UC Santa Barbara

Title

The Elusive, But Essential Juggling Act: Iterative Processes of Validation in Developing the Theory of Media Multitasking Intensity

Permalink

https://escholarship.org/uc/item/3rf174zz

Author

Zamanzadeh, Nicole Neda

Publication Date

2019

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA

Santa Barbara

The Elusive, But Essential Juggling Act: Iterative Processes of Validation in

Developing the Theory of Media Multitasking Intensity

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Communication

by

Nicole Neda Zamanzadeh

Committee in charge:

Professor Ronald Rice, Co-Chair

Professor Tamara Afifi, Co-Chair

Professor Andrew Maul

September 2019

The dissertation of Nicole Neda Zamanzadeh is approved.

_____

Andrew Maul

_____

Tamara D. Afifi, Committee Co-Chair

_____

Ronald E. Rice, Committee Co-Chair

June 2019

younger than me continues to blow me away with her maturity, patience, compassion, wisdom, and resilience. In many ways, it feels she has led the way for me, supported me in writing my own story free from the pity parties and stereotypes I've internalized. Thank you, Davina, you continue to be a light brightening my life and makes my victories and failures better. I can't wait until you have your own PhD! I also want thank my parents Niloufar and Behzad for cheerleading my often larger-than-life visions. Your stories, your adversities, your kindness, and your ambitions have made mine possible and even sweeter. I also want to thank my grandparents Habib, Iran, Ashraf, and Parviz who overcame the challenges of transitioning our family from Iran to the US and all instilled the importance of education.

I also would be amiss to not thank my boyfriend Jarrett Gorlick, who has put in innumerable solidarity-work nights, provided daily pep talks, and took care of my sanity with destressing walks, and served as a sounding board. Jarrett, you have asked me to question my stories, and helped me rewrite those that no longer served me. Your grace in the midst of difficulty and empathetic nurturing nature helped me persevere through the many challenges of the program, including two surgeries. You have been my rock, my mirror, my dream, my wish, my fantasy... I love you! Also, thank you for celebrating me with a unicorn piñata, swimsuit run around the block, cape and crown – it was the proper way to finish this!

There are also those friends who cross your path that change your life. I could not be more grateful for meeting Britney Lebaron in my first year of my PhD, her solid and consistent willingness to be authentic and vulnerable gave me the trust my path and gave me the courage to not give up. Thank you, Brit for the many juices, tissues, and

talks that helped me find compassion for myself and own my story. Similarly, Katie

Harrison your openness and integrity allowed me to voice my needs and pursue my

wildest dreams. I could not have imagined crossing the finish line without your daily

inspiration, R support, and venting sessions #Grittygirls. Thank you for inspiring me

and cheerleading me!

If it took a village to complete this project, the militia of this village were my

research asssistents. I do not know where I would be without them: Chrisine Ho,

William Yang, Steven Wu, Ashley Pike, Paula Wang, Steven Shi, Bryan Kim, Ricky

Banuelos and Nick Giosso also known as the Savages. Your bright, eager, and

determined savagery elevated my work not only in thought, but in meaning. Together,

we explored personally-relevant and professionally empowering ideas, supported each

other through challenges, and built a trust that allowed us to accomplish the five studies

covered in this dissertation in less than 3 years. I want to thank each of your for

bringing your vulnerable and total selves to our meetings and for diving deep into media

multitasking literature, coding, and R!

Finally, I also owe my sanity (or what's left of it) and the value of completing

my PhD to the graduate students and friends who uplifted me when I was stressed and

discouraged, who saw me for my intellect and beyond it (i.e., Karaoke), and helped me

celebrate my own path even if others were not. There are many to name, but for now I

want to thank the brave Avi McClelland-Cohen, the creative Audrey Abeyta, the wise

Richard Huskey, the bright Dajung Woo, the thoughtful Spencer Nicholls, and the

kindly-nosy Matt Giles. Ashley Bakhaj, Jasmine Shirazi, Nahal Akhavan, Bahar

Basseri, Sarah Ghodsi, Ashley Selki, and Tegan Brennan thank you for maintaining my

identity as more than a graduate student and supporting me through the marathon called graduate school. Whether it was a dinner, a dessert, or a coffee, each of you connected with me and reminded me to keep a bit of adventure and light-heartedness. Thank you.

