considerably in terms of the conceptualization of respondent multitasking and methodological approaches; consequently, reported shares of respondent who were multitasking range from 22% to 62%.

Overall, a synthesis of literature (as summarized in the previous two paragraphs) has revealed that more research is needed to get insight on the overall prevalence of respondent multitasking in web surveys, as well as what specific types of respondent multitasking behavior (e.g. secondary activity forms, types of interruptions etc.) are particularly

prevalent. In relation to this, a research question was developed as follows.

The complete interpretation of results related to this research question can be found in Section 4.3.2, which also contains comparisons with previously published studies on respondent multitasking in web surveys.

For the purposes of this summary, it is important to note that more than half of respondents were engaged in at least one instance of (concurrent and/or sequential) respondent multitasking and that media activities were more prevalent than non-media activities. It is also important to reiterate that interruptions due to same-device activities were relatively common (25% of respondents reported on having at least one such interruption; according to paradata, this share was 37%), although non-media interruptions were even more common according to self-reports (29%), while other-device interruptions were the least common (18%).

It is crucial to understand that only about one fifth of respondents were engaged in more intensive respondent multitasking behavior (e.g. engagement in more than one secondary activity form or having more than one interruption). This was also observed in the supplementary cluster analysis of secondary activity forms. Members of clusters that described a relatively non-intensive respondent multitasking behavior (i.e. Listeners, Eaters, F2F talkers and Watchers) represent 25% of the whole sample, while members of clusters that describe a more intensive respondent multitasking behavior (i.e. Phone talker, Common mixers and Complex mixers) represent 20% of the entire sample.

To conclude, the empirical research in this dissertation related to research question 2 ("What is the prevalence of respondent multitasking in the web survey?") confirmed that respondent multitasking is relatively common behavior and is regularly included in everyday media multitasking combinations. While this seems alarming, it is important to keep in mind that a much smaller share of respondents is engaged in intensive

respondent multitasking behavior. In the particular context of this dissertation (web surveys on PC), it is also important to note that a substantial share of secondary activities (e.g. about a third of all self-reported interruptions) occurred on the same device that was used for responding to the survey.

Factors

Based on literature, I presented four groups of factors of media multitasking (Section 2.2.4). Among socio-demographic factors age has been particularly exposed as one of the most important determinants of media multitasking behavior. Research shows that younger people tend to be substantially more involved in multitasking behavior, particularly with media activities on modern devices. Individual factors can be broadly separated into access, ownership and usage of technology and personality traits such as sensation seeking and polychronicity. Structural/technological factors relate to general technological development and penetration in society. Lastly, research on motives reveals that people are multitasking with media for different reasons such as entertainment (e.g. people engage in a secondary activity to fight boredom), efficiency (e.g. people believe that multitasking is a more effective approach to complete multiple activities), and addiction (e.g. people are multitasking routinely and cannot strongly control this habit).

In respondent multitasking literature (Section 2.3.4), only two factors have been confirmed as significant by more than one published study: age and the survey device. To expand on the latter, studies have shown that respondents tended to multitask more if they used a smartphone to respond to a web survey (when compared with PC).

While there has been some limited research on other socio-demographic (e.g. marital status) and individual (e.g. experience with web surveys) factors, the relative lack of insight into this aspect of respondent multitasking has also been identified as a research gap. Moreover, in

Section 2.3.2 I have theoretically elaborated on two individual factors that could importantly influence respondent multitasking and have not been empirically investigated thus far: everyday media multitasking behavior and respondents' engagement in the survey (i.e. how focused and motivated are respondents to respond to the survey).

Following the research question 3 (What are the socio-demographic and individual factors related to RM?), results confirm that age is an important factor of respondent multitasking. Younger respondents were more likely to be engaged in all types of multitasking according to all indicators, except for time-out events. They were particularly more likely to be engaged in a secondary media activity, which again confirms findings from more general literature od media multitasking.

Moreover, analysis has revealed everyday media multitasking and self-reported attention (which was understood as an indicator of one of the aspects of survey engagement) were also the most important factors of respondent multitasking in this survey. This is particularly true for the latter, which was strongly associated with all indicators of respondent multitasking. On the other hand, everyday media multitasking was associated only with indicators based on self-reports.

Other socio-demographic (gender and education) and individual (e.g. experiences with surveys, access to technology) factors were weakly associated with only some specific types of respondent multitasking. A more detailed discussion of the findings on factors can be found in Section 4.3.3.

The relationship between multitasking and quality of performance

The relationship between multitasking and quality of performance was firstly theoretically addressed in Section 2.1 with Threaded Cognition Theory by Salvucci and Taatgen (2011). This theory describes how multitasking can hinder the quality of performance in goal-oriented tasks. However, this relation depends on several factors such as the complexity

and similarity of the combined activities, the nature of interruptions, and general cognitive abilities.

Section 2.2.4 provides empirical evidences in the context of media multitasking literature and further expands on this relationship. In the same Section I also introduce another important stream of research pioneered by Ophir et al. (2009). This stream of research has provided empirical evidence that people who have a habit of intensive media multitasking in their everyday life (i.e. high media multitaskers) perform worse in cognitive activities that require a sustained level of attention, even if they are multitasking while doing such activities.

By combining these theoretical and empirical insights with survey methodology literature, I have elaborated on the proposed relationship between multitasking and response quality in web surveys (Section 2.3.2). While other authors have already elaborated on this relationship and expressed concerns that respondent multitasking could be importantly related to worse response quality, I further expand the discussion, mainly by proposing that this relationship can manifest on three levels. Level 1 describes the direct causal relationship between respondent multitasking and response quality (as explained with the Threaded Cognition Theory).

On level 2, the relationship involves another element – respondents' survey engagement. At this level, respondent multitasking and survey engagement are in an interdependent relationship where more multitasking could lead to a lower engagement while lower engagement could lead to more multitasking as well. Lower engagement could lead to worse response quality, independently of effects of respondent multitasking described in level 1.

Lastly, level 3 describes the relationship between everyday media multitasking habits and response quality. Based on research summarized in the previous paragraph, it is assumed that high media multitaskers

provide a worse response quality, regardless whether they are engaged in secondary activities during the survey or not.

In Section 2.3.2 I have also elaborated on several other key factors that could influence the relationship between multitasking and response quality. This includes characteristics of respondent multitasking (e.g. complexity of secondary activities), perceived difficulty of the survey and cognitive abilities.

In Section 2.3.4 I provide an overview of the scarce empirical evidence on this relationship. In most cases the authors could not confirm a significant relationship between respondent multitasking and response quality indicators – even when they did, effect sizes were small. I have also noted that several important aspects of this relationship (e.g. inclusion of some important factors; experimental research into the causal relationship as described at level 1; investigation of relationship at level 2) have not been empirically investigated so far. In relation to this identified research gap, research question 4 was formulated as "What is the relationship between RM indicators (based on paradata and self-reports) and RQ indicators?"

This dissertation explored the relationship of respondent multitasking with two types of respondent behavior associated with worse response quality: non-response behavior (observed with unit-level item non-response) and satisficing (observed with estimation as the response strategy for behavioral frequency question, trap item, non-differentiation, inconsistent responses and length of responses to open-ended questions). The research was focused on the relationship on level 2; however, the role of everyday media multitasking habits was explored as well.

The analysis in this dissertation did not find a strong relationship between respondent multitasking and response quality. Only time-out indicators predict a lower response quality according to two indicators (trap item and non-differentiation). An important insight was also that self-reported attention and everyday media multitasking were more strongly associated

with response quality than response quality indicators. This indicates that more investigation is warranted into the relationship at level 2 and 3. An extensive discussion of results can be found in Section 4.3.4.

To sum up, the results of this dissertation are in line with findings from previous empirical research in web surveys as well as other survey modes – respondent multitasking is not notably related to worse response quality.

On the first look, this makes concerns over respondent multitasking seem unwarranted. However, it is important to keep in mind that these concerns are founded on a strong theoretical foundation and supported by numerous empirical evidences from general multitasking and media multitasking literature. Moreover, a vast majority of studies (including this dissertation) was done on members of internet panels who can exhibit substantially different behavior than respondents in a non-panel environment. Thus far it is possible to conclude that while the overall prevalence of respondent multitasking among panel members is relatively large, the share of respondents who are intensively engaged in secondary activities is relatively small, as is also the share of respondents with notably low levels of observed response quality. These low shares also impede analysis related to the investigation of this relationship.

In conclusion, while a strong relationship between respondent multitasking and response quality has not thus far been statistically confirmed within the general context of internet panels, it is possible that this relationship is more evident in some more specific settings. Another motivation for further research can be found in reports on trends of increasing prevalence of everyday media multitasking. As shown in this dissertation, there are some indications that high media multitaskers have worse response quality than other respondents (regardless of whether they are multitasking during the survey or not). Moreover, this dissertation also suggests that more intensive everyday media

multitasking behavior increases the chances that respondent will be multitasking during the survey. Consequently, it is reasonable to assume that prevalence and intensity of respondent multitasking will increase in the future. Therefore, it is important that survey methodology keeps monitoring this respondent behavior and its possible relationship with response quality.

Comparison and evaluation of reactive (self-reports) and non-reactive (paradata) approaches

After I have elaborated on reactive and non-reactive approaches in media multitasking and respondent multitasking literature (already summarized here in Section 5.1.1.), I have also noted that there has been no empirical research that would systematically compare the two approaches within the context of respondent multitasking in web surveys (Section 2.3.4).

There are two general ways of how the two different data collection approaches can be compared. Firstly, there is a direct comparison of measurements obtained with the two approaches, e.g. by looking at what percentage of values match. In relation to this, research question 1 was stated as: "How are paradata indicators on respondent multitasking associated with self-reports?". Corresponding results are discussed in Section 4.3.1.

The other way of comparing the approaches is by looking whether relying on a single approach would lead to different conclusions in exploratory models e.g. by comparing results of models on the relationship of indicators of respondent multitasking with factors and response quality indicators. In relation to this, the main hypothesis was stated as: "Paradata-based procedures (focus-out and time-out events) can measure respondent multitasking in web surveys and they provide similar general findings on prevalence of multitasking behavior, its factors and relationship with response quality, as the approaches based on self-reports.". Corresponding results are discussed in Section 4.3.5.

In general, both data collection approaches give similar figures on the overall prevalence of respondent multitasking in the survey. The share of respondents who had at least one instance of multitasking was 63% according to self-reports and 59% according to paradata. However, only 44% of respondents had at least one instance of multitasking according to both approaches and 23% had none according to both indicators. This simple dichotomous measure thus matched two thirds of the sample.

A larger share of respondents was multitasking according to self-reports and not according to paradata (19%) than the opposite (14%). Discrepancies show the interplay of the main limitations of each approach: paradata can hardly measure all types of respondent multitasking behavior (e.g. concurrent secondary activities) while self-reports are affected by underreporting. Despite the differences, an important insight revealed by the analysis is that respondents who were observed as multitaskers by both approaches tended to be engaged in more intensive respondent behavior.

Comparison of separate models that included different respondent multitasking indicators as dependent variables and factors as independent variables did give notably different findings. While some key factors (age, self-reported attention) are significant predictors in models containing respondent multitasking indicators from both approaches, several factors (e.g. education, gender, everyday media multitasking) were mostly significant only for self-reports. The analysis of factors also showed that time-out events might be discriminative towards older respondents (i.e. there could be a relatively high amount of false positive time-out events for older respondents who are generally responding to web surveys at a slower pace, even though the calculation of time-out events did try to account for this).

A similar comparison of separate models that included response quality indicators as dependent variables and different sets of respondent

multitasking indicators as independent variables showed that substantive findings do not substantially change if I would rely only on one of the data collection approaches. While paradata perhaps showed marginally more meaningful associations due to time-out events, effects on response quality are small (and non-alarming) in both cases.

To conclude – on one hand, there were important differences between the approaches that were also the reason why the hypothesis was not confirmed. Namely, even when looking at the simple outcome such as whether a respondent has multitasked or not, the approaches match only in two thirds of cases. Moreover, indicators based on self-reports had overall more meaningful relationships with factors.

On the other hand, the approaches did give similar figures on the overall prevalence of respondent multitasking and were both successful at identifying intensive respondent multitaskers. The latter is particularly important since intensive respondent multitaskers are typically one of the main interests in this line of research. Furthermore, both approaches gave similar results on the relationship with response quality.

Thus, this dissertation has shown that paradata can be a viable approach for research on respondent multitasking. The appropriateness of a paradata approach depends on specific research objectives, particularly whether the key findings can be importantly affected by specific limitations of paradata (as described above). It is important to keep in mind that this dissertation is the first published attempt of extensive use of paradata in respondent multitasking research. As such, I believe that this research has shown the potentials and appeal of paradata, especially considering the fact that such approach is less demanding for respondents (no additional questions) and researchers (once the appropriate technical procedures are implemented, the approach can be used in different surveys without additional complications).

5.2 Limitations

The empirical research in this dissertation was limited in several aspects. Firstly, important limitations are inherent to the research design, as listed below.

- Research on a non-probability sample. It is important to stress anew that this study did not use a probability sample. Therefore, the reported p-values should be taken with reservations and the findings are not meant to be generalized to a general population.

- Research on internet panel. Panel members generally have more experiences with web surveys than people who are not panel members and this could importantly influence the findings of this dissertation. For example, since panel members typically have a greater knowledge of the primary activity (i.e. responding to a web survey), engagement in (certain types of) secondary activities could have less of an effect on response quality. For example, in a study done on a non-panel student sample (Sendelbah et al. 2016) focus-out events were in a more meaningful relationship with unit-level item non-response.

- Non-experimental design. This study did not use an experimental design; strictly speaking, this means that it is not possible to make conclusions on the causality of the relationship between respondent multitasking and response quality.

- Specifics of the data collection approaches for measuring respondent multitasking. Both implemented approaches represent only one of the possibilities for measuring respondent multitasking with paradata and self-reports. This is particularly true for the paradata approach where indicators partially depend on arbitrary criteria (the 5s minimum limit for focus-out event, the cut-off point definition for time-out events). In relation to this, I should mention that I have tried out some other criteria (e.g. 10s minimum limit for focus-out events) but this did not

substantially change the main findings. Nevertheless, more detailed investigation into these criteria is needed. Moreover, in the theoretical framework I have noted that other types of paradata events can be used for this purpose (e.g. question response times or inactivity of input devices). In case of self-reports there are numerous possible variations on how questions on respondent multitasking could be designed.

- Factors of respondent multitasking. Research design did not include several factors that could give an important insight into respondent multitasking or even influence some of the findings of this dissertation. This includes structural factors related to questionnaire properties (e.g. length, question topics, question types), survey device (e.g. respondents on smartphones), motives (i.e. why have respondents engaged in secondary activities) and other.

- Response quality indicators. While this dissertation included a wide variety of response quality indicators, some important indicators related to non-response (e.g. drop-outs) and satisficing (e.g. primacy effect, non-substantive responses, cognitive reflection test) were not a part of this research.

The second group of limitations is related to the performed analysis. This includes:

- Transformation of indicators. It could be beneficial to transform some indicators of respondent multitasking, response quality and factors. This is particularly true for models that assumed a generalized linear relationship between dependent and independent variables. Here it is important to note that some transformations were tried out in the initial stages of analysis; however, these attempts did not notably change the outcomes. In the end, I have decided to conduct a full analysis by using non-transformed variables in order to present more interpretable findings.

- More detailed use of indicators on respondent multitasking based on paradata. This dissertation mainly analyzed paradata indicators aggregated on a respondent level. This neglected one of the advantages of this data collection approach – the fact that paradata events are observed on a more detailed level. For example, it could be useful to see if focus-out and time-out events that occur on certain pages are tighter related to response quality. Another neglected additional insight provided by focus-out events is that they include information on the exact length of same-device interruptions – this information could for example be used to analyze whether the length of focus-out event is an important factor in the relationship with response quality.

- More detailed use of indicators on respondent multitasking based on self-reports. Similar to the previous bullet point, observations obtained with self-reports could also be examined in greater details. One way of doing this is demonstrated in this dissertation with clustering of secondary activity forms. While this cluster analysis did bring some additional insight, there is room for refinement. To give a specific example, it could be valuable to employ a more theory-driven approach by using existing literature to determine which secondary activity forms are particularly cognitively demanding (e.g. phone conversations, working on text documents) and subsequently using this information to contrast respondents who were engaged in such activities with other respondents.

- More elaborated analysis. In the theoretical framework I have elaborated on the complex interplay between respondent multitasking, response quality and factors. One particular approach that could account for this proposed relationship is structural equation modeling.

5.3 Directions for further research

Research on respondent multitasking in web surveys is still in its beginning stages. The newness and dynamic of research is evident from the fact that no studies have been published in the time of writing the disposition of this dissertation. In the meantime, four studies have been published (Zwarun and Hall 2014; Ansolabehere and Schaffner 2015; Antoun 2015; Sendelbah et al. 2016) and just before the finalization of the submission of this thesis I became aware of a new study that utilizes focus-out events (Höhne and Schlosser 2017).

Overall, research on respondent multitasking in web surveys topic still needs to address some fundamental problems. Based on the theoretical and empirical findings of this dissertation, I identify the following directions for further research:

- Further development and evaluation of data collection approaches. While this dissertation provided evidence that paradata can be a viable approach for measuring respondent multitasking, an open question remains on what is the most optimal specific method for research on this topic. Here I am not only talking about the distinction between paradata and self-reports, but about specifics such as which paradata events and question designs present the most appropriate way of measuring respondent multitasking. Furthermore, it would be also worthwhile to investigate whether a combined approach (i.e. combination of paradata and self-reports) yields additional insights. Overall, one of the main goals should be developing an approach that could and would be used across different applied research settings.

- Further investigation of factors of respondent multitasking. Literature on respondent multitasking (including this dissertation) has encompassed some important factors, however several important factors as identified by the theoretical framework have so far not been

investigated (e.g., see previous Section on limitations). Moreover, even those factors that have already been included in research projects still warrant further investigation. For example, this dissertation has provided some limited evidences on the importance of the respondents' engagement in the survey; however, this factor should be examined more holistically.

- Further elaboration on the analysis of the relationship between respondent multitasking and response quality. Based on the theoretical framework and some limited empirical evidences, it is reasonable to assume that some specific intensive types of respondent multitasking behavior could be related to worse response quality. Future research should focus on developing appropriate analytical procedures to identify cases of intensive respondent multitasking and to examine their relationship with response quality. This includes exploring the full potential of measurements obtained with data collection procedures (as explained in the previous Section).

- Applied research in non-panel surveys. As already noted, the majority of research on respondent multitasking so far has been limited to internet panels. It would be very beneficial to conduct research on respondent multitasking on surveys that use other sampling methods.

- Experimental and qualitative research design. All of the published empirical literature on respondent multitasking follows a quantitative non-experimental research design. Experimental and qualitative designs could yield new important insights. For example, experimental design can be used to thoroughly explore the causality of the relationship between respondent multitasking and response quality.

5.4 Original contribution of the dissertation

This dissertation has provided several original contributions within the context of research on respondent multitasking as well as in the broader context of research on media multitasking.