"The Bravest Thing You Can Be Is Yourself" - Unknown

NICOLE NEDA ZAMANZADEH

University of California, Santa Barbara
Department of Communication
4005 Social Sciences and Media Studies Bldg.,
Santa Barbara, CA 93106
n_zamanzadeh@ucsb.edu

Education

2016 – 2019   **University of California, Santa Barbara**
PhD in Communication
       Title: *The Elusive, But Essential Juggling Act: Iterative Processes of Validation in Developing the Theory of Media Multitasking Intensity*
Committee Chairs: Tamara D. Afifi and Ronald E. Rice

2013 – 2016   **University of California, Santa Barbara**
MA in Communication (June, 2016)
       Title: *Social and Nonsocial Media Multitasking's Effect on Cognitive Performance: Mood as a Mediator*
Committee Chair: Daniel Linz
Committee Members: Ronald E. Rice and Rene Weber

2009 – 2013   **University of California, Los Angeles**
BA in Communication Studies with Honors
       Title: *The Effect of Feedback in Computer Mediated Communication on Self-Concept and Self-Concept Clarity*
       Advisor: Travis L. Dixon

Research Interests
- Stress, Resilience, & Thriving
- Media Effects
- Family Systems
- Quantitative Methods and Measurement

Publications

1. **Zamanzadeh, N. N**. & Afifi, T. D. (in press). Prosocial relationship maintenance and resilience/thriving. In J. A. Muñiz Velázquez and C. Pulido (Eds.), *Routledge handbook of positive communication*. New York, NY: Routledge.

2. Afifi, T. D. **Zamanzadeh, N. N.**, Harrison, K., & Acevedo Callejas, M. (2018). WIRED:
  The impact of media and technology use on stress (cortisol) and inflammation (interleukin IL-6) in fast paced families. *Computers in Human Behavior, 81,* 265-273.

3. Rice, R. E., Hagen, I. **& Zamanzadeh, N. N.** (2018). College students' media mastery:
  Paradoxes in using computers and mobile phones. *American Behavioral Scientist.*

4. Afifi, T. D., **Zamanzadeh, N. N.,** Harrison, K., & Brandt, K. (forthcoming). Stress, coping, and thriving: The role of interpersonal communication. In S. Wilson and S. Smith (Eds.), Reflections on Interpersonal Communication. Cornella Academic Publishing.

5. Afifi, T. D., Harrison, K., & Zamanzadeh, N. N. (forthcoming). Communication, stress
  and resilience in close relationships. In K. Floyd and R. Weber, The handbook of physiology/biology in communication.

6. Rice, R. E., **Zamanzadeh, N. N.**, & Hagen, I. (forthcoming). Media Mastery by college students: A review and typology. In S. Yates and R. E. Rice (Eds.), *Oxford handbook of digital technology and society.* New York: Oxford University Press.

7. **Zamanzadeh, N. N.** & Afifi, T. D. (forthcoming). *Teen stress: Busting myths & answering your questions.* ABC-CLIO: Goleta, CA.

8. Dunbar, N.E., Giles, H., Bernhold, Q., Adams, A., Giles, M., **Zamanzadeh, N.N**., Gangi, K., Covaleski, S., & Fujiwara, K. (forthcoming). Micro-level Lies about In-group Affiliation: Exploring the Strategic Synchrony Hypothesis.

9. Afifi, T. D., **Zamanzadeh, N. N.,** Harrison, K., & Pérez Torres, D. (Revise & Resubmit). Preventing the "Trump Effect": Explaining the impact of voting patterns on resilience and relational load in romantic relationships. Journal of Social and Personal Relationships.

10. Afifi, T.D., Harrison, K., **Zamanzadeh, N. N.,** & Acevedo Callejas, M. (in review.) Relationship
  maintenance as stress management in fast paced families.