The theoretical framework developed in this dissertation provides the following original contributions:

- Integration of the Threaded Cognition Theory in the context of respondent multitasking. Integration of this theory by Salvucci and Taatgen (2011) was valuable for systematic and holistic elaboration on conceptualization of respondent multitasking as well as its relationship with response quality.

- Further development of taxonomy of respondent multitasking. Taxonomy by Zwarun and Hall (2014) was further extended based on key insights from literature on multitasking and media multitasking. Extended taxonomy can be particularly useful for designing methodological approaches for research on respondent multitasking.

- Elaboration on factors of respondent multitasking and its relationship with response quality. Based on insights from media multitasking literature, I have presented four major groups of factors that could influence respondent multitasking behavior. Building on multiple research streams, I have also elaborated on the relationship between multitasking, response quality and key factors; specifically, I proposed that this relationship could manifest on three distinct levels.

- Systematic overview of conceptualizations and research approaches in media multitasking literature. This systematic overview was a necessary by-product of the theoretical framework; nevertheless, it could prove to be valuable for further research on media multitasking.

While some other authors have already reviewed and organized conceptualizations of media multitasking, this dissertation was particularly focused on methodological aspects.

An inherent part of this dissertation was the development of specific data collection approaches with some novel elements:

- Respondent multitasking indicators based on paradata. While a limited pilot study utilizing focus-out and time-out events was previously published in Sendelbah et al. (2016), this dissertation has further developed this approach and can be understood as the first published large-scale attempt at using paradata in research on respondent multitasking. A particularly important innovation here was the strategy for determining time-out events. In a broader context, this approach follows the recommendations and calls for using non-reactive data in a different context of survey methodology and media multitasking research.

- Respondent multitasking indicators based on self-reports. This dissertation presented a new set of questions for prompting respondents to report on their respondent multitasking behavior. Question formation was mainly guided by taxonomy of respondent multitasking as well as by media multitasking literature (particularly by work by Ophir et al. (2009)) in order to have a well-rounded instrument for measuring respondent multitasking via self-reports.

Lastly, some key empirical findings of this dissertation also present original contributions:

- Comparison of reactive and non-reactive data collection approaches. To the best of my knowledge, this dissertation presents the first comparison of the two approaches within the research on respondent multitasking. In a broader context, I hope that the findings of this dissertation will be also a valuable addition to the research that

compares reactive and non-reactive approaches for measuring different types of media multitasking.

- New insight into prevalence of respondent multitasking in web surveys. Data collection approaches used in this dissertation allowed for a detailed analysis of prevalence of respondent multitasking. One specific novel way of investigating prevalence of respondent multitasking was the clustering of secondary activities.

- New insight into the relationship between multitasking, response quality and factors. Based on the theoretical framework, this dissertation investigated some previously unexplored aspects of the complex relationship between multitasking, response quality and related factors. Particularly important insights here are indications on the importance of respondents' survey engagement and everyday media multitasking habits. Both of these factors were related to the prevalence of respondent multitasking as well as to response quality. The role of media multitasking habits is also interesting within the context of research on this increasingly prevalent everyday behavior. So far, research on the relationship between intensive everyday media multitasking and performance in goal-oriented activities was mainly conducted in laboratory experiments; this dissertation brings new evidence on this relationship in an applied, real-world setting.

References

"1KA Survey on Media Multitasking in Web Surveys." 2017. https://www.1ka.si/a/93628&preview=on (July 15 2017).

Abeele, Mariek Vanden, Kathleen Beullens, and Keith Roe. 2013. Measuring Mobile Phone Use: Gender, Age and Real Usage Level in Relation to the Accuracy and Validity of Self-Reported Mobile Phone Use. *Mobile Media & Communication* 1 (2): 213–236. doi: 10.1177/2050157913477095.

Adler, Rachel F., and Raquel Benbunan-Fich. 2012. Juggling on a High Wire: Multitasking Effects on Performance. *International Journal of Human-Computer Studies* 70 (2): 156–68. doi:10.1016/j.ijhcs.2011.10.003.

Agosto, Denise E., and June Abbas. 2010. High School Seniors' Social Network and Other ICT Use Preferences and Concerns. *Proceedings of the American Society for Information Science and Technology* 47 (1): 1–10. doi:10.1002/meet.14504701025.

Alexopoulou, Peggy, Mark Hepworth, and Anne Morris. 2015. An Investigation of Multitasking Information Behavior and the Influence of Working Memory and Flow. *AIP Conference Proceedings*, 1644:37–43. doi:10.1063/1.4907815.

Alzahabi, Reem, and Mark W. Becker. 2013. The Association between Media Multitasking, Task-Switching, and Dual-Task Performance. *Journal of Experimental Psychology: Human Perception and Performance* 39 (5): 1485–95. doi:10.1037/a0031208.

Ansolabehere, Stephen, and Brian F. Schaffner. 2015. Distractions: The Incidence and Consequences of Interruptions for Survey Respondents. *Journal of Survey Statistics and Methodology* 3 (2): 216–39. doi:10.1093/jssam/smv003.

Antoun, Christopher. 2015. Mobile Web Surveys: A First Look at Measurement, Nonresponse, and Coverage Errors. https://deepblue.lib.umich.edu/handle/2027.42/116722 (July 15 2017).

Bardhi, Fleura, Andrew J. Rohm, and Fareena Sultan. 2010. Tuning in and Tuning out: Media Multitasking among Young Consumers. *Journal of Consumer Behaviour* 9 (4): 316–32. doi:10.1002/cb.320.

Barlas, F.M, R. K. Thomas, and N. R. Buttermore. 2016. Can Survey Data Be Too Clean?: Data Cleaning and Bias Reduction. Presented at *RC33 International Conference on Social Science Methodology*, Leicester, United Kingdom, September 16 2016. https://www.gfk.com/fileadmin/user_upload/dyna_content/US/documents/2016_RC332016_Barlas01_Can_Survey_Data_be_Too_Clean.pdf (July 15 2017).

Baron, Naomi. 2008. Adjusting the Volume: Technology and Multitasking in Discourse Control. In *Handbook of Mobile Communication Studies*,

ed. James E. Katz, 177–193. Massachusetts: The MIT Press. doi:10.7551/mitpress/9780262113120.003.0014.

Beckers, Tilo, Pascal Siegers, and Anabel Kuntz. 2011. Speeders in Online Value Research: Cross-Checking Results of Fast and Slow Respondents in Two Separate Samples Answering the 40 Item Portrait Value Questionnaire. Presented at the *General Online Research Conference*, Düsseldorf, Germany, March 15 2011. http://conftool.gor.de/conftool11/index.php?page=downloadPaper&filename=Beckers-Speeders_in_Online_Value_Research-172.pdf&form_id=172&form_index=&form_version=final (July 15 2017).

Belli, Robert, Norbert Schwarz, Eleanor Singer, and Jennifer Talarico. 2000. Decomposition Can Harm the Accuracy of Behavioral Frequency Reports. *Applied Cognitive Psychology* 14 (4):295-308. doi: 10.1002/1099-0720(200007/08)14:4<295::AID-0ACP646>3.0.CO;2-1.

Berolo, Sophia, Ivan Steenstra, Benjamin C. Amick III, and Richard P. Wells. 2015. A Comparison of Two Methods to Assess the Usage of Mobile Hand-Held Communication Devices. *Journal of Occupational and Environmental Hygiene* 12 (4): 276–285. doi: 10.1080/15459624.2014.973111.

Berzelak, Jernej. 2014. *Mode Effects in Web Survey*. University of Ljubljana. http://dk.fdv.uni-lj.si/doktorska_dela/pdfs/dr_berzelak-jernej.pdf (15 July 2017).

Bethlehem, Jelke, and Silvia Biffignandi. 2011. *Handbook of Web Surveys*. Hoboken NJ: Wiley.

Blair, Edward, and Scot Burton. 1987. Cognitive Processes Used by Survey Respondents to Answer Behavioral Frequency Questions. *Journal of Consumer Research*, 280–288. www.jstor.org/stable/2489417 (July 15 2017)

Blake, Christopher, and Christoph Klimmt. 2012. The Challenge of Measuring the Use of Computer Games. In *Computer Games and New Media Cultures*, ed. Johannes Fromme and Alexander Unger, 357–69. Netherlands: Springer. doi: 10.1007/978-94-007-2777-9_23.

Bolger, Niall, Angelina Davis, and Eshkol Rafaeli. 2003. Diary Methods: Capturing Life as It Is Lived. *Annual Review of Psychology* 54 (1): 579–616. doi:10.1146/annurev.psych.54.101601.145030.

Bosnjak, Michael, and Tracy L. Tuten. 2001. Classifying Response Behaviors in Web-Based Surveys. *Journal of Computer Mediated Communication* 6 (3). doi:10.1111/j.1083-6101.2001.tb00124.x.

Bouwman, Harry, Mark de Reuver, and N. Shahrokh. 2013. Mixed Method Research: Combining Mobile Log Data and Questionnaires. In *Proceedings of the 24th European Regional Conference of the International Telecommunication Society.* http://repository.tudelft.nl/assets/uuid:c36e1912-5e22-4877-8fba-698c61f9d675/302178.pdf (July 15 2017).

Bowman, Laura L., Laura E. Levine, Bradley M. Waite, and Michael Gendron. 2010. Can Students Really Multitask? An Experimental Study of Instant Messaging While Reading. *Computers & Education* 54 (4): 927–31. doi:10.1016/j.compedu.2009.09.024.

Brasel, S. Adam, and James Gips. 2011. Media Multitasking Behavior: Concurrent Television and Computer Usage. *Cyberpsychology, Behavior, and Social Networking* 14 (9): 527–534. doi:10.1089/cyber.2010.0350.

Burton, Scot, and Edward Blair. 1991. Task Conditions, Response Formulation Processes, and Response Accuracy for Behavioral Frequency Questions in Surveys. *Public Opinion Quarterly* 55 (1): 50–79. doi:10.1086/269241.

Cain, Matthew S, and Stephen R Mitroff. 2011. Distractor Filtering in Media Multitaskers. *Perception* 40 (10): 1183–92. doi:10.1068/p7017.

Callegaro, Mario. 2013. Paradata in Web Surveys. In *Improving Surveys with Paradata: Analytic Uses of Process Information*, ed. Frauke Kreuter, 261–79. Hoboken, NJ: John Wiley & Sons. 10.1002/9781118596869.ch11.

Callegaro, Mario, Katja Lozar Manfreda, and Vasja Vehovar. 2015. *Web Survey Methodology*. London, United Kingdom: SAGE Publications Ltd.

Callegaro, Mario, Yongwei Yang, Dennison S. Bhola, Don A. Dillman, and Tzu-Yun Chin. 2009. Response Latency as an Indicator of Optimizing in Online Questionnaires. *Bulletin de Méthodologie Sociologique* 103 (1): 5–25. doi:10.1177/075910630910300103.

Campbell, Danny, Morten Raun Mørkbak, and Søren Bøye Olsen. 2017. Response Time in Online Stated Choice Experiments: The Non-Triviality of Identifying Fast and Slow Respondents. *Journal of Environmental Economics and Policy* 6 (1): 17–35. doi:10.1080/21606544.2016.1167632.

Carrier, L. Mark, Nancy A. Cheever, Larry D. Rosen, Sandra Benitez, and Jennifer Chang. 2009. Multitasking across Generations: Multitasking Choices and Difficulty Ratings in Three Generations of Americans. *Computers in Human Behavior* 25 (2): 483–489. doi:10.1016/j.chb.2008.10.012.

Chang, LinChiat, and Jon A. Krosnick. 2003. Measuring the Frequency of Regular Behaviors: Comparing the Typical Week to the Past Week. *Sociological Methodology* 33 (1): 55–80. doi:10.1111/j.0081-1750.2003.t01-1-00127.x.

Chen, Lin, Sriram Venkataraman, and Sandy Jap. 2012. Media Multiplexing Behavior: Implications for Targeting and Media Planning. *Marketing Science* 32 (2): 310-24. doi: https://doi.org/10.1287/mksc.1120.0759.

Chinchanachokchai, Sydney, Brittany R. L. Duff, and Sela Sar. 2015. The Effect of Multitasking on Time Perception, Enjoyment, and Ad Evaluation. *Computers in Human Behavior* 45 (April): 185–91. doi:10.1016/j.chb.2014.11.087.

Circella, Giovanni, Patricia L. Mokhtarian, and Laura K. Poff. 2012. A Conceptual Typology of Multitasking Behavior and Polychronicity Preferences. *Electronic International Journal of Time Use Research* 9 (1): 59–107.

Collopy, Fred. 1996. Biases in Retrospective Self-Reports of Time Use: An Empirical Study of Computer Users. *Management Science* 42 (5): 758–767.

Cotten, Shelia R., Daniel B. Shank, and William A. Anderson. 2014. Gender, Technology Use and Ownership, and Media-Based Multitasking among Middle School Students. *Computers in Human Behavior*. doi:10.1016/j.chb.2014.02.041.

Couper, Mick P., and Frauke Kreuter. 2013. Using Paradata to Explore Item Level Response Times in Surveys. *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 176 (1): 271–86. doi:10.1111/j.1467-985X.2012.01041.x.

Courage, Mary L., Aishah Bakhtiar, Cheryll Fitzpatrick, Sophie Kenny, and Katie Brandeau. 2015. Growing up Multitasking: The Costs and Benefits for Cognitive Development. *Developmental Review, Special Issue: Living in the "Net" Generation: Multitasking, Learning, and Development, 35 (March)*: 5–41. doi:10.1016/j.dr.2014.12.002.

Cribari-Neto, Francisco, and Achim Zeileis. 2009. Beta Regression in R. *Journal of Statistical Software* 34 (2). www.jstatsoft.org/v34/i02/ (July 15 2017).

De Leeuw, Edith D. 2008. Choosing the Method of Data Collection. In *International Handbook of Survey Methodology*, ed. Edith D. De Leeuw, Joop Hox, and Don Dillman, 113-136. New York: Lawrence Erlbaum Associates. http://joophox.net/papers/SurveyHandbookCRC.pdf (July 15 2017).

Deane, Frank P., John Podd, and Ron D. Henderson. 1998. Relationship between Self-Report and Log Data Estimates of Information System Usage. *Computers in Human Behavior* 14 (4): 621–636. doi: 10.1016/S0747-5632(98)00027-2.

Downes-Le Guin, Theo, Reg Baker, Joanne Mechling, and Erica Ruylea. 2012. Myths and Realities of Respondent Engagement in Online Surveys. *International Journal of Market Research* 54 (5): 1–21. doi:10.2501/IJMR-54-5-000-000.

EnKlikAnketa - 1KA Web Surveys. 2016. http://english.1ka.si/c/694/General_description (July 15 2017).

Fay, Michael P. 2010. Two-Sided Exact Tests and Matching Confidence Intervals for Discrete Data. *R Journal* 2 (1): 53–58.

Ferguson, Douglas A. 1994. Measurement of Mundane TV Behaviors: Remote Control Device Flipping Frequency. *Journal of Broadcasting & Electronic Media* 38 (1): 35–47. doi:10.1080/08838159409364244.

Foehr, Ulla G. 2006. Media Multitasking among American Youth: Prevalence, Predictors and Pairings. Henry J. Kaiser Family Foundation.

http://kaiserfamilyfoundation.files.wordpress.com/2013/01/7592.pdf (July 15 2017).

Fried, Carrie B. 2008. In-Class Laptop Use and Its Effects on Student Learning. *Computers & Education* 50 (3): 906–914. doi:10.1016/j.compedu.2006.09.006.

Galesic, Mirta, and Michael Bosnjak. 2009. Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey. *The Public Opinion Quarterly* 73 (2): 349–60.

Ganassali, Stéphane. 2008. The Influence of the Design of Web Survey Questionnaires on the Quality of Responses. *Survey Research Methods* 2 (1): 21–32.

Gil de Zúñiga, Homero, Victor Garcia-Perdomo, and Shannon C. McGregor. 2015. What Is Second Screening? Exploring Motivations of Second Screen Use and Its Effect on Online Political Participation. *Journal of Communication* 65 (5): 793–815. doi:10.1111/jcom.12174.

Goldberg, Lewis R. 1990. An Alternative Description of Personality: The Big-Five Factor Structure. *Journal of Personality and Social Psychology* 59 (6): 1216.

Google. 2012. The New Multi-Screen World: Understanding Cross-Platform Consumer Behavior. https://ssl.gstatic.com/think/docs/the-new-multi-screen-world-study_research-studies.pdf (July 15 2017).

Gosling, Samuel D, Peter J Rentfrow, and William B Swann Jr. 2003. A Very Brief Measure of the Big-Five Personality Domains. *Journal of Research in Personality* 37 (6): 504–28. doi:10.1016/S0092-6566(03)00046-1.

Grace-Martin, Michael, and Geri Gay. 2001. Web Browsing, Mobile Computing and Academic Performance. *Educational Technology & Society* 4: 3.

Greenberg, Bradley S., Matthew S. Eastin, Paul Skalski, Len Cooper, Mark Levy, and Ken Lachlan. 2005. Comparing Survey and Diary Measures of Internet and Traditional Media Use. *Communication Reports* 18 (1–2): 1–8. doi:10.1080/08934210500084164.

Greszki, Robert, Marco Meyer, and Harald Schoen. 2015. Exploring the Effects of Removing Too Fast Responses and Respondents from Web Surveys. *Public Opinion Quarterly* 79 (2): 471–503. doi:10.1093/poq/nfu058.

Gummer, Tobias, and Joss Roßmann. 2015. Explaining Interview Duration in Web Surveys A Multilevel Approach. *Social Science Computer Review* 33 (2): 217–34. doi:10.1177/0894439314533479.

Gutierrez, Christina, Tom Wells, Kumar Rao, and David Kurzynski. 2011. Catch Them When You Can: Speeders and Their Role in Online Data Quality. Presented at the *annual meeting of Midwest Association for Public Opinion Research*, Chicago, IL, November 2011.

Heerwegh, Dirk. 2003. Explaining Response Latencies and Changing Answers Using Client-Side Paradata from a Web Survey. *Social Science Computer Review* 21 (3): 360–73. doi:10.1177/0894439303253985.

———. 2004. Using Progress Indicators in Web Surveys. Presented at the *annual conference of the American Association for Public Opinion Research*, Phoenix, AZ, May 2004.

———. 2005. *Web Surveys: Explaining and Reducing Unit Nonresponse, Item Nonresponse and Partial Nonresponse*. University of Leuven.

Hembrooke, Helene, and Geri Gay. 2003. The Laptop and the Lecture: The Effects of Multitasking in Learning Environments. *Journal of Computing in Higher Education* 15 (1): 46–64.

Henning, R. A., P. Jacques, G. V. Kissel, A. B. Sullivan, and S. M. Alteras-Webb. 1997. Frequent Short Rest Breaks from Computer Work: Effects on Productivity and Well-Being at Two Field Sites. *Ergonomics* 40 (1): 78–91. doi:10.1080/001401397188396.

Hill, Shawndra. 2014. TV Audience Measurement with Big Data. *Big Data* 2 (2): 76–86. doi:10.1089/big.2014.0012.

Höhne, Jan Karem, and Stephan Schlosser. 2017. Investigating the Adequacy of Response Time Outlier Definitions in Computer-Based Web Surveys Using Paradata SurveyFocus. *Social Science Computer Review*. doi:10.1177/0894439317710450.