Manuscripts Under Review
  1. Afifi, T. D., Harrison, K., & **Zamanzadeh, N. N.** (Revise & Resubmit) Preventing relational load in families with an adolescent type I diabetes.

Manuscripts in Progress

  1. Harrison, K., **Zamanzadeh, N.N**., & Afifi, T. D. (in data analysis) Relationship maintenance in dating couples, stress and emotional eating behaviors.
  2. **Zamanzadeh, N. N**. & Rice, R. E. (in writing). Individual factors and media task characteristics
     as predictors of media multitasking with difficult academic tasks.

3. **Zamanzadeh, N.N.,** Harrison, K., & Afifi, T. D. (in data analysis) Gratitude and resilience in young adults: Emotional reserves in thwarting depressive symptoms and self-harm behaviors
4. Pérez Torres, D., **Zamanzadeh, N. N** & Afifi, T. D., (in data analysis) The role of media habits in relational stress.
5. Flanagin, A., **Zamanzadeh, N. N.,** & Metzger, M. L. (in data analysis) Credibility assessments of online information and Google page ranking.

Conference Presentations

1. **Zamanzadeh, N. N.,** Linz, D., & Dixon, T. L. (2015, November). *Is it me you're looking for? The effects of CMC feedback.* Poster presented at the meeting of the National Communication Association, Las Vegas, NV.
2. **Zamanzadeh, N. N.,** & Linz, D. (2016, June). *The mood regulatory function of media multitasking,* Paper presented at the meeting of the International Communication Association, Fujioka, Japan.
3. Rice, R. E., **Zamanzadeh, N**. **N.,** & Hagen, I. (2017, May). *College students' media mastery: Themes of tensions, contradictions, and paradoxes in using computers and mobile phones.* Presentation on panel Mapping digital divides in communication practices across major life realms. International Communication Association, San Diego, CA.
4. Rice, R. E & **Zamanzadeh, N. N.** (2017, October). *Media mastery by college students: A review and proposed narrative sequence*. Paper presented at the meeting of the WBDA2017: Ways of Being in a Digital Age: A Review Conference. UK: University of Liverpool.
5. **Zamanzadeh, N. N.** (2017, November). *An initial validation of the media multitasking intensity instrument (MMTII).* Paper presented at the meeting of the National Communication Association, Dallas, TX
6. Afifi, T. D., **Zamanzadeh, N. N.**, Harrison, K., & Acevedo Callejas, M. (2017, November). *WIRED: The impact of media and technology use on stress (cortisol) and inflammation (interleukin IL-) in fast paced families*. Paper presented at the meeting of the National Communication Association. Dallas, TX.
7. Rice, R. E., **Zamanzadeh, N**. N., & Hagen, I. (2017, November). *College students' media mastery: Themes of tensions, contradictions, and paradoxes in using computers and mobile phones.* Paper presented at the meeting of the National Communication Association, Dallas, TX.
8. **Zamanzadeh, N. N.** (2018, May). *Getting the hamster off the wheel: The search for meaningful media effects research.* Paper presented at the meeting of the International Communication Association, Prague, The Czech Republic.
9. Afifi, T., Harrison, K., **Zamanzadeh, N. N.,** & Acevedo Callejas, M. (2018, May). *Relationship maintenance as stress management in fast paced families*. Paper to be presented at the annual meeting of the International Communication Association, Prague, Czech Republic. * Top Paper Award in Interpersonal Communication Division *

10. Afifi, T. D., **Zamanzadeh, N. N.**, Harrison, K., & Pérez Torres, D. (2018, November). *Preventing the "Trump Effect": Explaining the impact of voting patterns on resilience and relational load in romantic relationships*. Paper to be presented at the meeting of the National Communication Association, Salt Lake City, UT.

11. Afifi, T. D., **Zamanzadeh, N. N.**, Harrison, K., & Acevedo Callejas, M. (2018, October). *WIRED: The impact of media and technology use on stress (cortisol) and inflammation (interleukin IL-) in fast paced families*. Poster presented at the meeting of Digital Media and Developing Minds, Cold Springs Harbor, NY.