Holbrook, Allyson L., Melanie C. Green, and Jon A. Krosnick. 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. doi:10.1086/346010.

Holmes, M. E., R. A. Papper, M. N. Popovich, and M. Bloxham. 2005. Middletown Media Studies II: Concurrent Media Exposure. *Muncie, IN: Ball State University Center for Media Design*.

Hwang, Yoori, HyoungJee Kim, and Se-Hoon Jeong. 2014. Why Do Media Users Multitask?: Motives for General, Medium-Specific, and Content-Specific Types of Multitasking. *Computers in Human Behavior* 36: 542–48. doi:10.1016/j.chb.2014.04.040.

IAB Europe. 2012. Mediascope Europe Bulletin: Media Multi-Tasking Means More Active Consumers. http://www.iabeurope.eu/research-and-papers/media-multi-tasking-means-more-active-consumers-full-report (July 15 2017).

Ie, Amanda, Chiara S. Haller, Ellen J. Langer, and Delphine S. Courvoisier. 2012. Mindful Multitasking: The Relationship between Mindful Flexibility and Media Multitasking. *Computers in Human Behavior* 28 (4): 1526–32. doi:10.1016/j.chb.2012.03.022.

Iqbal, Shamsi T., and Eric Horvitz. 2007. "Disruption and Recovery of Computing Tasks: Field Study, Analysis, and Directions." *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 677–686. doi:10.1145/1240624.1240730.

Jacobsen, Wade C., and Renata Forste. 2011. The Wired Generation: Academic and Social Outcomes of Electronic Media Use among University Students. *Cyberpsychology, Behavior, and Social Networking* 14 (5): 275–80.

Jeong, Se-Hoon, and Martin Fishbein. 2007. Predictors of Multitasking with Media: Media Factors and Audience Factors. *Media Psychology* 10 (3): 364–84. doi:10.1080/15213260701532948.

Jeong, Se-Hoon, Weiyu Zhang, Elisabeth Davis, Amy Jordan, Martin Fishbein, Michael Hennessy, and Suzanne Martin. 2005. Multiple Media Use and Multitasking with Media among High School and College Students. Presented at the *General Online Research Conference*, Zurich, Switzerland, May 25. http://www.academia.edu/227042/Multiple_media_use_and_multitasking_with_media_among_high_school_and_college_students_A_Diary_method (July 15 2017).

Judd, Terry. 2013. Making Sense of Multitasking: Key Behaviours. *Computers & Education* 63: 358–67. doi:10.1016/j.compedu.2012.12.017.

Judd, Terry, and Gregor Kennedy. 2011. Measurement and Evidence of Computer-Based Task Switching and Multitasking by Net Generation Students. *Computers & Education* 56 (3): 625–31. doi:10.1016/j.compedu.2010.10.004.

Junco, Reynol, and Shelia R. Cotten. 2012. No A 4 U: The Relationship between Multitasking and Academic Performance. *Computers & Education* 59 (2): 505–14. doi:10.1016/j.compedu.2011.12.023.

Kaczmirek, Lars. 2008. Human-Survey Interaction: Usability and Nonresponse in Online Surveys. University of Mannheim. https://ub-madoc.bib.uni-mannheim.de/2150/ (July 15 2017).

Kaye, Barbara K., and Barry S. Sapolsky. 1997. Electronic Monitoring of In-home Television RCD Usage. *Journal of Broadcasting & Electronic Media* 41 (2): 214–28. doi:10.1080/08838159709364402.

Kennedy, Courtney Kathryn. 2010. *Nonresponse and Measurement Error in Mobile Phone Surveys*. University of Michigan. http://deepblue.lib.umich.edu/handle/2027.42/75977 (July 15 2017).

Kenyon, Susan. 2010. What Do We Mean by Multitasking?–Exploring the Need for Methodological Clarification in Time Use Research. *Electronic International Journal of Time Use Research* 7 (1).

Kobayashi, Tetsuro, and Jeffrey Boase. 2012. No Such Effect? The Implications of Measurement Error in Self-Report Measures of Mobile Communication Use. *Communication Methods and Measures* 6 (2): 126–43. doi:10.1080/19312458.2012.679243.

König, Cornelius J., Lourdes Oberacher, and Martin Kleinmann. 2010. Personal and Situational Determinants of Multitasking at Work. *Journal of Personnel Psychology* 9. doi:10.1027/1866-5888/a000008.

König, Cornelius J., and Mary Waller. 2010. Time for Reflection: A Critical Examination of Polychronicity. *Human Performance* 23 (2): 173–190.

Kononova, Anastasia. 2013. Multitasking Across Borders: A Cross-National Study of Media Multitasking Behaviors, Its Antecedents, and

Outcomes. *International Journal of Communication* 7 (23): 1688–1710.

Kononova, Anastasia, and Saleem Alhabash. 2012. When One Medium Is Not Enough: Media Use and Media Multitasking among College Students in Kuwait. *Journal of Middle East Media* 8 (1): 1–28.

Kononova, Anastasia, and Yi-Hsuan Chiang. 2015. Why Do We Multitask with Media? Predictors of Media Multitasking among Internet Users in the United States and Taiwan. *Computers in Human Behavior* 50 (September): 31–41. doi:10.1016/j.chb.2015.03.052.

Kraushaar, James M., and David C. Novak. 2010. Examining the Effects of Student Multitasking with Laptops during the Lecture. *Journal of Information Systems Education* 21 (2): 241–51.

Krosnick, Jon A. 1991. Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys. *Applied Cognitive Psychology* 5 (3): 213–36. doi:10.1002/acp.2350050305.

Lagerstedt, Christina, Andreas Aurelius, Olle Findahl, Hemamali Pathirana, and Claus Popp Larsen. 2012. Understanding Internet User Behavior: Towards a Unified Methodology. *International Journal On Advances in Telecommunications* 5 (3 & 4): 153–162.

Lavrakas, Paul J., Trevor N. Tompson, Robert Benford, and C. Fleury. 2010. Investigating Data Quality in Cell Phone Surveying. Presented at the *annual conference of the American Association for Public Opinion Research*, Chicago, IL, May 20101.

Loh, Kep Kee, and Ryota Kanai. 2014. Higher Media Multi-Tasking Activity Is Associated with Smaller Gray-Matter Density in the Anterior Cingulate Cortex. *PLoS ONE* 9 (9). doi:10.1371/journal.pone.0106698.

Lottridge, Danielle M., E. Marschner, E. Wang, M. Romanovsky, and C. Nass. 2012. Browser Design Impacts Multitasking. *Proceedings of the Human Factors and Ergonomics Society Annual Meeting*, 1957–61. doi:10.1177/1071181312561289.

Lottridge, Danielle M., Christine Rosakranse, Catherine S. Oh, Sean J. Westwood, Katherine A. Baldoni, Abrey S. Mann, and Clifford I. Nass. 2015. The Effects of Chronic Multitasking on Analytical Writing. *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, 2967–2970. doi:10.1145/2702123.2702367.

Lui, Kelvin F. H., and Alan C.-N. Wong. 2012. Does Media Multitasking Always Hurt? A Positive Correlation between Multitasking and Multisensory Integration. *Psychonomic Bulletin & Review* 19 (4): 647–53. doi:10.3758/s13423-012-0245-7.

Lynn, Peter, and Olena Kaminska. 2011. The Impact of Mobile Phones on Survey Measurement Error. ISER Working Paper Series 2011-07. http://hdl.handle.net/10419/65992.

———. 2012. Factors Affecting Measurement Error in Mobile Phone Interviews. In *Telephone Surveys in Europe*, ed. Sabine Häder, Michael Häder, and Mike Kühne, 211–228. Berlin: Springer. doi:10.1007/978-3-642-25411-6_14.

Lynn, Peter, and Gerry Nicolaas. 2010. Making Good Use of Survey Paradata. *Survey Practice* 3 (2). http://surveypractice.org/index.php/SurveyPractice/article/view/127 (July 15 2017).

Maechler, Martin, Peter Rousseeuw, Anja Struyf, Mia Hubert, and Kurt Hornik. 2017. *Cluster: Cluster Analysis Basics and Extensions*.

Malhotra, Neil. 2008. Completion Time and Response Order Effects in Web Surveys. *Public Opinion Quarterly* 72 (5): 914–34. doi:10.1093/poq/nfn050.

Mark, Gloria, Stephen Voida, and Armand Cardello. 2012. A Pace Not Dictated by Electrons: An Empirical Study of Work without Email. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 555–564. doi:10.1145/2207676.2207754.

Mark, Gloria, Yiran Wang, and Melissa Niiya. 2014. Stress and Multitasking in Everyday College Life: An Empirical Study of Online Activity. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, 41–50. doi:10.1145/2556288.2557361.

Matjašič, Miha. 2015. *Parapodatki v spletnih anketah in pristopi za odkrivanje časovnih anomalij*. Univerza v Ljubljani, Fakulteta za družbene vede. https://repozitorij.uni-lj.si/IzpisGradiva.php?id=72928 (July 15 2017).

McCarthy, Jordan. 2013. *Enhanced Media Multitasking: The Restorative Cognitive Effects of Temporarily Escaping the Multitasking Mindset*. Stanford University. http://comm.stanford.edu/wp-content/uploads/2013/01/JordanMcCarthyMAThesis.pdf (July 15 2017).

Meng, Jingbo, and Daniel McDonald. 2009. Predictors and Impacts of TV Multitasking and Simultaneous Multiple Media Use. Presented at the annual meeting of the *International Communication Association*, Chicago, IL, May 2009. http://citation.allacademic.com//meta/p_mla_apa_research_citation/2/9/8/1/0/pages298106/p298106-1.php (June 15 2017).

Meyer, David E., and David E. Kieras. 1997. A Computational Theory of Executive Cognitive Processes and Multiple-Task Performance: Part I. Basic Mechanisms. *Psychological Review* 104 (1): 3–65. doi:10.1037/0033-295X.104.1.3.

Mijović, Pavle, Vanja Ković, Ivan Mačužić, Petar Todorović, Branislav Jeremić, Miloš Milovanović, and Ivan Gligorijević. 2015. Do Micro-Breaks Increase the Attention Level of an Assembly Worker? An ERP Study. *6th International Conference on Applied Human Factors and Ergonomics (AHFE 2015) and the Affiliated Conferences*: 5074–80. doi:10.1016/j.promfg.2015.07.521.

Milavsky, J. Ronald. 1992. How Good Is the A.C. Nielsen People-Meter System? A Review of the Report by the Committee on Nationwide Television Audience Measurement. *The Public Opinion Quarterly* 56 (1): 102–15.

Minear, Meredith, Faith Brasher, Mark McCurdy, Jack Lewis, and Andrea Younggren. 2013. Working Memory, Fluid Intelligence, and

Impulsiveness in Heavy Media Multitaskers. *Psychonomic Bulletin & Review* 20 (6): 1274–81. doi:10.3758/s13423-013-0456-6.
Moisala, M., V. Salmela, L. Hietajärvi, E. Salo, S. Carlson, O. Salonen, K. Lonka, K. Hakkarainen, Katariina Salmela-Aro, and K. Alho. 2016. Media Multitasking Is Associated with Distractibility and Increased Prefrontal Activity in Adolescents and Young Adults. *NeuroImage* 134: 113–121.
Mokhtari, Kouider, Julie Delello, and Carla Reichard. 2015. Connected Yet Distracted: Multitasking Among College Students. *Journal of College Reading and Learning* 45 (2): 164–80. doi:10.1080/10790195.2015.1021880.
Möller, Andreas, Matthias Kranz, Barbara Schmid, Luis Roalter, and Stefan Diewald. 2013. Investigating Self-Reporting Behavior in Long-Term Studies. *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*: 2931–2940. doi:10.1145/2470654.2481406.
Newell, Allen. 1994. *Unified Theories of Cognition*. Massachusetts: Harvard University Press.
Nielsen. 2012. *State of the Media: The Cross Platform Report Q2 2012*. http://www.nielsen.com/us/en/insights/reports/2012/state-of-the-media--cross-platform-report-q2-2012.html (July 15 2017).
Nielsen, Arthur C. 1942. Trends toward Mechanization of Radio Advertising. *Journal of Marketing* 6 (3): 217–28. doi:10.2307/1245867.
Nooner, Kate, and Lauren Schaefer. 2015. A Pilot Investigation of Media Multitasking and Trauma Symptoms Among Adolescents. *Journal of Human Behavior in the Social Environment* 25 (6): 687–96. doi:10.1080/10911359.2015.1013659.
Ofcom. 2010. The Consumer's Digital Day. https://www.ofcom.org.uk/__data/assets/pdf_file/0019/51913/consumers-digital-day.pdf (July 15 2017).
———. 2013. Communications Market Report 2013. http://stakeholders.ofcom.org.uk/market-data-research/market-data/communications-market-reports/cmr13/ (July 15 2017).
———. 2015. Media Multitasking: Report on Findings of Analysis of Our Digital Day Research. https://www.ofcom.org.uk/__data/assets/pdf_file/0020/47261/media_multitaskingl.pdf (July 15 2017).
———. 2016. Communications Market Report 2016. https://www.ofcom.org.uk/__data/assets/pdf_file/0024/26826/cmr_uk_2016.pdf (July 15 2017).
Olson, Kristen, and Bryan Parkhurst. 2013. Collecting Paradata for Measurement Error Evaluations. In *Improving Surveys with Paradata: Analytic Uses of Process Information*, ed. Frauke Kreuter, 43–72. Hoboken, NJ: John Wiley & Sons. doi:10.1002/9781118596869.ch3.
Ophir, Eyal, Clifford Nass, and Anthony D. Wagner. 2009. Cognitive Control in Media Multitaskers. *Proceedings of the National Academy*

of Sciences 106 (37): 15583–15587. doi:10.1073/pnas.0903620106.

Pea, Roy, Clifford Nass, Lyn Meheula, Marcus Rance, Aman Kumar, Holden Bamford, Matthew Nass, Aneesh Simha, Benjamin Stillerman, and Steven Yang. 2012. Media Use, Face-to-Face Communication, Media Multitasking, and Social Well-Being among 8-to 12-Year-Old Girls. *Developmental Psychology* 48 (2): 327–36. doi:10.1037/a0027030.

Petrič, Gregor, Andraž Petrovčič, and Vasja Vehovar. 2011. Social Uses of Interpersonal Communication Technologies in a Complex Media Environment. *European Journal of Communication* 26 (2): 116–32. doi:10.1177/0267323111402654.

Pew Research Center. 2006. *The Cell Phone Challenge to Survey Research*. http://www.people-press.org/files/legacy-pdf/276.pdf.

Pilotta, Joseph J., Don E. Schultz, Gary Drenik, and Philip Rist. 2004. Simultaneous Media Usage: A Critical Consumer Orientation to Media Planning. *Journal of Consumer Behaviour* 3 (3): 285–292.

R Core Team. 2017. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. https://www.R-project.org/ (July 15 2017).

Ralph, Brandon CW, David R. Thomson, James Allan Cheyne, and Daniel Smilek. 2014. Media Multitasking and Failures of Attention in Everyday Life. *Psychological Research* 78 (5): 661–669.

Ralph, Brandon CW, David R. Thomson, Paul Seli, Jonathan SA Carriere, and Daniel Smilek. 2015. Media Multitasking and Behavioral Measures of Sustained Attention. *Attention, Perception, & Psychophysics* 77 (2): 390–401.

Reuver, Mark de, and Harry Bouwman. 2015. Dealing with Self-Report Bias in Mobile Internet Acceptance and Usage Studies. *Information & Management* 52 (3): 287–94. doi:10.1016/j.im.2014.12.002.

Rideout, Victoria J., Ulla G. Foehr, and Donald F. Roberts. 2010. *Generation M2: Media in the Lives of 8-to 18-Year-Olds*. Henry J. Kaiser Family Foundation. https://kaiserfamilyfoundation.files.wordpress.com/2013/04/8010.pdf (July 15 2017).

Rivron, Vassili, Mohammad Irfan Khan, Simon Charneau, and Isabelle Chrisment. 2015. Refining Smartphone Usage Analysis by Combining Crowdsensing and Survey. *2015 IEEE International Conference on Pervasive Computing and Communication Workshops (PerCom Workshops)*. doi: 10.1109/PERCOMW.2015.7134065.

Roberti, Jonathan W. 2004. A Review of Behavioral and Biological Correlates of Sensation Seeking. *Journal of Research in Personality* 38 (3): 256–79. doi:10.1016/S0092-6566(03)00067-9.

Roberts, Donald F., Ulla G. Foehr, and Victoria J. Rideout. 2005. *Generation M: Media in the Lives of 8-18 Year-Olds*. Henry J. Kaiser Family Foundation. https://kaiserfamilyfoundation.files.wordpress.com/2013/01/generation-m-media-in-the-lives-of-8-18-year-olds-report.pdf (July 15 2017).

Salvucci, Dario D, and Niels A Taatgen. 2011. *The Multitasking Mind*. New York City, NY: Oxford University Press.

Samsonovich, Alexei V. 2010. Toward a Unified Catalog of Implemented Cognitive Architectures. *Proceedings of the 2010 Conference on Biologically Inspired Cognitive Architectures 2010: Proceedings of the First Annual Meeting of the BICA Society*, 195–244. Amsterdam, The Netherlands, The Netherlands: IOS Press. http://dl.acm.org/citation.cfm?id=1893313.1893352.

Sanbonmatsu, David M., David L. Strayer, Nathan Medeiros-Ward, and Jason M. Watson. 2013. Who Multi-Tasks and Why? Multi-Tasking Ability, Perceived Multi-Tasking Ability, Impulsivity, and Sensation Seeking. *PLoS ONE* 8 (1): e54402. doi:10.1371/journal.pone.0054402.

Schober, Michael F., Frederick Conrad, Christopher Antoun, Patrick Ehlen, Stefanie Fail, Andrew L. Hupp, Michael Johnston, Lucas Vickers, H. Yanna Yan, and Chan Zhang. 2015. Precision and Disclosure in Text and Voice Interviews on Smartphones. *PLoS ONE* 10 (6): e0128337. doi:10.1371/journal.pone.0128337.

Sendelbah, Anže. 2015. Investigating Respondent Multitasking in Web Surveys with Paradata. Presented at the *annual conference of American Association of Public Opinion Research*, Hollywood, Florida, May 2015.

Sendelbah, Anže, Vasja Vehovar, Ana Slavec, and Andraž Petrovčič. 2016. Investigating Respondent Multitasking in Web Surveys Using Paradata. *Computers in Human Behavior* 55, Part B (February): 777–87. doi:10.1016/j.chb.2015.10.028.

Smith, Aaron, and Jan Lauren Boyles. 2012. *The Rise of the Connected Viewer*. http://wwww.pewinternet.com/~/media/Files/Reports/2012/PIP_Connected_Viewers.pdf (July 15 2017).

Spink, Amanda, Charles Cole, and Mary Waller. 2008. Multitasking Behavior. *Annual Review of Information Science and Technology* 42 (1): 93–118. doi:10.1002/aris.2008.1440420110.

Spink, Amanda, Minsoo Park, Bernard J. Jansen, and Jan Pedersen. 2006. Multitasking during Web Search Sessions. *Information Processing & Management* 42 (1): 264–75. doi:10.1016/j.ipm.2004.10.004.