Teaching Experience

University of California, Santa Barbara (September 2013–Present)
Teaching Associate (instructor of record)
- COMM 133: Children and Media (Spring 2018)
- COMM 88: Research Methods (Summer 2018)

Teaching Assistant
- COMM 1: Introduction to Communication (Fall 2013, Fall 2015, Winter 2015)
- COMM 87: Statistics for Social Scientists (Winter 2016)
- COMM 88: Communication Research Methods (Spring 2015, Spring 2016, Summer 2016, Winter 2019)
- COMM 89: Communication Theory (Winter 2014, Spring 2017, Fall 2017)
- COMM 107: Interpersonal Communication (Winter 2017)
- COMM 126: Gender and Communication (Fall 2016)
- COMM 133: Children and Media (Winter 2018)
- COMM 156: Health Communication (Fall 2014)

Santa Barbara City College (SBCC), Santa Barbara (September 2017- March 2018)
- COMM 101: Intro to Communication (Fall 2017)
- COMM 131: Public Speaking (Spring 2018)

Honors and Recognition

2018   ICA Communication and Technology & Mobile Communication Doctoral Consortium
2018   ICA Interpersonal Communication Division Top Paper Award with Dr. Tamara D. Afifi, Kathryn Harrison, & Michelle Acevedo Callejas
2013   Graduated with Honors in Communication Studies from UCLA

Awards and Grants

2018	Outstanding Graduate Service Award
2012 –2013	Wasserman Foundation Grant $5,000 awarded to complete one-year independent research project on cyberbullying.

Service

2017 –	All-Grad representative in Graduate Student Advisors to the Chair Committee (GSACC)
	• Designed Professional Development Series
2017	Designed and organized speaker series for Quantitative Methods and Social Sciences (QMSS) emphasis
2017	Served on Early Career and Student Committee for SRCD special topics, Technology, Media and Child Development (TMCD) conference
2016 –	Reviewer for Communication and Technology Division, Children, Adolescents and Media, Information Systems, ICA
2015 –	Reviewer for Communication and the Future Interest Group, NCA

Professional Association Membership

2016 –	Society of Research in Child Development (SRCD)
2013 –	International Communication Association (ICA)
2012 –	National Communication Association (NCA)

Skills

Software: Proficient in SPSS, SSPS Amos, R, Mplus, and NVIVO
Statistical Analysis: Familiar with Item-Response Theory, Rasch Models, Factor Analysis, Structural Equation Modeling, Mixture Modeling.
Language: Proficient in Farsi/Persian and Hebrew

ABSTRACT

The Elusive, But Essential Juggling Act: Iterative Processes of Validation in

Developing the Theory of Media Multitasking Intensity

by

Nicole Neda Zamanzadeh

The purpose of the current dissertation was multi-faceted: 1) propose the value of

an iterative process of measure validation for theoretical advances, 2) iteratively

develop and evaluate validity of the Media Multitasking Intensity Questionnaire

(MMTIQ), and 3) propose and examine a theory of media multitasking intensity.

The current dissertation proposes a relationship between current measure

development and validation practices, inconsistent empirical evidence, and stalled

theoretical development. The inability to replicate seminal empirical findings and thus

evidence theoretical validity has led the social sciences to begin to re-examine norms

for empirical research. Of the various solutions that have been proposed, the role of

measurement and validation has been unexamined and underestimated. Validation is

often limited to a single-effort or justified via citation and significant findings, leaving

various potential threats to the security (i.e., evidenced confidence) of validity claims.

Iterative validation efforts would allow for issues with the definition and measurement

can be identified, examined, and adjusted based on evidence. Thus, all research can be

considered a form of measure validation effort. This iterative process of validation is

exemplified via the development and validation of an experience sampling measure of

media multitasking intensity called the Media Multitasking Intensity Questionnaire (MMTIQ). Five studies provided evidence to evaluate the security (i.e., confidence of claims) of conceptual validity, interpretational or useful validity, criterion validity, generalizable validity across context and populations or measurement invariance, and response process validity.

Finally, the current dissertation involves a theory of media multitasking intensity. This theory of media multitasking intensity was investigated by examining parent-adolescent executive functioning, self-regulation, media multitasking intensity, and stress. Media multitasking was redefined as the perception of the co-occurrence of or interference between two or more tasks, when at least one of these tasks' stimuli is a form of mediated information (i.e., media). Media multitasking intensity refers to four dimensions of media multitasking that vary the intensity or demand of resource allocation: task co-occurrence or interference, task difficulty, task intentionality, and task relevancy. In a one-week longitudinal intensive study of 324 parent-adolescent dyads' media multitasking intensity, the current dissertation found that parents engage in higher media multitasking intensity than adolescents. Increased media multitasking intensity was associated with greater general stress at the end of the week and lower executive functioning and self-regulation capacities. Yet, adolescents with diminished self-regulation engaged in less media multitasking and more intentionally irrelevant media use in comparison to parents, which may be indicative of procrastination. These findings demonstrate the nuance that media multitasking intensity can observe, and illustrate that the nuances of media multitasking warrants further research especially upon the dimensions of intentionality and relevance.

TABLE OF CONTENTS

LIST OF FIGURES

CHAPTER ONE

INTRODUCTION

The current dissertation will first demonstrate the need for new theory development, measurement practices, and validation in the field of media effects research, then propose a theory of media multitasking and media multitasking intensity, detail the construction and validation of a measure of media multitasking intensity, and evaluate the relationship between media multitasking intensity and theoretically relevant constructs.

Rationale

Multifaceted Internet-connected devices have become ubiquitous and essential tools for accomplishing goals in daily life. The implications of this increasingly media-saturated environment have only begun to unfold. Further investigation is necessary to understand if the human mind and body have the abilities to adapt and manage the constant stimulation and new demands that this media ecology has produced (Afifi, Zamanzadeh, Harrison, & Acevedo Callejas, 2017; Levitin, 2014; Rice, Hagen, & Zamanzadeh, 2018). The effects on the physical and mental well-being of a growing population who use and increasingly depend on multi-capable devices to be productive at work, create and maintain connections with family and friends, and to relax when alone is uncertain. The existing work has suggested that the evolving media ecology has created a novel range of possible uses and behaviors with media that vary in the ways they impact well-being (Rice, Zamanzadeh, & Hagen, 2019).

One of these behaviors is media multitasking, which has traditionally been defined as the involvement in two or more tasks (i.e., multitasking), at least one of which is a media task (Xu, Wang, & David, 2016). Media multitasking involves a wide range of behaviors, including switching between media tasks and simultaneously engaging in multiple media tasks, but it also includes switching and simultaneously engaging with media and other tasks such as eating, walking, or speaking face-to-face. Though investigations of media multitasking have flourished in the last decade, much debate remains about the conceptualization of media multitasking as well as the causal process of its effects (Lang & Chrzan, 2015; Van der Schurr, Baumgartner, Sumter, & Valkenburg, 2015; Wang, Cooper, Irwin, & Srivastava, 2015; Yeykelis, Cummings, & Reeves, 2014). The lack of cohesive theorizing about the construct of media multitasking potentially generates the inconsistent and potentially null findings in the field (Jeong & Hwang, 2017; Van der Schurr et al., 2015; Wiradhany & Nieuwenstein, 2017). Another potential explanation for the inconclusive nature of the media multitasking literature include the fragmented ways of measuring media multitasking and the lack of validation efforts for these measures (Fisher & Keene, 2019; Segijn, Xiong, & Duff, 2019). The current dissertation explicates and develops a cohesive theory of media multitasking and its relationship with cognitive, emotional, and social processes. It also reviews the proposal, development, and evidence for validity of a measure of the key attributes of media multitasking. In advancing theory and

measurement, the current dissertation aims to improve the interpretability of empirical findings.