Srivastava, Jaideep, Robert Cooley, Mukund Deshpande, and Pang-Ning Tan. 2000. Web Usage Mining: Discovery and Applications of Usage Patterns from Web Data. *ACM SIGKDD Explorations Newsletter* 1 (2): 12–23.

Srivastava, Jatin. 2010. *Media Multitasking and Role of Visual Hierarchy and Formatting Cues in Processing of Web Content*. The Ohio State University. https://etd.ohiolink.edu/ap/10?0::NO:10:P10_ACCESSION_NUM:osu1281718384 (July 15 2017).

Stieger, Stefan, and Ulf-Dietrich Reips. 2010. What Are Participants Doing While Filling in an Online Questionnaire: A Paradata Collection Tool

and an Empirical Study. *Computers in Human Behavior* 26 (6): 1488–95. doi:10.1016/j.chb.2010.05.013.

Stone, Arthur A., Christine A. Bachrach, Jared B. Jobe, Howard S. Kurtzman, and Virginia S. Cain. 1999. *The Science of Self-Report: Implications for Research and Practice*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sweller, John. 2012. Human Cognitive Architecture: Why Some Instructional Procedures Work and Others Do Not. In *APA Educational Psychology Handbook, Vol 1: Theories, Constructs, and Critical Issues*, ed. by K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, and J. Sweller, 295–325. Washington, DC, US: American Psychological Association. doi:10.1037/13273-011.

Thomas, R. K., and F.M Barlas. 2014. Respondents Playing Fast and Loose?: Antecedents and Consequences of Respondent Speed of Completion. Presented at the *annual conference of American Association of Public Opinion Research*, Anaheim, California, May 2014. http://www.aapor.org/AAPOR_Main/media/AnnualMeetingProceedings/2014/Session-C-8-2-Thomas-Y.pdf (July 15 2017).

Timotijevic, Lada, Julie Barnett, Richard Shepherd, and Victoria Senior. 2009. Factors Influencing Self-Report of Mobile Phone Use: The Role of Response Prompt, Time Reference and Mobile Phone Use in Recall. *Applied Cognitive Psychology* 23 (5): 664–683.

Tokan, Fatih, Pekka Mattila, and Elina Koivisto. 2011. *Media as Multitasking*. http://www.aaltomediamark.org/wp-content/uploads/2011/04/Media-as-Multitasking-NextMedia.pdf (May 1 2013).

Tourangeau, Roger, Lance J. Rips, and Kenneth Rasinski. 2000. *The Psychology of Survey Response*. Cambridge: Cambridge University Press doi:10.1017/CBO9780511819322.

Tran, Phuoc, Rogelio Carrillo, and Kaveri Subrahmanyam. 2013. Effects of Online Multitasking on Reading Comprehension of Expository Text. *Cyberpsychology: Journal of Psychosocial Research on Cyberspace* 7 (3). doi:10.5817/CP2013-3-2.

Uncapher, Melina R., Monica K. Thieu, and Anthony D. Wagner. 2016. Media Multitasking and Memory: Differences in Working Memory and Long-Term Memory. *Psychonomic Bulletin & Review* 23 (2): 483–490.

Van Cauwenberge, Anna, Gabi Schaap, and Rob van Roy. 2014. TV No Longer Commands Our Full Attention: Effects of Second-Screen Viewing and Task Relevance on Cognitive Load and Learning from News. *Computers in Human Behavior* 38 (September): 100–109. doi:10.1016/j.chb.2014.05.021.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S*. New York: Springer. http://www.stats.ox.ac.uk/pub/MASS4.

Voorveld, Hilde A. M., and Margot van der Goot. 2013. Age Differences in Media Multitasking: A Diary Study. *Journal of Broadcasting &*

Electronic Media 57 (3): 392–408. doi:10.1080/08838151.2013.816709.

Voorveld, Hilde A. M., Claire M. Segijn, Paul E. Ketelaar, and Edith G. Smit. 2014. Investigating the Prevalence and Predictors of Media Multitasking across Countries. *International Journal of Communication* 8 (0): 23.

Wallis, Claudia. 2006. The Multitasking Generation. *Time Magazine*, March 27 2006. http://content.time.com/time/magazine/article/0,9171,1174696,00.html (July 15 2017)

———. 2010. *The Impacts of Media Multitasking on Children's Learning and Development: Report from a Research Seminar*. http://multitasking.stanford.edu/MM_FinalReport_030510.pdf (June 15 2017).

Wang, Zheng, Prabu David, Jatin Srivastava, Stacie Powers, Christine Brady, Jonathan D'Angelo, and Jennifer Moreland. 2012. Behavioral Performance and Visual Attention in Communication Multitasking: A Comparison between Instant Messaging and Online Voice Chat. *Computers in Human Behavior* 28 (3): 968–975. doi:10.1016/j.chb.2011.12.018.

Wang, Zheng, and John M. Tchernev. 2012. The Myth of Media Multitasking: Reciprocal Dynamics of Media Multitasking, Personal Needs, and Gratifications. *Journal of Communication* 62 (3): 493–513. doi:10.1111/j.1460-2466.2012.01641.x.

WEBDATANET. 2010. *Memorandum of Understanding for the Implementation of a European Concerted Research Action Designated as COST Action IS1004: WEBDATANET: Web- Based Data-Collection - Methodological Challenges, Solutions and Implementations*. http://webdatanet.cbs.dk/images/presentations/mou%20is1004-webdatanet.pdf (April 9 2014).

Weiss, Gary M. 2005. Data Mining in Telecommunications. In *Data Mining and Knowledge Discovery Handbook*, ed. by Oded Maimon and Lior Rokach, 1189–1201. New York City, NY: Springer. doi:10.1007/0-387-25465-X_56.

Welch, Chris. 2014. Samsung Goes Back to Trashing Apple with Latest Commercials. *The Verge*, February 19. https://www.theverge.com/2014/2/19/5425888/samsung-back-to-trashing-apple-with-galaxy-tab-pro-commercial (June 15 2017).

West, Brady T. 2011. Paradata in Survey Research. *Survey Practice* 4 (4). http://surveypractice.org/index.php/SurveyPractice/article/view/112 (June 15 2017).

Weyandt, Kai. 2014. Using Longitudinal Paradata to Explain Page-Level Response Times in Web Surveys: A Multi-Level Approach. University of Mannheim.

Wickens, Christopher D. 2008. Multiple Resources and Mental Workload. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 50 (3): 449–455. doi:10.1518/001872008X288394.

Wood, Eileen, Lucia Zivcakova, Petrice Gentile, Karin Archer, Domenica De Pasquale, and Amanda Nosko. 2012. Examining the Impact of Off-Task Multi-Tasking with Technology on Real-Time Classroom Learning. *Computers & Education* 58 (1): 365–74. doi:10.1016/j.compedu.2011.08.029.

Yan, Ting, and Roger Tourangeau. 2008. Fast Times and Easy Questions: The Effects of Age, Experience and Question Complexity on Web Survey Response Times. *Applied Cognitive Psychology* 22 (1): 51–68.

Yeykelis, Leo, James J. Cummings, and Byron Reeves. 2014. Multitasking on a Single Device: Arousal and the Frequency, Anticipation, and Prediction of Switching between Media Content on a Computer. *Journal of Communication* 64 (1): 167–92. doi:10.1111/jcom.12070.

Zacher, Hannes, Holly A. Brailsford, and Stacey L. Parker. 2014. Micro-Breaks Matter: A Diary Study on the Effects of Energy Management Strategies on Occupational Well-Being. *Journal of Vocational Behavior* 85 (3): 287–97. doi:10.1016/j.jvb.2014.08.005.

Zhang, Chan, and Frederick Conrad. 2014. Speeding in Web Surveys: The Tendency to Answer Very Fast and Its Association with Straightlining. *Survey Research Methods* 8 (2): 127–35.

Zhang, Tongda, Xiao Sun, Yueting Chai, and Hamid Aghajan. 2015. A Look at Task-Switching and Multi-Tasking Behaviors: From the Perspective of the Computer Usage among a Large Number of People. *Computers in Human Behavior* 49 (August): 237–44. doi:10.1016/j.chb.2015.03.012.

Zhang, Weiyu, Se-Hoon Jeong, and Martin Fishbein. 2010. Situational Factors Competing for Attention: The Interaction Effect of Multitasking and Sexually Explicit Content on TV Recognition. *Journal of Media Psychology* 22 (1): 2–13. doi:10.1027/1864-1105/a000002.

Zimbardo, Philip G., and John N. Boyd. 2015. Putting Time in Perspective: A Valid, Reliable Individual-Differences Metric. In *Time Perspective Theory; Review, Research and Application*, ed. by Maciej Stolarski, Nicolas Fieulaine, and Wessel van Beek, 17–55. New York, NY: Springer. doi:10.1007/978-3-319-07368-2_2.

Zwarun, Lara, and Alice Hall. 2014. What's Going on? Age, Distraction, and Multitasking during Online Survey Taking. *Computers in Human Behavior* 41 (December): 236–44. doi:10.1016/j.chb.2014.09.041.

Subject index

Author index

Appendices

Appendix A Web questionnaire

Q2 - V nekaterih primerih želimo raziskovalci bolj poglobljeno razumeti določen problem in zato izvedemo spletne pogovore (chat), kjer imajo anketiranci priložnost izraziti svoja mnenja in predstaviti svoje osebne izkušnje. Spletni pogovor je anonimen in traja približno 30 minut. Vsak anketiranec, ki bi sodeloval v takšnem pogovoru, bi prejel 10 € nagrade. V spletnem pogovoru bi lahko sodelovali takoj po zaključku ankete ali pa kasneje, v času, ki bi ga predlagali sami.

Q3 - Ali bi vi osebno bili pripravljeni sodelovati v takšnem spletnemupogovoru?

◯ Da
◯ Ne

Q4 - Če bi vas povabili k sodelovanju, kdaj bi bili pripravljeni sodelovati v spletnem pogovoru?

☐ V spletnem pogovoru bi bil pripravljen sodelovati kar takoj, s klikom na povezavo, ki bi vodila v spletni pogovor.
☐ V spletnem pogovoru bi sodeloval kasneje, v času, ki bi ga predlagal sam.

Q5 - Ali vi osebno posedujete naslednje naprave ali pa jih souporabljate z drugimi člani vašega gospodinjstva?

	Da	Ne	Ne vem
Mobilni telefon	◯	◯	◯
E-bralnik (npr. Kindle, Nook)	◯	◯	◯
Tablični računalnik (npr. iPad, Samsung Galaxy Tab, Google Nexus, Kindle Fire)	◯	◯	◯
Namizni ali prenosni računalnik	◯	◯	◯
Igralna konzola (npr. Xbox, Play Station)	◯	◯	◯
Prenosni MP3 predvajalnik (npr. iPod)	◯	◯	◯
Prenosna igralna konzola (npr. PSP ali Sega Genesis)	◯	◯	◯
Televizor	◯	◯	◯

Q6 - Ali imate "pametni telefon" - mobilni telefon z naprednimi funkcionalnostmi, npr. iPhone, Android, Blackberry ali Windows telefon.

◯ Da
◯ Ne
◯ Ne vem

Q7 - Naslednja tri vprašanja se nanašajo na vašo uporabo MOBILNEGA TELEFONA. Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje uporabljate **glasovno ali video klicanje** po mobilnem telefonu za pogovore o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q8 - Ocenite, kako pogosto se doma, v službi, v šoli ali kje drugje komunicirate prek SMS ali MMS s pomočjo mobilnega telefona o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q9 - Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje preko interneta na mobilnem telefonu izmenjujete **besedilna ali slikovna sporočila** za komunikacijo o…

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q7a - Med tipičnim tednom, koliko **ur** uporabljate internet?Pri tem ne upoštevajte časa, ki ga porabite za e-mail.

[ ]

Q7b - Med tipičnim tednom, koliko **ur** uporabljate internet?

[ ]

Q10 - Naslednji dve vprašanji se nanašata na vašo uporabo TABLIČNEGA RAČUNALNIKA oziroma TABLICE. Prosimo, da pri vaših odgovorih ne upoštevate vaše uporabe stacionarnih ali prenosnih računalnikov.Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje **preko interneta** na tablici izmenjujete **besedilna ali slikovna sporočila** za komunikacijo o…

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih	○	○	○	○	○	○

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
temah?						

Q11 - Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje uporabljate tablico za glasovne ali video klice (npr. Skype) za pogovore o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

IF (3) (Media)
IF (8) (IC: tablet + PC)
Q12 - Naslednja dve vprašanji se nanašata na vašo uporabo OSEBNEGA (STACIONARNEGA ali PRENOSNEGA) RAČUNALNIKA.Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje **preko interneta** na osebnem računalniku **izmenjujete besedilna ali slikovna sporočila** za komunikacijo o…

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q13 - Ocenite, kako pogosto doma, v službi, v šoli ali kje drugje uporabljate osebni računalnik za glasovne

ali video klice (npr. Skype) za pogovore o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q14 - Zamislite si vaš povprečen dan. Zanima nas, **koliko ur na dan porabite za naštete aktivnosti**.Prosimo, vnesite številko med 0 in 24, lahko uporabite tudi decimalna števila (npr. število 1.5 pomeni 1 uro in 30 minut).

	Število ur	Te dejavnosti ne izvajam na povprečen dan
Brskanje po spletu in branje spletnih vsebin		
Kratka sporočila (SMS), takojšnje sporočanje (instant messaging) ali elektronska pošta		
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)		
Pogovor preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)		
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)		
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)		
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti		

Q15 - Medtem ko brskate po spletu ali berete spletne vsebine, kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Brskanje ali branje drugih vsebin na spletu	○	○	○	○
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	○	○	○	○
Poslušanje glasbe, radia, podcastov ali	○	○	○	○

	Večino časa	Nekaj časa	Malo časa	Nikoli
drugih audio vsebin (npr. TV v ozadju)				
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	◯	◯	◯	◯
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	◯	◯	◯	◯
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	◯	◯	◯	◯
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	◯	◯	◯	◯

Q16 - Medtem ko pošiljate ali berete kratka sporočila (SMS), sporočila takojšnjega sporočanja (IM) ali e-pošto, kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Druga seansa kratkih sporočil, takojšnjega sporočanja ali e-pošte	◯	◯	◯	◯
Brskanje po spletu, branje spletnih novic	◯	◯	◯	◯
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)	◯	◯	◯	◯
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	◯	◯	◯	◯
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	◯	◯	◯	◯
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	◯	◯	◯	◯
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	◯	◯	◯	◯

Q17 - Medtem, ko poslušate glasbo, radio, podcaste ali druge audio vsebine (npr. TV v ozadju), kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Poslušanje drugih audio vsebin (glasbe, radia, podcastov...)	◯	◯	◯	◯
Brskanje po spletu, branje spletnih novic	◯	◯	◯	◯
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	◯	◯	◯	◯
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	◯	◯	◯	◯
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	◯	◯	◯	◯
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	◯	◯	◯	◯
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	◯	◯	◯	◯

Q18 - Medtem ko se pogovarjate po telefonu ali drugih napravah (vključuje video klepete, npr. Skype), kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Druga seansa pogovorjanja preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	◯	◯	◯	◯
Brskanje po spletu, branje spletnih vsebin	◯	◯	◯	◯
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	◯	◯	◯	◯
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)	◯	◯	◯	◯
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	◯	◯	◯	◯
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	◯	◯	◯	◯
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	◯	◯	◯	◯

Q19 - Medtem ko uporabljate spletna socialna omrežja (npr. Facebook, Twitter, itd.), kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Uporaba drugih spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	○	○	○	○
Brskanje po spletu, branje spletnih vsebin	○	○	○	○
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	○	○	○	○
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)	○	○	○	○
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	○	○	○	○
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	○	○	○	○
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	○	○	○	○

Q20 - Medtem ko gledate TV ali video vsebine (npr. filme, nadaljevanke, poročila, YouTube posnetke), kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Gledanje drugih TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	○	○	○	○
Brskanje po spletu, branje spletnih vsebin	○	○	○	○
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	○	○	○	○
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	○	○	○	○
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)	○	○	○	○
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	○	○	○	○
Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti	○	○	○	○

Q21 - Medtem ko delate na besedilnih dokumentih, prezentacijah, preglednicah ali podobnih aktivnostih, kako pogosto počnete še katero od drugih aktivnosti?

	Večino časa	Nekaj časa	Malo časa	Nikoli
Delo na drugih besedilnih dokumentih, prezentacijah, preglednicah ali podobnih aktivnostih	○	○	○	○
Brskanje po spletu, branje spletnih vsebin	○	○	○	○
Pošiljanje kratkih sporočil, takojšnje sporočanje ali pošiljanje elektronske pošte	○	○	○	○
Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)	○	○	○	○
Pogovorjanje preko telefona ali drugih naprav (vključuje video klepete, npr. Skype)	○	○	○	○
Uporaba spletnih socialnih omrežij (npr. Facebook, Twitter, itd.)	○	○	○	○
Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)	○	○	○	○

Q22 - Ocenite, kako pogosto se doma, v službi, v šoli ali kje drugje pogovarjate po stacionarnem telefonu o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q23 - Ocenite, kako pogosto se doma, v službi, v šoli ali kje drugje osebno (ko ste v živo fizično skupaj) pogovarjate o ...

	Nikoli	Nekajkrat na leto ali manj	Nekajkrat na mesec	Nekajkrat na teden	Vsak dan ali skoraj vsak dan	Večkrat na dan
službenih, poslovnih ali šolskih zadevah?	○	○	○	○	○	○
vsakdanjih praktičnih stvareh (dogovarjanja, informacije, opravki)?	○	○	○	○	○	○
neobveznih stvareh (klepetanje, zabava, druženje)?	○	○	○	○	○	○
zaupnih osebnih temah?	○	○	○	○	○	○

Q24 - Zaključili ste s sklopom vprašanj o vaši uporabi medijev. Nahajate se približno na polovici vprašalnika.Sedaj sledi nekaj vprašanj o vaših stališčih in različnih vidikih vašega življenja.

Q25 - Zame je najboljši/a prijatelj/ica tisti/a, ki ...

Q26 - Zame je dober/a prijatelj/ica tisti/a, ki ...

Q27 - Navedenih je nekaj osebnostnih značilnosti, ki morda veljajo za vas. Za vsako trditev označite, v kolikšni meri se strinjate z njo. Ocenite, koliko za vas velja posamezen par značilnosti, tudi če ena izmed značilnosti velja bolj od druge.

Q28 - Sebe vidim osebo, ki je ...

	Močno se ne strinjam	Ne strinjam se	Delno se ne strinjam	Niti - niti	Delno se strinjam	Strinjam se	Močno se strinjam
ekstravertirana, polna navdušenja.	○	○	○	○	○	○	○
kritična, prepirljiva.	○	○	○	○	○	○	○
zanesljiva, disciplinirana.	○	○	○	○	○	○	○
napeta, zlahka razdražljiva.	○	○	○	○	○	○	○
odprta za nove izkušnje, večplastna.	○	○	○	○	○	○	○
zadržana, tiha.	○	○	○	○	○	○	○
sočutna, topla.	○	○	○	○	○	○	○
neorganizirana, lahkomiselna.	○	○	○	○	○	○	○
mirna, čustveno stabilna.	○	○	○	○	○	○	○
preprosta, brez domišljije.	○	○	○	○	○	○	○

Q29 - Preberite vsako trditev in odgovorite, kako značilno je opisano vedenje za vas.