Media Multitasking in the Family

The family system provides a unique context in which to examine the impacts of media multitasking. As media-technology (i.e., devices that connect people to media) has become more ubiquitous, parents and children have also increasingly reported that the uses of these technologies have become a source of anxiety and conflicts within the home (McDaniel & Radensky, 2017; Warren, 2016). Parents or primary caregivers who are digital immigrants (i.e., they are still adopting and assimilating to a digital environment) are posed with a difficult task of successfully mastering technology and socializing children who are digital natives (i.e., immersed in a digital environment and potentially more assimilated), in a media-saturated world. Thus, parents and children may experience the tensions of media mastery (Rice et al., 2018), intra- and interpersonally generating stress within the family system (Afifi et al., 2017). To further complicate the issue, scholarship in the area of child development and media theorizes that the best uses and the most harmful impacts of technology are nuanced. They fluctuate depending on family structure, parental efficacy and attitudes, and socioeconomic background (Jennings, 2017; Nathanson, 2015; Slater, Peters, & Valkenburg, 2016).

Nonetheless, concerns about the effects of media-technology use on children, and adolescents in particular, appear to be warranted. Media multitasking is the most

prevalent among, and harmful for, youth (Pea et al., 2009; Reinecke et al., 2017). Yet, it is also perceived as necessary skills and even vital centerpieces of socioemotional development especially amongst adolescents (boyd, 2014; Rice et al., 2018). Adolescents undergo vital changes, including the process of obtaining autonomy from their parents or primary caregivers and developing their own identity. They also experience immature impulse control, greater emotional intensity, salience of their social networks, and inexperience with self-monitoring and perspective-taking (Steinberg et al., 2017). These developmental characteristics of adolescence may explain why teens more frequently engage in media multitasking and experience more media-related stress than do other ages (Afifi et al., 2017; Baumgartner & Sumter, 2017; Brasel & Gips, 2011; Judd & Kennedy, 2011; Reinecke et al., 2017; Voorveld & van der Goot, 2013). Therefore, after developing a theory, construct, and measure of media multitasking, the current dissertation examines media multitasking amongst parents and adolescents. Adolescents are hypothesized to be a theoretically significant population potentially vulnerable to risky or harmful media multitasking. In comparing adolescent developmental characteristics, media multitasking behaviors by them and their parents, and stress to them and their parents, the current dissertation aims to illuminate the causal processes through which media multitasking creates stress. Importantly, adolescence refers to the general category of a developmental stage, and thus refers to physiological and psychological shifts that are not inherently bound to age. This dissertation involves teenagers, the socially constructed age range of 13-19 as

reflective of the average adolescent experiences, and therefore, uses the terms adolescent and teen interchangeably.

This dissertation will begin with proposing the value of an iterative process of measure validation for theory development and advancement (Chapter 2). Then, commencing in Chapter 3, it will exemplify the value of this iterative process by explicating the development and refinement of the Media Multitasking Intensity Questionnaire (MMTIQ). The iterations of validation efforts in Chapter 3 will result in the proposed theory of media multitasking intensity in Chapter 4. The remainder of the dissertation will describe the most recent iteration in validating the MMTIQ (Chapter 5), explicate the statistical analyses chosen (Chapter 6), present the results (Chapter 7), and finally discuss the implications of these findings for the validity of the MMTIQ and the proposed theory (Chapter 8).

CHAPTER TWO

THEORY DEVELOPMENT, MEASUREMENT PRACTICES, AND

VALIDATION

While new media literature is often criticized for atheoretical research, few have clarified the mechanisms stalling theoretical advancement and proposed paths for theory development. Yet, new media literature is faced with investigating an evolving ecology in which the volume, variety, accessibility, availability, and complexity of media increase. Media multitasking is a nascent area of research in the new media effects literature that involves investigating the impacts of this evolution, which has facilitated the ability to simultaneously engage in multiple media or engage in media along with other daily activities. The media multitasking literature should develop from foundational research within the media effects domain, but has been primarily guided by psychological theories focused on information *processing* which apply to general cognition (van Der Schurr et al., 2015; Wang et al., 2015; Yeykelis, Cummings & Reeves, 2014), with the exception of one *media psychology* theory (i.e., LC4MP; Lang, 2008). Thus, before proposing a theory of media multitasking, it is necessary to explain and describe challenges to theory development, as well as the explain and describe the process of theory development in this dissertation which will depend on the development, iterative calibration, and evidence of validity of a measure of media multitasking. Thus, this chapter will develop the fundamental language and framework in which the resulting dissertation has developed.