	Povsem neznačilno	Redko značilno	Enako pogosto značilno kot neznalično	Pogosto značilno	Povsem značilno
Podobe, zvoki, vonjave, ki me spominjajo na otroštvo, v meni pogosto vzbudijo lepe spomine.	○	○	○	○	○
Pogosto razmišljam o tem, kaj bi moral v življenju storiti drugače.	○	○	○	○	○
Uživam v razmišljanju o svoji preteklosti.	○	○	○	○	○
Kadar želim nekaj doseči, si postavim cilje in razmislim o sredstvih, s katerimi jih lahko dosežem.	○	○	○	○	○
Življenje bo pač prineslo svoje, ne glede na to, kaj počnem.	○	○	○	○	○
Veseli spomini na dobre čase mi zlahka prihajajo na misel.	○	○	○	○	○
Pomembno mi je, da v življenje vnašam	○	○	○	○	○

	Povsem neznačilno	Redko značilno	Enako pogosto značilno kot neznalično	Pogosto značilno	Povsem značilno
zabavo.					
Razmišljam o priložnostih, ki sem jih zamudil.	○	○	○	○	○
Nima smisla skrbeti o prihodnosti, saj nanjo tako ali tako nimam nobenega vpliva.	○	○	○	○	○
S stalnim in postopnim napredovanjem mi uspe pravočasno zaključevati projekte.	○	○	○	○	○
Iščem tveganja, da bi prinesel vznemirjenje v svoje življenje.	○	○	○	○	○
Zmožen sem se upreti skušnjavam, če vem, da je nekaj treba postoriti.	○	○	○	○	○
Zgodi se mi, da me vznemirjenje povsem prevzame.	○	○	○	○	○
Pogosto razmišljam o slabih stvareh, ki so se mi zgodile.	○	○	○	○	○
Moje življenje nadzorujejo sile, na katere nimam vpliva.	○	○	○	○	○
Preverjamo vašo pozornost, prosimo označite odgovor povsem značilno.	○	○	○	○	○

Q28a - Zdaj bi vas radi vprašali o zmernih fizičnih aktivnostih, kot so hitra hoja, kolesarjenje, sesanje, delo na vrtu in druge aktivnosti, ki povzročijo manjši porast bitja srca in dihanja.Na dneve, ko izvajate vsaj 10 minut takšnih zmernih aktivnostih na dan, koliko celotnega časa v minutah na dan porabite za takšne aktivnosti?

[ ]

Q28b - Zdaj bi vas radi vprašali o zmernih fizičnih aktivnostih, kot so hitra hoja, kolesarjenje, sesanje, delo na vrtu in druge aktivnosti, ki povzročijo manjši porast bitja srca in dihanja.Na dneve, ko izvajate vsaj 10 minut takšnih zmernih aktivnostih na dan, koliko celotnega časa v minutah na dan porabite za takšne aktivnosti?Pri tem prosim ne upoštevajte aktivnosti, ki jih izvajate na delu ali na poti na delo.

[ ]

Q30 - Kolikokrat ste v zadnjih 14 dneh bili <u>neprekinjeno</u> fizično aktivni vsaj 20 minut ali več?
Upoštevajte tudi dela v gospodinjstvu ali na vrtu, v kolikor ste jih neprekinjeno izvajali vsaj 20 minut ali več.

- ◯ Noben dan
- ◯ En dan
- ◯ Dva dni
- ◯ Tri dni
- ◯ Štiri dni
- ◯ Pet dni
- ◯ Šest dni
- ◯ Sedem dni
- ◯ Osem dni
- ◯ Devet dni
- ◯ Deset dni
- ◯ Enajst dni
- ◯ Dvanajst dni
- ◯ Trinajst dni
- ◯ Štirinajst dni

Q31 - Katera možnost najbolje opiše, kako ste odgovorili na prejšnje vprašanje o fizičnih aktivnostih v zadnjih 14-ih dneh?

- ◯ Vedel/a sem odgovor na pamet
- ◯ Razmislil/a sem o fizičnih aktivnostih po posameznih dneh in jih seštel/a
- ◯ Ocenil/a sem glede na moj splošni vtis
- ◯ Drugo:

Q32 - Pred vami se nahaja še nekaj vprašanj o vaših izkušnjah z anketami.

Q33 - Poleg askGfK, ali sodelujete še v kakšnih drugih spletnih panelih?

- ◯ Ne
- ◯ Da, sodelujem še v enem spletnem panelu
- ◯ Da, sodelujem še v dveh spletnih panelih
- ◯ Da, sodelujem še v treh ali več spletnih panelih

Q34 - Ali lahko ocenite, kolikokrat ste bili v zadnjih 12 mesecih kontaktirani oziroma osebno povabljeni po pošti, telefonu, e-pošti ali v osebnem srečanju...Pri tem ne upoštevajte svojega sodelovanja v askGfK in morebitnih drugih spletnih panelih.

	Število kontaktov	Kolikokrat od tega ste v anketi tudi sodelovali?
za sodelovanje v osebnem anketiranju		
za sodelovanje v anketi po fiksnem telefonu		
za sodelovanje v anketi po mobilnem telefonu		
za sodelovanje v anketi po internetu		

Q35 - Poleg osebnih povabil, kolikokrat ste v zadnjih 12 mesecih na internetu opazili tudi splošna vabila za sodelovanje v spletnih anketah (npr. oglasi na spletnih straneh, objave na Facebooku ipd.)?

Število opaženih vabil	
Kolikokrat od tega ste v anketi tudi sodelovali?	

Q36 - Kolikokrat pa so vas v zadnjih 12 mesecih nepovabljeno osebno kontaktirali in vam prodajali storitve ali blago...

	Število kontaktov	Kolikokrat od tega ste pri tem tudi sodelovali oziroma se odzvali na ponudbo (npr. naročili, kupili, včlanili)?
po telefonu (telemarketing)		
osebno (direktno prodajanje na vaših vratih)		
po e-pošti v slovenščini		

Q37 - Ljudje sodelujejo v spletnih panelih iz različnih razlogov.Prosimo, označite koliko so za vaše sodelovanje v spletnih panelih pomembni spodnji razlogi.

	Sploh ni pomembno	Ni pomembno	Niti ni pomembno, niti nepomembno	Pomembno	Zelo pomembno
Denar	○	○	○	○	○
Praktične nagrade	○	○	○	○	○
Možnost izražanja mnenj	○	○	○	○	○
Možnost vplivanja na rezultate	○	○	○	○	○
Radovednost	○	○	○	○	○
Darovanje v dobrodelne namene	○	○	○	○	○
Loterija	○	○	○	○	○
Ogled rezultatov	○	○	○	○	○

Q38 - Zanimajo nas vaši vtisi o današnji anketi.

Q39 - Prosimo, ocenite današnjo anketo. 1 predstavlja najnižjo, 10 pa najvišjo oceno.

○1 ○2 ○3 ○4 ○5 ○6 ○7 ○3 ○9 ○10

Q40 - Kako bi vi ocenili današnjo anketo?

	Močno se ne strinjam	Se ne strinjam	Delno se ne strinjam	Niti - niti	Delno se strinjam	Se strinjam	Močno se strinjam
Anketa je bila drugačna od ostalih	○	○	○	○	○	○	○
Anketa je bila enostavna za izpolnjevanje	○	○	○	○	○	○	○
Anketa je bila zanimiva	○	○	○	○	○	○	○
Anketa je bila predolga	○	○	○	○	○	○	○

Q41 - Za naše raziskovanje je ključnega pomena, da v analizo vključimo le odgovore anketirancev, ki so tej anketi posvetili vso svojo pozornost. Sicer bi lahko leta truda (raziskovalcev in ostalih anketirancev) bila zapravljena. *GfK točke boste prejeli v vsakem primeru,* vendar vas kljub temu prosimo, da nam poveste, koliko truda ste namenili tej anketi.

○ Skoraj nič truda
○ Zelo malo truda
○ Nekaj truda
○ Kar veliko truda
○ Zelo veliko truda

Q42 - Pogosto so med izpolnjevanjem spletnih vprašalnikov prisotne motnje (npr. drugi ljudje, televizija, glasba ipd.). Prosimo označite koliko pozornosti ste posvetili tej anketi. Ponovno poudarjamo, da boste prejeli GfK točke ne glede na vaš odgovor. Cenimo vašo iskrenost!

○ Skoraj nič pozornosti
○ Zelo malo pozornosti
○ Nekaj pozornosti
○ Večino pozornosti
○ Vso pozornost

Q43 - Kaj ste počeli, če sploh kaj, na katerikoli napravi med izpolnjevanjem te ankete?

☐ Brskanje po spletu, branje spletnih novic, branje e-knjig
☐ Kratka sporočila (SMS), takojšnje sporočanje (instant messaging) ali elektronska pošta
☐ Poslušanje glasbe, radia, podcastov ali drugih audio vsebin (npr. TV v ozadju)
☐ Pogovor preko telefona ali druge naprave (vključuje video klepete, npr. Skype)
☐ Igranje iger (računalnik, video, splet)
☐ Uporaba socialnih omrežij (npr. Facebook, Twitter, itd.)
☐ Gledanje TV ali video vsebin (npr. filmi, nadaljevanke, poročila, YouTube posnetki)
☐ Delo na besedilnih dokumentih, prezentacijah, preglednicah ali podobne aktivnosti
☐ Drugo
☐ Nisem bil zaposlen z nobeno drugo aktivnostjo na katerikoli napravi.

Q44 - Poleg že naštetega v prejšnjem vprašanju, kaj ste še počeli, če sploh kaj, med izpolnjevanjem te ankete?

☐ Hranjenje, pitje ali priprava obroka

☐ Gospodinjska opravila (čiščenje, pomivanje posode, itd.)
☐ Skrb za druge ljudi (npr. otroke)
☐ Pogovarjanje z drugo osebo iz oči v oči
☐ Poslušanje govora druge osebe (npr. biti prisoten na predavanju)
☐ Nakupovanje in podobna opravila (npr. banka, pošta)
☐ Hoja (npr. sprehajanje)
☐ Uporabljanje prevoznih sredstev (npr. avto, avtobus, vlak)
☐ Drugo:
☐ Nisem bil zaposlen z nobeno drugo aktivnostjo.

Q45 - Kolikokrat vas je, medtem ko ste izpolnjevali to anketo, nekaj prekinilo za več kot 5 sekund?

	0 krat	1 krat	2 krat	3 krat	4 krat	5-9 krat	10+ krat
Prekinitve za aktivnosti na tej napravi	○	○	○	○	○	○	○
Prekinitve za aktivnosti na drugih napravah (npr. TV, tablica, telefon)	○	○	○	○	○	○	○
Druge prekinitve, ki niso vezane na neko elektronsko napravo (npr. gospodinjska opravila, pogovor z drugo osebo iz oči v oči ipd.)	○	○	○	○	○	○	○

Q49 - Sledi še na nekaj sociodemografskih vprašanj.

Q50 - Najvišja stopnja dosežene izobrazbe:

○ Osnovna šola ali manj.
○ Dveletna ali triletna poklicna šola.
○ Štiriletna ali petletna srednja šola.
○ Višja, visoka šola,
○ Univerzitetna izobrazba.
○ Specializacija, magisterij, doktorat.

Q51 - Kako bi opisali svoj sedanji status?

○ Zaposleni
○ Samozaposleni
○ Kmet
○ Druga oblika dela (pogodba o delu, avtorska pogodba, priložnostno delo)
○ Pomagajoči družinski član
○ Brezposelni
○ Učenec, dijak, študent
○ Upokojenec
○ Gospodinjska dela
○ Nezmožen za delo
○ Drug status

Q52 - Koliko oseb, vključno z vami, živi v vašem gospodinjstvu?

Q53 - Koliko oseb vašega gospodinjstva je starih 16 let ali več?

Q54 - V katerega od spodnjih razredov bi uvrstili skupni neto dohodek vašega gospodinjstva?

- ◯ Do vključno 400 evrov
- ◯ Do vključno 700 evrov
- ◯ Do vključno 1100 evrov
- ◯ Do vključno 1500 evrov
- ◯ Do vključno 1800 evrov
- ◯ Do vključno 2100 evrov
- ◯ Do vključno 2500 evrov
- ◯ Do vključno 3000 evrov
- ◯ Do vključno 4000 evrov
- ◯ Več kot 4000 evrov
- ◯ Ne želim odgovoriti

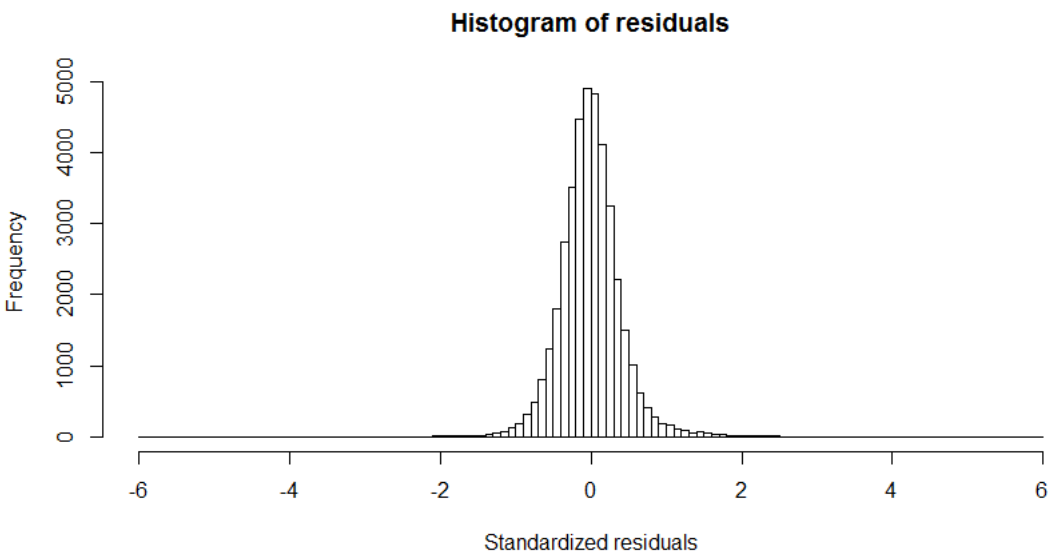
Appendix B Linear mixed model for time-out events

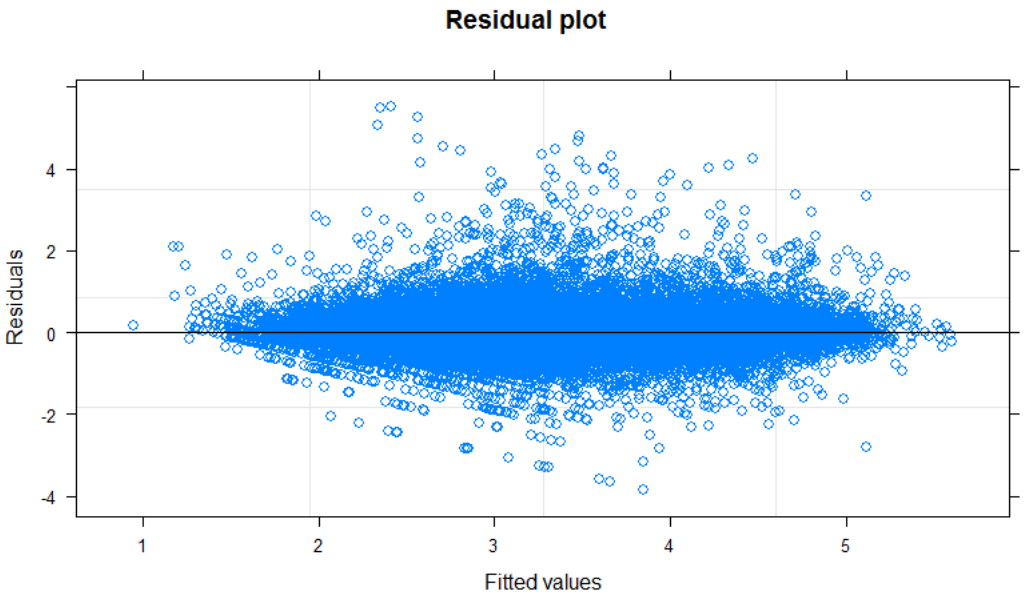
In total, 40,264 units (i.e. page response times of individuals) are included in this analysis. The model is specified as $y_{ij} = \beta_0 + u_i + u_j + \varepsilon_{ij}$, where:

- y_{ij} denotes a log-transformed response time of individual *i* on page *j*
- u_i is a random effect of individual *i*
- u_j is a random effect of page *j*
- ε_{ij} is a residual for individual *i* on page *j*

Variance of random effects for individuals is 0.10, for random effects for pages it is 0.38, while the variance of standardized residuals is 0.22.

Bellow I present the residual plot and histogram of standardized residuals. Overall, 2.21% of units had residuals with value over one.




Residual plot
Residuals
Fitted values
4
2
0
-2
-4
1
2
3
4
5

Appendix C Additional tables from analyses

Appendix C.1 Models summarized in Table 24

Appendix C.1.1 Models with paradata indicators

	Estim. behav. freq.	Trap item	Straight-lining	Incon-sistent resp-onses	Short open-ended answer	Any	Any without estim. & incons.
Intercept	-0.54	1.30*	-0.14	-2.94	-1.45	2.10‡	1.47*
P_FOCUS	-0.09	0.05	0.09	-0.33	-0.15	-0.07	0.02
P_TIME	-0.11	0.26†	0.36*	-0.10	0.09	0.03	0.26†
Age	-0.01*	0.01	0.00	-0.01	0.03*	-0.01	0.01
Gender: female	-0.04	-0.23	-1.17†	0.00	-1.40‡	-0.08	-0.16
Education: middle	0.17	-0.42	-1.18*	-0.06	-0.35	-0.11	-0.31
Education: high	0.30	-0.52*	-1.06*	-0.84	-0.03	-0.16	-0.48
Effort	-0.02	-0.20	-0.24	-0.10	-0.21	-0.15	-0.21
Attention	0.01	-0.57‡	-0.49†	0.06	-0.32*	-0.30‡	-0.54‡
Time (first level coef.)	6.30*	-15.40‡	-20.75†	5.48	-8.16	-1.07	-13.53‡
Time (sec. level coef.	-1.73	7.77†	21.16‡	11.63	0.21	3.56	8.00†

* $p < 0.05$; † $p < 0.01$; ‡ $p < 0.001$

Appendix C.1.2 Models with indicators on self-reported ranges of secondary activity forms and interruptions

	Estim. behav. freq.	Trap item	Straight-lining	Incon-sistent resp-onses	Short open-ended answer	Any	Any without estim. & incons.
Intercept	-0.71	1.46*	-1.01	-4.63*	-1.35	1.85†	1.42*
SR_RM_MM	-0.07	-0.07	-0.10	0.66*	-0.06	-0.01	0.01
SR_RM_NM	-0.01	-0.02	0.24	0.37	0.19	-0.05	-0.03
S_SEQ_SD	-0.04	0.09	0.24	-0.15	-0.04	0.02	0.08
S_SEQ_OD	0.03	0.17	0.20	-0.59	0.06	0.05	0.05
S_SEQ_NM	0.01	0.09	0.00	-0.58	-0.44*	-0.01	0.09
Age	-0.01	0.01	0.03	0.01	0.03*	-0.01	0.01
Gender: female	-0.03	-0.22	-1.39†	-0.05	-1.41‡	-0.05	-0.14
Education: middle	0.15	-0.38	-1.35*	-0.14	-0.27	-0.16	-0.33
Education: high	0.29	-0.51*	-0.86	-0.92	0.02	-0.24	-0.53*
Effort	-0.03	-0.21*	-0.24	-0.06	-0.21	-0.14	-0.20
Attention	0.02	-0.58‡	-0.57†	0.22	-0.36*	-0.29‡	-0.54‡
Time (first level coef.)	2.92	-12.92‡	-22.64†	0.47	-6.97	-2.33	-10.81‡
Time (sec. level coef.	-1.81	7.52†	20.97‡	10.71	0.46	2.78	6.95*

* $p < 0.05$; † $p < 0.01$; ‡ $p < 0.001$

Appendix C.1.3 Models with indicators on clusters of self-reported secondary activity forms

	Estim. behav. freq.	Trap item	Straight-lining	Incon-sistent resp-onses	Short open-ended answer	Any	Any without estim. & incons.
Intercept	-1.19*	1.98†	-0.45	-5.64†	-0.99	1.99‡	2.08†
Common mixers	-0.25	-0.44	-0.61	0.67	-1.55	-0.48	-0.46
Complex mixers	0.13	0.22	0.12	0.87	-0.05	0.09	0.17
Listeners	0.34	-0.60*	-0.32	1.60†	-0.36	-0.03	-0.48
F2F talkers	0.33	-0.31	0.81	1.86*	-0.74	0.01	-0.45
Watchers	0.50	-0.20	0.85	1.22	-0.41	-0.11	-0.91
Eaters	-0.04	-0.53	-15.05	0.26	-15.71	-0.28	-0.76
Phone talkers	0.34	-1.00	0.79	1.23	-0.93	-0.27	-0.91
Age	-0.01	0.00	0.02	0.01	0.03*	-0.01	0.00
Gender: female	-0.02	-0.27	-1.42†	0.03	-1.42‡	-0.07	-0.18
Education: middle	0.13	-0.33	-1.42*	-0.26	-0.24	-0.13	-0.26
Education: high	0.25	-0.44	-0.87	-1.14	0.07	-0.22	-0.46
Effort	-0.02	-0.18	-0.19	-0.13	-0.22	-0.14	-0.19
Attention	0.08	-0.64‡	-0.57†	0.40	-0.40†	-0.31‡	-0.61‡
Time (first level coef.)	1.99	-8.59†	-18.33†	-6.17	-7.19	-1.36	-6.85*
Time (sec. level coef.	-1.59	7.36†	22.35‡	12.97*	-1.46	2.64	6.31*

* $p < 0.05$; † $p < 0.01$; ‡ $p < 0.001$

Appendix C.1.4 Models with indicator on high media multitasking

	Estim. behav. freq.	Trap item	Straight-lining	Incon-sistent resp-onses	Short open-ended answer	Any	Any without estim. & incons.
Intercept	-0.66	1.20	-0.28	-1.04	-1.78	2.34‡	1.21
HMM	-0.47*	0.87‡	1.10†	0.20	0.26	0.22	0.77‡
Age	-0.01*	0.01	0.02	0.00	0.04*	-0.01	0.02*
Gender: female	-0.02	-0.25	-1.17†	0.27*	-1.39‡	0.01	-0.42†
Education: middle	0.17	-0.42	-1.13*	-0.53*	-0.37	-0.18	-0.29
Education: high	0.28	-0.51*	-1.03*	-0.61†	-0.06	-0.28	-0.28
Effort	-0.02	-0.19	-0.19	0.05	-0.21	-0.13	-0.18
Attention	0.03	-0.60‡	-0.59‡	0.04	-0.32*	-0.25†	-0.57‡
Time (first level coef.)	1.95	-8.87†	-12.49*	-4.06	-10.03	-2.36	-9.38†
Time (sec. level coef.	-1.43	7.10*	20.33‡	3.18	-0.52	2.71	3.56

* $p < 0.05$; † $p < 0.01$; ‡ $p < 0.001$

Appendix C.2 Models on length of non-differentiation

Appendix C.2.1 Models with indicators on self-reported ranges of secondary activity forms and interruptions

	Null model	+ Self-reports	+ MMI	+ Controls
Intercept	-1.36***	-1.38***	-1.41***	-0.39
SR_RM_MM		0.04	0.03	0.02
SR_RM_NM		-0.01	-0.01	-0.02
S_SEQ_SD		0.04	0.04	0.05
S_SEQ_OD		-0.01	0.00	0.03
S_SEQ_NM		-0.04	-0.04	-0.02
HMM			0.22**	0.24***
Age				0.01***
Gender: female				0.00
Education: middle				-0.43***
Education: high				-0.44***
Effort				-0.07*
Attention				-0.20***
Time (first level coef.)				-6.07***
Time (sec. level coef.)				4.80***
2 * log likelihood	1.529.2		1,445.2	1.728.6
Pseudo R^2		0.0058	0.0149	0.1399
AIC	-1,528.1	-1.521.0	-1,529.2	-1,696.7

[a] the decrease is statistically significant according to the likelihood ratio test

* p < 0.05; ** p < 0.01; *** p < 0.00

Appendix C.2.1 Models with indicators on clusters of self-reported secondary activity forms

	Null model	+ Self-reports	+ MMI	+ Controls
Intercept	-1.36***	-1.35***	-1.38***	-0.28
Common mixers		-0.11	-0.12	-0.04
Complex mixers		0.09	0.07	0.05
Listeners		-0.11	-0.13	-0.15*
F2F talkers		-0.04	-0.02	0.02
Watchers		0.12	0.14	0.14
Phone talkers		-0.11	-0.15	-0.04
Age		-0.22	-0.24	-0.17
HMM			0.24***	0.25***
Age				0.01***
Gender: female				-0.01
Education: middle				-0.42***
Education: high				-0.43***
Effort				-0.07*
Attention				-0.21***
Time (first level coef.)				-5.52***
Time (sec. level coef.)				4.98***
2 * log likelihood	1529.2	1,538.0	1,550.4[a]	1,790.2[a]
Pseudo R^2		0.0089	0.0197	0.1404
AIC	-1,528.1	-1,520.0	-1,530.4	-1,694.3

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

Appendix C.3 Models on length of open-ended responses

Appendix C.3.1 Models with paradata indicators

	Null model	+ Self-reports	+ MMI	+ Controls
Intercept	2.71***	2.68***	2.61***	2.37***
P_FOCUS		0.02**	0.01*	0.00
P_TIME		0.02	0.02	-0.03
HMM			0.13***	0.07*
Age				-0.01***
Gender: female				0.22***
Education: middle				0.08
Education: high				0.16**
Effort				0.05*
Attention				0.09***
Time (first level coef.)				4.92***
Time (sec. level coef.)				-2.66***
2 * log likelihood	-8,853.1	-8,845.4	-8,832.0	-8,608.9
Pseudo R^2		0.0009	0.0029	0.0276
AIC	8,857.1	8,853.4	8,842.0	8,634.9

[a] the decrease is statistically significant according to the likelihood ratio test

* p < 0.05; ** p < 0.01; *** p < 0.00

Appendix C.3.2 Models with indicators on self-reported ranges of secondary activity forms and interruptions

	Null model	+ Self-reports	+ MMI	+ Controls
Intercept	2.71***	2.67***	2.67***	2.35***
SR_RM_MM		0.09**	0.08**	0.08**
SR_RM_NM		0.04	0.04	0.01
S_SEQ_SD		-0.03	-0.03	-0.03
S_SEQ_OD		-0.01	-0.01	-0.03
Age		0.00	0.00	-0.01
HMM			0.02	0.00
Age				-0.01***
Gender: female				0.22***
Education: middle				0.06
Education: high				0.15**
Effort				0.06*
Attention				0.09***
Time (first level coef.)				4.36***
Time (sec. level coef.)				-2.51***
2 * log likelihood	-8,853.1	-8,840.3 [a]	-8,840.2	-8,605.9 [a]
Pseudo R^2		0.0014	0.0015	0.0280
AIC	8,857.1	8,854.3	8,856.2	8,637.9

[a] the decrease is statistically significant according to the likelihood ratio test

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.00$

Razširjeni povzetek v slovenskem jeziku

Večopravilnost se nanaša na hkratne ali zaporedne kombinacije vsaj dveh aktivnosti (Salvucci in Taatgen 2011). *Medijska večopravilnost* vključuje vsaj eno medijsko aktivnost (tj. prebiranje ali ustvarjanje medijskih vsebin in komunikacija prek elektronskih naprav) (Foehr 2006; Ophir in dr. 2009; Wallis 2010). Raziskave kažejo, da je medijska večopravilnost vedno pogostejša oblika obnašanja v vsakdanjem življenju (npr. Rideout in drugi 2010; Voorveld in van der Goot 2013; Ofcom 2016). Raziskave prav tako nakazujejo, da so določene oblike večopravilnosti povezane s slabšo kakovostjo izvedbe ciljno usmerjenih aktivnosti (Meyer in Kieras 1997; Ophir in dr. 2009; Salvucci in Taatgen 2011).

Medijska večopravilnost predstavlja dva splošna izziva za družboslovno znanost. Prvi izziv je, kako opredeliti in meriti ta kompleksen in večplasten pojav, drugi pa, kako večopravilnost sodelujočih v raziskavah vpliva na kakovost podatkov, pridobljenih z raziskavo. Ta disertacija obravnava zlasti drugi izziv v kontekstu spletnih anket.

Medijska večopravilnost v spletnih anketah je redko predmet empiričnih raziskav (Zwarun in Hall 2014; Ansolabehere in Schaffner 2015; Sendelbah in drugi 2016). Iz tega izhajata dva izziva, ki jih obravnava pričujoča disertacija: kako meriti *večopravilnost anketirancev* med spletnimi anketami in kakšen je odnos med večopravilnostjo in *kakovostjo odgovorov*? Pri tem se večopravilnost anketirancev nanaša na kombiniranje odgovarjanja na vprašalnik z drugimi (tj. *sekundarnimi*) opravili (Zwarun in Hall 2014), kakovost odgovorov pa se nanaša na tiste vidike kakovosti podatkov, ki so odvisni od anketirančevega zaznavanja, razumevanja in interakcije z vprašalnikom.

Ob sledenju tema dvema specifičnima izzivoma sem po uvodni obravnavi problematike v prvem poglavju v drugem poglavju disertacije najprej predstavil in razvil teoretični okvir na širokem, multidisciplinarnemu naboru literature. Na podlagi tega teoretičnega okvirja sem v tretjem

poglavju disertacije identificiral štiri raziskovana vprašanja in eno hipotezo. V četrtem poglavju disertacije sem predstavil metodologijo empirične študije na spletni anketi ter rezultate in diskusijo. Peto poglavje vsebuje povzetek glavnih ugotovitev ter razpravo o omejitvah, smernicah za nadaljnjo raziskovalno delo in izvirnem prispevku te disertacije.

V preostanku tega razširjena povzetka bom najprej povzel ključne koncepte in opredelitve teoretičnega okvirja (Poglavje 1 tega razširjenega povzetka). Nato bom strnil poglavitne empirične ugotovitve iz obravnavane literature o medijski večopravilnosti in večopravilnosti anketirancev ter raziskovalne cilje te disertacije, ki so oblikovani glede na ključne raziskovalne vrzeli obstoječe literature o večopravilnosti anketirancev (Poglavje 2).

Sledi predstavitev novih pristopov za merjenje večopravilnosti anketirancev, ki so bili uporabljeni v empiričnem delu te disertacije (Poglavje 3), ter ključnih ugotovitev empiričnega dela (Poglavje 4). Sklenem s krajšo obravnavo ključnih omejitev disertacije (Poglavje 5), smernic za nadaljnje raziskovanje (Poglavje 6) in izvirnih znanstvenih prispevkov te disertacije (Poglavje 7).

1 Pregled teoretičnih opredelitev in konceptov

Glavni namen teoretičnega okvirja je vzpostavitev celostnega razumevanja večopravilnosti v spletnih anketah. Teoretični okvir temelji na razumevanju, da je večopravilnost nadpomenka medijske večopravilnosti in da je medijska večopravilnost nadpomenka večopravilnosti anketirancev. Temu razumevanju sledi nabor literature – v teoretičnem okvirju vključujem raziskave na področju večopravilnosti (zlasti iz kognitivnih znanosti), medijske večopravilnosti (iz širokega nabora disciplin) in večopravilnosti anketirancev (zlasti iz anketne metodologije).

Definicije večopravilnosti (e.g. König in Waller 2010; Salvucci in Taatgen 2011; Circella in dr. 2012) in medijske večopravilnosti (Foehr 2006; Wallis 2010; Tokan in dr. 2011) se lahko izrazito razlikujejo med posameznimi disciplinami ali celo znotraj posamezne discipline. V začetnih fazah razvoja teoretičnega okvirja je zato pomembna opredelitev (a) večopravilnosti in (b) medijske večopravilnosti. Na podlagi tega nato (c) opredelim večopravilnost anketirancev in predstavim taksonomijo sekundarnih aktivnosti.

V teoretičnem okvirju je poseben poudarek na teoretični sintezi literature, ki ponuja izsledke o dveh specifičnih izzivih večopravilnosti v spletnih anketah, ki sta bila opredeljena v uvodu tega razširjenega povzetka. Najprej obravnavam izziv (d) odnosa med večopravilnostjo in kakovostjo odgovorov v spletnih anketah, nato pa še izziv (e) merjenja večopravilnosti.

(a) Disertacija sledi široko zastavljeni opredelitvi večopravilnosti, ki jo pri svojem delu uporabljata Salvucci in Taatgen (2011). Avtorja razlagata večopravilnost in njeno raziskovanje s tremi kontinuumi.

Kontinuum večopravilnosti opisuje večopravilnost na spektru med *sočasno večopravilnostjo* (npr. dve aktivnosti se izvajata hkrati, pozornost je razdeljena med obe) in *zaporedno večopravilnostjo* (npr. dve aktivnosti se izmenjujeta v daljših kombinacijah, kjer je večina pozornosti naenkrat namenjena eni izmed teh dveh aktivnosti). Pri zaporedni večopravilnosti je pomemben koncept *preklopa* med aktivnostmi, ki označuje, kdaj se pozornost preusmeri z ene aktivnosti na drugo.

Kontinuum abstrakcije opisuje časovno ločljivost večopravilnosti. Drugače povedano, večopravilnost je lahko opredeljena in/ali merjenja na štirih različnih osnovnih ravneh: biološki pas (nevrološki in fiziološki procesi, merjeni v milisekundah ali manjših časovnih enotah), kognitivni pas (specifične faze aktivnosti, merjene v sekundah), racionalni pas (dokončane aktivnosti, merjene v minutah ali urah) in družbeni pas

(obnašanje v daljšem obdobju, merjenjem v dnevih, tednih ali večjih časovnih enotah).

Kontinuum aplikacije opisuje, kako blizu je empirična raziskava večopravilnosti dejanskemu vsakodnevnemu scenariju. Na eni strani tega spektra so manj aplikativne raziskave, kot so npr. eksperimenti v laboratorijih, na drugi strani so bolj aplikativne raziskave, kot so opazovalne študije.

(b) Opredelitev medijske večopravilnosti je prav tako široko zastavljena in temelji predvsem na Wallisu (2010). Glede na vpetost v medijsko okolje ločim med *zunanjo* medijsko večopravilnostjo (kombinacije medijskih in nemedijskih aktivnosti), *notranjo* medijsko večopravilnostjo (kombinacije zgolj medijskih aktivnosti) in medijsko večopravilnostjo *na eni napravi* (kombinacije medijskih aktivnosti zgolj na eni napravi).

(c) Opredelitev večopravilnosti anketirancev je osnovana na opredelitvah večopravilnosti in medijski večopravilnosti. Večopravilnost anketirancev je vsaka večopravilnost, kjer je ena od aktivnosti odgovarjanje na anketo. Odgovarjanje na spletno anketo je medijska aktivnost, torej je večopravilnost anketirancev v spletnih anketah oblika medijske večopravilnosti.

Taksonomija sekundarnih aktivnosti temelji na taksonomiji iz Zwarun in Hall (2014), ki pa je dopolnjenja z relevantnimi koncepti iz širšega teoretičnega okvirja. Sekundarne aktivnosti so najprej ločene glede na to, ali so kombinirane z odgovarjanjem na anketo sočasno ali zaporedno. Zaporedne sekundarne aktivnosti se naprej delijo na nemedijske aktivnosti ter medijske aktivnosti. Pri zadnjih ločim med aktivnostmi na isti napravi (torej napravi, ki jo anketiranec uporablja za dostop do spletne ankete) in na drugih napravah.

Dva pomembna koncepta, povezana s taksonomijo, sta *instanca* in *oblika* sekundarnih aktivnosti. Instanca se nanaša na dejansko posamezno sekundarno aktivnost (npr. obisk spletne strani o vremenski napovedi).

Oblika sekundarnih aktivnosti opisuje skupine sorodnih sekundarnih aktivnosti (npr. obisk spletnih strani).

Prav tako je pomemben koncept tudi *prekinitev* spletne ankete, ki se zgodi, kadar je večina anketirančeve pozornosti namenjena instancam sekundarnih aktivnosti oz. aktivnostim. Pri tem je lahko ena prekinitev povezana z več instancami in oblikami zaporednih sekundarnih aktivnosti. Velja tudi obratno – ena sekundarna aktivnost je lahko povezana z več prekinitvami.

(d) Disertacija je osredotočena na dva ključna aspekta kakovosti odgovorov: *neodgovorjanje* in *zadostovanje*. Neodgovarjanje vključuje različne oblike odsotnosti odgovorov na vprašanja (npr. neodgovor na posamezna vprašanja, zapustitev ankete pred koncem). Izraz zadostovanje opredeljuje obnašanje anketirancev, ki sicer odgovorijo na vprašanja, vendar pri tem delajo kognitivne bližnjice oz. procesu odgovarjanja ne namenijo optimalne količine kognitivnih virov.

Na podlagi dognanj različnih raziskovalnih tokov, vključenih v teoretični okvir, predlagam, da se odnos med večopravilnostjo in kakovostjo odgovorov lahko kaže na treh ravneh.

Prva raven opisuje vzročni odnos med sekundarnimi aktivnostmi in kakovostjo odgovorov. Osnovno izhodišče razumevanja tega odnosa temelji na *Threaded cognition theory*, ki sta jo utemeljila Salvucci in Taatgen (2011). Njuna teorija opisuje, kako lahko večopravilnost poslabša kakovost izvedbe opravil, vendar pa sta intenziteta ali celo obstoj negativnih učinkov odvisna od številnih dejavnikov, kot so kompleksnost in podobnost kombiniranih aktivnosti, karakteristike prekinitev in splošne kognitivne sposobnosti.