Media Effects Research as Context

Over the last 60-70 years, the mass communication and media effects literature has evidenced hundreds of effects with decent effect sizes (Rains, Levine, & Weber, 2018), produced a vast amount of research becoming a cornerstone of the field (Nabi & Oliver, 2009; Potter & Riddle, 2007; Perse & Lambe, 2016; Sparks, 2015). Yet, the field is facing challenges with replication and answering fundamental questions about the nature of the construct of media and the mechanisms of its effects (Lang, 2013; Potter, 2011). Current obstacles in measurement construction and validation practices are proposed to be fundamentally stalling theory development and cohesiveness within the field (Potter 2018; Slater, Peter, & Valkenburg, 2016).

This chapter will first review the evidence of these obstacle to cohesiveness within media effects research, review the paradigmatic theories and practices within measurement that thwart theoretical development, and contend that an *iterative* validation practice in which measures can vary in the *security (i.e.,* degree of credibility or certainty) of their interpretations, uses, and consequences would improve the state of media and new media theory development and research. The term security introduced by Tal (2013) is introduced as a framework for interpreting evidence of validity as it supports and disaffirms hypotheses and the confidence that can be placed into it the presented evidence.

Obstacles to Cohesive Media Effects Research and Theory

Numerous concerns have surfaced in the last decade of meta-syntheses and meta-analyses reviewing the history of media effects research. Many researchers have brought into question the progress of media effects scholarship by critiquing current construction and use of theory, dominant methodologies, and the use of validity processes (Bucy & Tao, 2007; Fishbein & Hornick, 2008; Lang, 2013; Niederdeppe, 2016; Newman & Guggenheim, 2011; Perse & Lambe, 2016; Potter, 2011; Reeves, Yeykelis & Cummings, 2016; Reinecke & Eden, 2017; Reeves, Yeykelis, & Cummings, 2015; Slater, 2004; Slater et al., 2016; Valkenburg, Peter & Walther, 2016). The commonalities in these critiques, which span a decade, suggest that the symptoms of a problematic paradigm have persisted. The following subsections will describe and provide examples of these symptoms: conceptual fragmentation, and ambiguity in measurement and validity, leading to some proposed solutions.

Conceptual Fragmentation

Media and the *processing, presence of, or perception* of mediated messages are fragmented theoretical constructs within the field. There are various definitions for media and the *processing, presence* of*, or *perception* of media that differ in scope and specificity (de Vreese & Neijens, 2016; Fishbein & Hornick, 2008; Greenwald & Leavitt, 1984; Neiderdeppe, 2014; Schneider, Reich, & Reinecke, 2017; Slater, 2004). For instance, these concepts include: media exposure (Slater, 2004), media use (Katz & Foulkes, 1962; Reinecke & Eden, 2017; Valkenburg and Peters, 2016), media consumption (Boukes & Vliegenhart, 2017; Cingel & Krcmar, 2013; Cardoso-Leite,

8

Ludt, Ma, Shawn & Daphne, 2015), media attention (Drew & Weaver, 1990), media experiences (Reeves, Yeykelis, & Cummings, 2016), and media interaction (Reeves & Nass, 1996). These constructs describe ways of *processing* media, forms of media *presence*, or *perceptions* of media. It is clear they are often imbued with assumptions and hypotheses about the mechanisms of media effects and the construct of media. Though, these are rarely expressed.

In the current paradigm of the field, the central construct of *media* is taken for granted despite the widely changing landscape of media. Hence, it is difficult to ascertain if these concepts or constructs (such as media attention and consumption) and their various measures are hypothesizing various ontologies of the same theoretical attributes (targeted quality) or hypothesizing various theoretical attributes. In other words, one can wonder are attention and consumption two distinct attributes of the phenomenon of *media's* effect or are they two hypotheses about the *mechanisms* of media effects? It is even more arduous to determine the degree to which their measures should capture shared variance in order to evaluate the validity of these measures. While it may be attractive to consider that these terminologies are only semantically different whose variations have no empirical impact, these constructs often connotate unstated hypotheses and theories of media, which frame empirical findings. Without stating these theories or hypotheses, the field may conflate constructs that are derived from opposing or competing theories of media effects. The next section will

demonstrate how this conflation extends to the ambiguity in measurement and validation practices.