Druga raven v odnos vključi še en element – angažma oz. *vključenost anketirancev* v anketni proces. Vključenost anketirancev se nanaša na to, koliko so anketiranci osredotočeni, zavzeti in motivirani (angl. engagement) za sodelovanje v anketnem vprašalniku (Callegaro in dr.

2015). Glede na obstoječo literaturo predpostavljam, da sta večopravilnost in vključenost anketirancev v medsebojno odvisnem odnosu – pri manj angažiranih anketirancih je večja verjetnost, da bodo začeli opravljati druge, sekundarne aktivnosti med anketo, večja vpetost v sekundarne aktivnosti pa lahko vodi k manjši stopnji vključenosti v anketo. Pri tem je pomembno dodati, da je manjša vključenost sama po sebi povezana s slabšo kakovostjo odgovorov ne glede na prisotnost in obliko sekundarnih aktivnosti.

Tretja raven opisuje odnos med kakovostjo odgovorov in vsakodnevnimi navadami, povezanimi z medijsko večopravilnostjo. Pri utemeljevanju tega odnosa se naslanjam predvsem na delo Ophirja in drugih (2009). Avtorji odmevnega članka so bili osredotočeni na ljudi, ki v vsakodnevnem življenju zelo pogosto kombinirajo medijske aktivnosti (tj. notranja medijska večopravilnost); ljudi z intenzivnimi oblikami takšnega obnašanja so poimenovali »*high media multitaskers*« oz. *HMM* (Ophir in dr. 2009). Njihovi rezultati so nakazali, da majo HMM slabše sposobnosti osredotočanja na eno aktivnost, tudi kadar ni bilo prisotnih sekundarnih aktivnosti. To ugotovitev so dokazale tudi nadaljnje raziskave drugih avtorjev (npr. Cain in Mitroff 2011; Lui in Wong 2012; Uncapher in dr. 2016). Na podlagi tega je mogoče domnevati, da HMM podajajo odgovore slabše kakovosti ne glede na to, ali med anketo izvajajo še kakšne sekundarne aktivnosti.

(e) Kot nakazuje do zdaj obravnavana literatura, se večopravilnost nanaša na kompleksen in večplasten skupek obnašanj, ki predstavljajo znaten metodološki izziv. V empirični literaturi lahko najdemo zelo raznolike pristope merjenja večopravilnosti, ki so vsaj delno pod vplivom različnih teoretičnih izhodišč posameznih disciplin, ki preučujejo to področje. Več avtorjev (npr. Jeong in dr. 2005; Wallis 2006; Tokan in dr. 2011) je opozorilo na pomanjkanje širše razprave o primernosti in primerljivosti posameznih metodoloških pristopov.

V kontekstu te disertacije metode merjenja medijske večopravilnosti ločujem na *reaktivne* in *nereaktivne* pristope, pri čemer se opiram na definicijo Leeja in drugih (2008). Reaktivni pristopi so tisti pristopi, pri katerih udeleženci raziskave aktivno sodelujejo: se odzivajo na pozive izvajalcev raziskave – tipični primer je, da odgovorijo oz. *samoporočajo* na postavljeno vprašanje. Pri nereaktivnih pristopih se takšno aktivno sodelovanje ne dogaja – namesto tega se obnašanje udeležencev opazuje in beleži. V kontekstu te disertacije sem osredotočen na nereaktivne pristope, ki temeljijo na *elektronskih sledeh*, ki se zbirajo na medijskih napravah in beležijo njihovo uporabo.

Pregled empirične literature nakazuje, da ima vsak pristop specifične prednosti in slabosti. Glavna omejitev samoporočanja je, da ljudje pogosto podcenjujejo količino svoje večopravilnosti (npr. Kaye in Sapolsky 1997; Iqbal in Horvitz 2007; Brasel in Gips 2011). Po drugi strani je implementacija reaktivnih pristopov, ki temeljijo na elektronskih sledeh, pogosto zahteven tehnični poseg. Posledično so ti pristopi omejeni na merjenje večopravilnosti na zgolj eni napravi (npr. Kraushaar in Novak 2010; Hill 2014; Yeykelis in dr. 2014). Pri tem je treba upoštevati, da je uporaba takšnih nereaktivnih pristopov v znanosti še vedno v relativno začetni fazi. Več avtorjev (Greenberg in dr. 2005; Wallis 2010; Möller in dr. 2013) je pozvalo k širši uporabi takšnih pristopov v raziskovanju medijske večopravilnosti in uporabe medijev na splošno.

Tudi v specifičnem kontekstu raziskovanja večopravilnosti anketirancev so reaktivni pristopi pogosteje uporabljeni kot nereaktivni. Reaktivni pristopi se tukaj tipično pojavijo v obliki vprašanj na koncu anketnega vprašalnika, ki pozovejo anketirance, naj poročajo o svojih sekundarnih aktivnosti med anketo (Zwarun in Hall 2014; Ansolabehere in Schaffner 2015; Antoun 2015). Pri tem je treba poudariti, da trenutno ni diskusije o tem, kakšna vprašanja so najprimernejša za raziskovanje večopravilnosti anketirancev. Takšna diskusija oz. evalvacija bi bila zlasti potrebna zaradi zgoraj opisane ključne omejitve pristopov, ki temeljijo na samoporočanju.

V spletni metodologiji – in v anketni metodologiji nasploh – se je za elektronske sledi, ki se neposredno zbirajo skozi interakcijo med anketirancem in spletno anketo, uveljavil izraz *parapodatki*. Parapodatki lahko beležijo več različnih tipov dogodkov oz. informacij o obnašanju anketirancev (Callegaro 2013). Izmed njih sta dva tipa dogodkov relevantna za raziskovanje večopravilnost anketirancev. Prvi temelji na *odzivnih časih*, tj. čas, ki ga anketiranec porabi za to, da napreduje, konča oz. odgovori na vprašalnik oz. del vprašalnika (npr. posamezno vprašanje ali stran). Na podlagi ugotovitev iz literature o bolj splošnih oblikah večopravilnosti je mogoče predpostavljati, da izvajanje sekundarnih aktivnosti ponavadi podaljša odzivne čase. Znatno daljši odzivni časi torej lahko nakazujejo, da je anketiranec počel sekundarne dejavnosti. Daljše odzivne čase je mogoče identificirati z določitvijo *praga* oz. časovne točke – če je odzivni čas daljši od praga, v kontekstu te disertacije to označujem z izrazom dogodek *time-out*. Drugi tip dogodkov iz parapodatkov se nanaša na *neaktivnost anketirancev*, ki se lahko odraža v daljših obdobjih pomanjkanja aktivnosti vhodnih naprav (npr. miška ali tipkovnica) ali prek dogodkov *focus-out*. Zadnji se zgodijo, kadar anketiranec preklopi iz okna oz. zavihka, v katerem je spletni vprašalnik, v drugo okno oz. zavihek na svoji napravi.

Specifična omejitev pristopov, ki temeljijo na parapodatkih za merjenje večopravilnosti anketirancev v spletnih anketah, je pomanjkanje vpogleda, v kakšne oblike sekundarnih aktivnosti so bili anketiranci vpleteni. Poleg tega takšni pristopi težje identificirajo nekatere oblike sekundarnih aktivnosti (npr. sočasnih). Posledično obstaja nezanemarljiva možnost lažnih negativnih in pozitivnih zaznavanj večopravilnosti anketirancev.

2 Pregled empiričnih ugotovitev iz literature

Teoretični okvir vključuje tudi pregled in integracijo ključnih empiričnih ugotovitev iz literature o medijski večopravilnosti in večopravilnosti anketirancev. Pri tem je treba poudariti, da je širok nabor empirične literature o medijski večopravilnosti izrazito multidiciplinaren, posledično so vključene raziskave izrazito variirale v konceptualnih in metodoloških okvirjih, ki so bili obravnavani v prejšnjem poglavju. Po drugi strani je pregled empirične literature o večopravilnosti anketirancev omejen z dejstvom, da so bile do zdaj izdane zgolj štiri študije, ki v večjem obsegu raziskujejo točno takšno obnašanje anketirancev – preostale študije obravnavajo širše oziroma posredne vidike (npr. splošen pojav večopravilnosti, obravnava v primeru telefonskih anket ipd.) (Zwarun in Hall 2014; Ansolabehere in Schaffner 2015; Sendelbah in dr. 2016; Antoun 2015) –, sicer pa se tudi te štiri študije pomembno razlikujejo v konceptualnih in metodoloških pristopih.

Pregled je organiziran v tri tematske sklope o večopravilnosti: (a) razširjenost (tkako pogoste so različne oblike večopravilnosti), (b) dejavniki (dejavniki, ki vplivajo na razširjenost večopravilnosti) in (c) odnos s kakovostjo izvedbe ciljno-usmerjenih aktivnosti.

(a) Večina sistematičnih raziskav razširjenosti medijske večopravilnosti v vsakdanjem življenju je osredotočenih na notranjo medijsko večopravilnost (torej zgolj na kombinacije medijskih aktivnosti) (npr. Rideout in drugi 2010; Voorveld in van der Goot 2013; Ofcom 2016). Rezultati teh raziskav nakazujejo, da je medijska večopravilnost vedno bolj prisotna v vsakdanjem življenju in predstavlja približno 30 odstotkov celotnega časa, ki ga ljudje namenijo medijskim aktivnostim. Avtorji teh raziskav poudarjajo, da se aktivnosti, ki so vezane na uporabo interneta in/ali modernih medijskih naprav (npr. osebni računalniki in pametni telefoni), zlasti pogosto pojavljajo v različnih kombinacijah večopravilnosti. To dodatno potrjujejo tudi raziskave, ki so osredotočene

na medijsko večopravilnost na eni medijski napravi. Na primer, več publikacij poroča, da so udeleženci v teh raziskavah med aktivnostmi v povprečju preklopili na vsake tri minute ali pogosteje (Yeykelis in dr. 2014; Judd 2015; Zhang in dr. 2015).

Tri od štirih raziskav o večopravilnosti v spletnih anketah poročajo o deležu anketirancev (Ansolabehere in Schaffner 2015; Antoun 2015; Sendelbah in dr. 2016), za katere je bila opazovana vsaj ena instanca sekundarnih aktivnosti med anketo. Poročani deleži nihajo od 22 odstotkov do 62 odstotkov. Na to verjetno vpliva raznolikost metodoloških pristopov. To vključuje različne dolžine in vsebine anketnih vprašalnikov v raziskavah ter različne opredelitve in pristope k merjenju večopravilnosti anketirancev. Zaradi teh dejavnikov globlja sinteza rezultatov o razširjenosti tega pojava v spletnih anketah žal ni mogoča.

(b) Na podlagi literature predstavim štiri skupine dejavnikov, ki vplivajo na pojavnost in intenziteto medijske večopravilnosti v vsakdanjem življenju: socio-demografski, individualni in strukturni/tehnološki dejavniki ter motivi. Med socio-demografskimi dejavniki izstopa vpliv starosti. Številne raziskave (npr. Voorveld in van der Goot 2013; Hwang in dr. 2014; Ofcom 2016) nakazujejo, da mlajši ljudje pogosteje kombinirajo različne medijske aktivnosti, zlasti tiste, ki so vezane na moderne medijske naprave. Preostale individualne dejavnike se lahko loči na dve podskupini. Prva se nanaša na dostop, lastništvo in uporabo informacijsko-komunikacijskih (IKT) tehnologij. Večja vpetost v IKT-okolje povečuje možnost medijske večopravilnosti. Druga podskupina se nanaša na osebnostne lastnosti. Raziskave so na primer nakazale, da ljudje, ki čutijo večjo potrebo po dražljajih (sensation seeking), pogosteje opravljajo več medijskih aktivnosti hkrati (e.g. Jeong in Fishbein 2007; Sanbonmatsu in dr. 2013; Hwang in dr. 2014). Strukturni/tehnološki dejavniki se nanašajo na raven razvoja in penetracije IKT v družbi – višja ko je ta raven, večja je razširjenost medijske večopravilnosti v družbi, tudi po tem, ko se upošteva vpliv socio-demografskih in individualnih

dejavnikov. Raziskovanje motivov za medijsko večopravilnost razkriva, da so ljudje v takšno obnašanje vpeti iz različnih razlogov – npr. zaradi iskanja razvedrila, občutka povečanja učinkovitosti ali zasvojenosti (medijska večopravilnost je za določene ljudi tako močna navada, da jim je ne uspe nadzorovati).

V empirični literaturi o večopravilnosti anketirancev je več kot ena študija kot pomembna potrdila samo dva dejavnika. Zwarun in Hall (2014) ter Ansolabehere in Schaffner (2015) so potrdili, da starost vpliva tudi na večopravilnost v spletnih anketah – mlajši anketiranci so bili pogosteje vpeti v sekundarne aktivnosti. Ansolabehere in Schaffner (2015) ter Antoun (2015) poročajo o vplivu naprave, ki se uporablja za odgovarjanje na spletno anketo. Večopravilnost anketirancev je bila bolj razširjena med tistimi, ki so do ankete dostopali prek pametnih telefonov, kot med tistimi, ko so za ta namen uporabljali osebne računalnike.

Posamezne publikacije (zlasti Ansolabehere in Schaffner (2015)) poročajo tudi o vplivu drugih dejavnikov (npr. zakonski stan, izkušnje s spletnimi anketami), celostno gledano pa v trenutni literaturi primanjkuje več sistematičnega raziskovanja vpliva dejavnikov. V povezavi s tem bi spomnil, da sem v Poglavju 1 tega razširjenega povzetka (pod točko (d)) podal teoretično podkrepljeno razlago o potencialnih vplivih, ki jih na večopravilnost anketirancev lahko imajo vključenost anketirancev in vsakodnevne navade, povezane z medijsko večopravilnostjo.

(c) Raziskovanje odnosa med večopravilnostjo in kakovostjo izvedbe ciljno usmerjenih aktivnosti ima relativno dolgo znanstveno tradicijo, ki traja že skoraj stoletje (Meyer in Kieras 1997). V kontekstu raziskav o medijski večopravilnosti je bil ta odnos raziskan tako v smislu vzročnosti (npr. Salvucci in Taatgen 2011) kot drugih oblik odnosa (npr. raziskave o HMM, ki so že bile na kratko povzete v točki (e) Poglavja 1). Različne oblike odnosov so bile preverjane tako z manj aplikativnimi raziskavami (npr. laboratorijski eksperiment v Ophir in drugi (2009)) kot tudi na bolj

aplikativne načine (npr. šolski uspeh v Junco in Cotten (2012)). Celostno gledano literatura potrjuje, da je večopravilnost lahko v močnem odnosu s kakovostjo. Kot že rečeno, pa na ta odnos vplivajo različni dejavniki (npr. kompleksnost in podobnost kombiniranih aktivnosti, karakteristike prekinitev in splošne kognitivne sposobnosti).

Dozdajšnje empirične študije niso potrdile močnega odnosa med večopravilnostjo anketirancev in kakovostjo podatkov. To ne velja le za spletne ankete (Ansolabehere in Schaffner 2015; Antoun 2015; Sendelbah in dr. 2016), ampak tudi za telefonske (Kennedy 2010; Lavrakas in dr. 2010). Te empirične ugotovitve vsaj na prvi pogled niso skladne s teoretično utemeljitvijo tega odnosa (točka (d) v Poglavju 1) in empiričnimi dokazi v kontekstu medijske večopravilnosti (povzeti v prejšnjem odstavku). To neskladje nakazuje potrebo po bolj poglobljenih raziskavah o odnosu med večopravilnostjo anketirancev in kakovostjo odgovorov.

3 Raziskovalni cilji in raziskovalna metodologija

Kot že omenjeno v uvodu tega razširjenega povzetka, poskuša disertacija dati vpogled v dva poglavitna izziva večopravilnosti anketirancev v spletnih anketah. Prvi izziv je, kako meriti večopravilnost anketirancev. Po pregledu literature sem se odločil v povezavi s tem izzivom primerjati reaktivni (v obliki vprašanj ob koncu vprašalnika, ki anketirance pozovejo, da samoporočajo o morebitni vpetosti v sekundarne aktivnosti) in nereaktivni pristop (ta temelji na parapodatkih, natančneje na dogodkih focus-out in time-out).

Drugi izziv, ki ga večopravilnost anketirancev predstavlja za spletno metodologijo, pa je, kako (in če sploh) je pojav povezan s kakovostjo podatkov. V povezavi s tem izzivom sem se odločil za poglobljeno analizo odnosa med večopravilnostjo in kakovostjo podatkov, ki vključuje tudi vplive pomembnih dejavnikov, ki niso bili obravnavani v dozdajšnji empirični literaturi.

Pomembno je še poudariti, da je raziskava osredotočena samo na anketirance, ki na anketo odgovarjajo prek osebnih računalnikov. Vključitev in primerjava anketirancev, ki do ankete dostopajo prek drugih naprav (npr. prek pametnega telefona ali tablice), ni bila mogoča glede na razpoložljiva sredstva in opredeljene ključne raziskovalne cilje.

Na podlagi zgoraj opisanih splošnih izzivov in ciljev ter teoretičnega okvirja in specifičnih raziskovalnih vrzeli sem (a) razvil štiri glavna raziskovalna vprašanja in eno hipotezo. Raziskovalnim ciljem sledi (b) zasnova reaktivnega in (c) nereaktivnega raziskovalnega pristopa merjenja večopravilnosti anketirancev.

(a) Vsak od glavnih raziskovalnih ciljev (torej štiri raziskovalna vprašanja in ena hipoteza) ima opredeljene še natančnejše podcilje. Vendar pa podrobnejša razlaga presega omejitve tega razširjenega povzetka, zato tukaj zgolj naštejem glavne raziskovalne cilje:

> **Raziskovalno vprašanje 1:** Kako so indikatorji večopravilnosti anketirancev, ki temeljijo na parapodatkih, povezani z indikatorji večopravilnosti anketirancev, ki temeljijo na samoporočanju?

Bistvo tega vprašanja je primerjava dogodkov focus-out in time-out z indikatorji, ki temeljijo na samoporočanju. Koncept dogodkov focus-out (tj. prekinitev ankete zaradi aktivnosti na isti napravi, ki se uporablja za izpolnjevanje ankete) je mogoče ubesediti v anketnem vprašanju. Po drugi strani je koncept dogodkov time-out (tj. odstopanje od oz. preseganje določene kritične časovne točke, ki je običajno določena s statističnimi postopki) preveč abstrakten, da bi o njem bilo mogoče povprašati anketirance. Neposredna primerjava dogodkov time-out z indikatorji samoporočanja torej ni mogoča, vseeno pa je mogoče analizirati bolj splošne oblike povezanosti.

> **Raziskovalno vprašanje 2:** Kakšna je razširjenost različnih oblik večopravilnosti anketirancev v spletni anketi glede na indikatorje, ki temeljijo na parapodatkih in samoporočanju?

Obravnava tega vprašanja sledi taksonomiji sekundarnih aktivnosti, ki je bila opisana v točki (c) v Poglavju 1. Specifične oblike večopravilnosti anketirancev vključujejo različne medijske in nemedijske forme sekundarnih aktivnosti ter število začasnih prekinitev izpolnjevanja ankete zaradi sekundarnih aktivnosti.

> **Raziskovalno vprašanje 3:** Kakšen je vpliv socio-demografskih in individualnih dejavnikov na razširjenost večopravilnosti anketirancev?

Tukaj je poglavitna osredotočenost na vlogo starosti, vpetosti anketirancev in vsakodnevnih navad medijske večopravilnosti. Drugi socio-demografski (starost in izobrazba) in individualni (npr. izkušnje s spletnimi anketami) dejavniki so prav tako vključeni v raziskavo.

> **Raziskovalno vprašanje 4:** Kako so indikatorji večopravilnosti anketirancev povezani s kakovostjo odgovorov?

Kakovost odgovorov v spletnih anketah je povezana z dvema poglavitnima oblikama obnašanja: neodgovarjanje in zadostovanje. Neodgovarjanje se v tej raziskavi meri z indikatorjem števila neodgovorov na anketiranca (tj. frekvenca vprašanj, na katera anketiranec ni odgovoril). Zadostovanje se meri z več indikatorji, med drugim nediferenciacijo odgovorov v matričnih vprašanjih, vsebinsko konsistentnostjo odgovarjanja in dolžino odgovorov na odprta vprašanja.

> **Hipoteza:** Pristop s parapodatki poda podobne ključne ugotovitve o razširjenosti in dejavnikih večopravilnosti anketirancev ter o odnosu s kakovostjo podatkov kot pristop merjenja s samoporočanjem.

Hipoteza v bistvu trdi, da med pristopoma merjenja ni večjih razlik glede ključnih vsebinskih ugotovitev, ki so zajete v raziskovalnih vprašanjih 2, 3 in 4. Glede na obravnavno literaturo v teoretičnem okvirju velja izpostaviti dve pomembni dinamiki za to hipotezo: pristop, ki temelji na samoporočanju, je lahko pod močnim vplivom podcenjevanja anketirancev

o prisotnosti sekundarnih aktivnosti, pri pristopu, ki temelji na parapodatkih, pa obstaja znatna verjetnost, da vse oblike večopravilnosti anketirancev (npr. sočasne sekundarne aktivnosti) ne bodo zajete v meritvah. Vpliva teh dveh dinamik še nista bila raziskana na primeru večopravilnosti anketirancev, zaradi tega ni mogoče razviti hipoteze, ki bi bila usmerjena (na primer v smislu trditve, da bo pristop, ki temelji na parapodatkih, izmeril večjo prisotnost večopravilnosti anketirancev).

(b) Empirična študija (reaktivni pristop) je bila izvedena na 20-minutni spletni anketi, anketiranci so bili rekrutirani med člani slovenskega spletnega panela askGfK. Anketiranje je potekalo od 2. junija do 7. julija 2016. Spletni panel ne uporablja verjetnostnega vzorčenja, uporabljeno pa je bilo kvotno vzorčenje na podlagi spola, starostnih skupin (starostni razredi 15–24, 25–34, 35–44 in 45–55) in statistične regije. Panelisti, ki so se odločili za sodelovanje v raziskavi, so bili preusmerjeni na spletni vprašalnik na strežniku orodja 1KA, kjer je potekalo zbiranje podatkov (to vključuje tudi parapodatke).

Vprašalnik je bil sestavljen posebej za to raziskavo. Prvi del vprašalnika vsebuje vprašanja o medijskih navadah, vključno s prilagojeno različico instrumenta Media use questionnaire (Ophir in dr. 2009), ki se uporablja za merjenje indeksa intenzitete vsakodnevne medijske večopravilnosti (MMI – glej Poglavje 1). Drugi del vprašalnika vsebuje vprašanja o različnih vidikih anketirančevih življenjskih navad, stališč in osebnostnih lastnosti. Tretji del vprašalnika sestavljajo vprašanja o anketirančevih izkušnjah in stališčih do anket (tako na splošno kot o tej točno določeni anketi), vprašanja o anketirančevi večopravilnosti ter socio-demografska vprašanja. V celoti vprašalnik sestavlja 61 vprašanj, ki vsebujejo 202 spremenljivki in so razporejena na 37 straneh.

Skupno 2.059 panelistov je prispelo na uvodno stran vprašalnika. Iz analize so bili nato izločeni anketiranci, ki: niso prispeli do zadnje strani vprašalnika; niso odgovorili na vsaj polovico spremenljivk v vprašalniku;

niso uporabljali osebnega računalnika za dostop do ankete. V naslednjem koraku so bili iz analize izločeni anketiranci, katerih čas odgovarjanja na celotno anketo je bil med dvema odstotkoma najpočasnejših ali dvema odstotkoma najhitrejših v vzorcu. Anketiranci, ki močno odstopajo v času odgovarjanja, namreč predstavljajo netipične ekstremne oblike interakcije z vprašalnikom. Končni vzorec, ki je vključen v analizo, vsebuje 1.366 anketirancev. Socio-demografska struktura končnega vzorca je primerljiva z uradnimi podatki Statističnega urada Republike Slovenije, kar zadeva spol in starostne skupine, obstajajo pa pomembne razlike v izobrazbeni strukturi (glede na uradne podatke je v tej raziskavi več visoko izobraženih anketirancev in manj nižje izobraženih, medtem ko je delež srednje izobraženih primerljiv).

(c) Nereaktivni pristop temelji na dveh tipih parapodatkov: dogodki focus-out in time-out, oboji so že bili opisani v prejšnjem podpoglavju. Pri tem velja dodatno izpostaviti, kako so določeni pragi za dogodke time-out. V moji raziskavi dogodki time-out temeljijo na odzivnih časih na straneh vprašalnika. Vsi odzivni časi po straneh vseh anketirancev so analizirani z mešanimi modeli kot odvisna spremenljivka. Vplivi posameznih strani vprašalnika in anketirancev so modelirani kot naključnostni dejavniki za posamezne strani oz. anketirance. Dogodek time-out na posamezni strani za posameznega anketiranca je identificiran, če je vrednost njegovega ostanka v modelu večja od ena.

Reaktivni pristop temelji na samoporočanju na dve vprašanji na koncu vprašalnika. Prvo vprašanje pozove anketirance, da poročajo o tem, ali so izvajali posamezne oblike sekundarnih aktivnosti ali ne. Seznam ponujenih sekundarnih aktivnosti deloma temelji na enem najbolj uporabljenih instrumentov za merjenje medijske večopravilnosti (Ophir in dr. 2009). Drugo vprašanje pozove anketirance, naj poročajo o frekvenci prekinitev med anketo zaradi aktivnosti na isti napravi, drugi napravi ali zaradi nemedijskih sekundarnih aktivnosti.

4 Glavne empirične ugotovitve te disertacije

V tem poglavju najprej obravnavam rezultate o razširjenosti (4.1) in dejavnikih razširjenosti (4.2) večopravilnosti anketirancev. Nato povzamem ugotovitve o odnosu večopravilnosti s kakovostjo odgovorov (4.3) ter primerjam pristopa k merjenju večopravilnosti anketirancev glede na ključne vsebinske ugotovitve (4.4). V zadnjem podpoglavju (4.5) glavne ugotovitve nato na kratko tudi strnem.

4.1 Razširjenost

Več kot polovica anketirancev je med spletno anketo počela še nekaj drugega. Točen delež je odvisen od uporabe oz. interpretacij posameznih indikatorjev oz. njihovega kombiniranja, vendar pa podobne ugotovitve ponujajo tako parapodatki kot samoporočanje.

Sodeč po samoporočanju so najpogostejši razlog za prekinitve nemedijske sekundarne aktivnosti (29 %), ki ji sledijo sekundarne aktivnosti na isti napravi (25 %) in drugih medijskih napravah (18 %). Na splošno 35 odstotkov anketirancev poroča o vsaj eni medijski sekundarni aktivnosti, medtem ko je pri nemedijskih ta delež 26-odstotni.

Analiza (tako parapodatkov kot samoporočanja) prav tako nakazuje, da je zgolj petina (ali celo manj) anketirancev vpletena v bolj intenzivne oblike večopravilnosti med anketo (npr. več kot ena prekinitev ali pa več kot ena sekundarna aktivnost). To potrjujejo tudi najpogosteje poročane oblike sekundarnih aktivnosti, kjer prevladujejo relativno preprosti primeri aktivnosti, kot so poslušanje glasbe, priprava obroka, pitje in prehranjevanje ter pogovor iz oči v oči.

4.2 Dejavniki

Povezanost med dejavniki se je preverjala tako z deskriptivnim statistikami kot z regresijskimi modeli, kjer so bili indikatorji večopravilnosti anketirancev obravnavani kot odvisne spremenljivke, dejavniki pa kot neodvisne.

Rezultati disertacije potrjujejo, da je starost eden izmed pomembnejših dejavnikov večopravilnosti. Mlajši anketiranci so pogosteje vpleteni v večopravilnost, sodeč po vseh indikatorjih razen dogodkov time-out. Glede na starejše anketirance so med mlajšimi zlasti razširjene medijske sekundarne aktivnosti.

Pri dejavnikih izstopa tudi stopnja pozornosti, ki jo je anketiranec namenil izpolnjevanju ankete. Ta dejavnik temelji na odgovorih anketirancev na vprašanje »Koliko pozornosti ste namenili tej anketi?«. Merjen je na petstopenjski lestvici in ga lahko razumemo kot enega izmed vidikov vključenosti anketirancev v anketo. Samoporočana pozornost je povezana z vsemi indikatorji večopravilnosti anketirancev.

Za pomemben dejavnik se je pokazal tudi indeks vsakodnevne medijske večopravilnosti (MMI), ta je povezan z vsemi indikatorji večopravilnosti anketirancev, ki temeljijo na samoporočanju. Preostali dejavniki, ki so bili vključeni v analizo (npr. spol, izobrazba, motivacija za sodelovanje v anketi, dostop do IKT), so bili povezani z omejenim naborom indikatorjev večopravilnosti anketirancev.

4.3 Odnos s kakovostjo

Odnos s kakovostjo sem preverjal z regresijskimi modeli, kjer so bili indikatorji kakovosti odgovorov obravnavani kot odvisne spremenljivke, indikatorji večopravilnosti anketirancev in drugi dejavniki, ki bi lahko vplivali na ta odnos, pa kot neodvisne.

Kljub podrobni analizi, ki je vključevala različne oblike indikatorjev in regresijskih modelov, disertacija ni pokazala močne povezanosti med večopravilnostjo anketirancev in različnimi indikatorji kakovosti odgovorov. Na splošno so torej ugotovitve podobne kot v prejšnjih raziskavah o odnosu s kakovostjo v spletnih in telefonskih anketah (glej točko (c) Poglavja 2).

Med indikatorji večopravilnosti anketirancev so bili le dogodki time-out opazneje povezani z več kot enim indikatorjem kakovosti odgovorov (s frekvenco neodgovorov in nediferenciacijo). Vendar pa je tudi v tem primeru bila statistična moč povezave razmeroma nizka.

Pomembna ugotovitev v kontekstu večstopenjskega odnosa med večopravilnostjo in kakovostjo (glej točko (d) Poglavja 1) je, da sta tako samoporočana povezanost kot MMI močneje povezana z indikatorji kakovosti odgovorov kot katerikoli indikator večopravilnosti anketirancev. Za anketirance, ki so poročali, da so anketi posvetili več pozornosti, je bila ugotovljena večja verjetnost, da bodo podajali tudi kakovostne odgovore. Za anketirance, ki so bili glede na MMI kategorizirani v skupino »kronične« vsakodnevne medijske večopravilnosti, pa je bila večja verjetnost, da bodo njihovi odgovore slabše kakovosti.

4.4 Primerjava nereaktivnega (parapodatki) in reaktivnega (samoporočanje) pristopa

Pristopa se primerjata glede ključne ugotovitve, ki jih podajata za tri glavne vsebinske sklope te disertacije: razširjenost, dejavniki razširjenosti in povezava s kakovostjo odgovorov.

Pristopa podajata podobni oceni deleža anketirancev, ki so med anketo počeli vsaj eno sekundarno aktivnost, vendar pa se razlike med njima izkažejo na podrobnejši ravni. Kot primer, pri tretjini anketirancev se zgodi, da so bili vpeti – sodeč po enem izmed pristopov – v vsaj eno sekundarno dejavnost, po drugem pristopu pa ne.

Pri povezavi z dejavniki se pokažejo še večje razlike med pristopoma. Indikatorji večopravilnosti anketirancev, ki temeljijo na samoporočanju, so bolj vsebinsko povezani z različnimi dejavniki kot indikatorji, ki temeljijo na parapodatkih. Nekateri dejavniki (npr. izobrazba, spol, MMI) so tako povezani le z indikatorji večopravilnosti, ki temeljijo na samoporočanju.

Glede odnosa s kakovostjo podatkov oba pristopa podajata podobne vsebinske ugotovitve. Kot že omenjeno, za nobenega od indikatorjev večopravilnosti anketirancev ni mogoče trditi, da je močno povezan s kakovostjo podatkov.

Če povzamem, pomembnejše razlike med pristopoma se pokažejo predvsem pri identifikaciji anketirancev, ki so bili vpeti v sekundarne aktivnosti med anketo, in pri povezanosti z določenimi dejavniki. Vse to pomeni, da osnovna hipoteza disertacije ni bila potrjena.

5 Omejitve

V pogledu omejitev velja izpostaviti predvsem empirično raziskavo, za katero velja več omejitev. Če izpostavim najbolj ključne, so to naslednje:

- Raziskava temelji na neverjetnostnem vzorcu. Ker pridobljeni podatki ne temeljijo na verjetnostnem vzorcu, rezultati načeloma ne morejo biti posplošeni na splošno populacijo.

- Raziskava temelji na spletnem panelu. Panelisti imajo ponavadi več izkušenj z izpolnjevanjem anket kot ljudje, ki niso člani panelov. Razlikujejo se lahko tudi v drugih pomembnih lastnostnih (npr. motivacija za sodelovanje v anketi). Te lastnosti lahko pomembno vplivajo na končne vsebinske ugotovitve o večopravilnosti anketirancev. Kot primer, ker je pri panelistih večja stopnja poznavanja spletnih anket, je mogoče, da jih vpetost v sekundarne aktivnosti manj moti pri izpolnjevanju ankete. Dodati velja, da so na drugi strani panelisti seveda najbolj zanimivi, saj v raziskavah splošne populacije spletni paneli postajajo prevladujoči način zbiranja anketnih podatkov.

- Neeksperimentalni raziskovalni načrt. Ker raziskava ni potekala v obliki eksperimenta, strogo gledano ni mogoče delati sklepov o vzročnostnem odnosu med večopravilnostjo in kakovostjo podatkov.

- Dejavniki večopravilnosti anketirancev. Raziskovalni načrt ni omogočal raziskave nekaterih dejavnikov (npr. oblikovanje in vsebina vprašalnika ter vpliv naprave, ki se uporablja za izpolnjevanje ankete), ki bi lahko predstavljali pomemben vpogled v večopravilnost anketirancev in pomembno vplivali na ključne vsebinske ugotovitve.

- Poleg omejitev empirične raziskave velja omeniti še določene tehnične omejitve pri zajemu parapodatkov in konceptualne omejitve oziroma zožitve fokusa disertacije.

6 Smernice za nadaljnje raziskovanje

Na podlagi obravnavane literature in empiričnega dela izpostavljam naslednje ključne smernice za nadaljnje raziskovanje večopravilnosti anketirancev:

- Nadaljnji razvoj in evaluacija pristopov merjenja večopravilnosti anketirancev. Disertacija ponuja nekaj odgovorov o ključnih vidikih primerjave merjenja večopravilnost anketirancev prek parapodatkov in samoporočanja. Odprto vprašanje pa ostaja, katere specifične metode so optimalne za raziskovanje tega področja. To vključuje obravnavo drugih dogodkov, ki se jih lahko zajema v parapodatkih, ter oblikovanje novih anketnih vprašanj za merjenje večopravilnosti anketirancev.

- Aplikativna raziskava v spletnih anketah, ki temeljijo na podatkih zunaj spletnih panelov. Z izjemo manjše pilotne študije v Sendelbah in drugi (2016) so vse raziskave o večopravilnosti anketirancev v spletnih anketah (vključno s to disertacijo) do zdaj bile omejene na člane spletnih panelov. Kot že orisano v prejšnjem poglavju, bi bilo pomembno preveriti, ali se izkazujejo pomembne razlike v drugih kontekstih spletnega anketiranja.

- Eksperimentalne in kvalitativne raziskave. Vse raziskave (vključno s to disertacijo) so potekale v obliki kvantitativno usmerjenih,

neeksperimentalnih študij. Eksperimentalno in/ali kvalitativno usmerjene raziskave lahko ponudijo dodatne vpoglede v večopravilnost anketirancev ter njenega odnosa z različnimi dejavniki in kakovostjo podatkov.

7 Izvirni prispevek

V teoretičnem delu velja kot izviren prispevek izpostaviti predvsem integracijo literature o večopravilnosti, medijski večopravilnosti in anketni večopravilnosti. V tem okviru velja še posebej poudariti naslednje:

- Nadaljnji razvoj taksonomije večopravilnosti anketirancev. Prvotna taksonomija, ki sta jo predstavila Zwarun in Hall (2014), je nadgrajena s ključnimi koncepti iz literature o širših oblikah večopravilnosti.

- Elaboracija dejavnikov večopravilnosti anketirancev in njenega odnosa s kakovostjo podatkov. Na podlagi literature o medijski večopravilnosti sem izpostavil štiri skupine dejavnikov, ki lahko vplivajo na večopravilnost anketirancev. Teoretični okvir prav tako vzpostavi tri različne ravni odnosa med večopravilnostjo anketirancev in kakovostjo podatkov.

Eden izmed ključnih korakov raziskovalnega dela za to disertacijo je bil tudi razvoj specifičnih postopkov merjenja večopravilnosti anketirancev:

- Parapodatki. Zasnova tega pristopa je bila sicer objavljena že v Sendelbah in drugi (2016), vendar pa je v tej disertaciji pristop dodatno razvit. Pri tem velja izpostaviti zlasti novo strategijo določevanja dogodkov time-out.

- Samoporočanje. Za namen te disertacije so bila oblikovana nova anketna vprašanja o večopravilnosti anketirancev, ki temeljijo na teoretičnem okvirju in stremijo k temu, da je pristop k merjenju tega pojava čim bolj celosten.

Pri empiričnem delu velja izpostaviti zlasti naslednje izvirne prispevke:

- Primerjava reaktivnih in nereaktivnih pristopov. Disertacija predstavlja prvo sistematično primerjavo teh dveh pristopov v kontekstu večopravilnosti spletnih anketirancev. Upam, da bodo rezultati koristni tudi v širšem kontekstu raziskav, ki primerjajo reaktivne in nereaktivne pristope za merjenje različnih medijskih in/ali večopravilnostnih obnašanj.

Novi vpogledi v odnos med večopravilnostjo, dejavniki in kakovostjo podatkov. Ob izhajanju iz teoretičnega okvirja so bili v tej disertaciji raziskani nekateri prej še neobravnavani vidiki tega kompleksnega niza odnosov. Pri tem velja predvsem izpostaviti vlogo vključenosti anketirancev in navad v povezavi z medijsko večopravilnostjo v vsakdanjem življenju.