Ambiguity in Measurement and Validity

The existing role of ambiguity in standards by which the field determines an whether instrument, such as those used to observe media or the *processing, interacting with, or experiencing of* media, deserve the status of measures warrants attention and the consequent intentions of measurement and quantitative structures (Mari, 2013; Mari, Maul, Irribarra, & Wilson, 2017). There is also ambiguity in the features of an instrument of observation that determine whether it is a *valid* measure and often a lack of attempts to *validate* a measure. The consequences of the ambiguities of the purpose of the term measure and the differences between those that are valid and those that are not, create ambiguity in the interpretations of the empirical evidence these instruments (or tools) of observation provide (Maul, 2017). For the purposes of this dissertation, this ambiguity will be exemplified within Media Effects literature.

Measurement and thus validation issues have been identified as problematic within both experimental and survey methodology. First, it is important to begin with the issues in measuring and validating measures of the concept of media. Experimental methodology is often used to examine the effects of a particular medium or the impact of variation in an attribute of mediated content. Lang (2013) and Reeves et al. (2016) discuss the measurement and validation issues that arise in stimuli selection within experimental media research. In their critiques, they both address the issue of "post-

10

hoc" and unstated definitions of media. The lack of conceptualizing, operationalizing, and validating stimuli within experiments as *measures* co-occur. Stimuli are often selected without definitions of media or media content they represent. Instead, they are frequently substantiated through manipulation checks, which involve one or two items asking participants to rate stimuli or simply demonstrate they attended to stimuli by recalling their exposure. This provides one type of evidence of validity which depends on the response process validity and accurate interpretations of participants responses to the manipulation check. These items usually cannot assess stimuli as representative of the larger range of possible stimuli.

In addition to the scarcity of clarity in conceptualizations of media and the subsequent lack of valid claims and generalizations based on the chosen stimuli, there are troubling issues with validating measures of *processing,* being in the *presence* of, and *perceiving* media. Potter (2018) conducted a content analyses of media effects literature between 2010 and 2015 through which he quantifies existing issues with measurement and methodology. He found that the majority of measures (64.8%) are self-report in nature and ask participants to report on mundane behaviors. Self-report measures of media *exposure* (e.g., involving questions about typical media use) have threats to validity because of the various heuristics that could influence response processes and thus invite bias or error into measurement outcomes. Response processes involve the cognitive and communicative mechanisms that allow participants to respond to questions in ways that align with the intended interpretations of those answers. He

also found a dearth of attempts to justify the validity of measures. Less than half (46%) of papers published provided any justification. Of the papers that provided justification, 90% employed citations to previous studies using that measure as evidence of validity. Thus, justification is primary made via face-validity or consensus. Previous scholars have also identified that validation efforts are rare in media effects research (Fishbein & Hornick, 2008; Gentile & Bushman, 2012; Niederdeppe, 2014; Slater 2004). Thus, it seems the complex and vital task of developing a valid measure is undervalued or rather assumed to be too complex for scholars to complete, while the observations and data obtained from these measures are potentially (i.e., without more evidence it is difficult to evaluate) overvalued, misunderstood, and quickly generated. Thus, the conceptual fragmentation of media and the *interaction with, experience of,* and *processing of* media is mirrored with ambiguity in its measurement and validation practices in ways that threaten the existing security, or trusted-value, of the current claims in the literature.

Proposing Solutions

Though change may be slow within media effects literature (as well as across social science fields), one of the obstacles to transcending existing challenges is the lack of concrete actions that could allow junior and senior scholars to overcome them, and the norms that do not invite them to *invest* in these actions. *The explication and demonstration of potential actions and shifts in norms surrounding measurement and theory development that would provide inertia to change are one of the essential purposes of this dissertation.*

Scholars have proposed that media effects research would improve by increasing the role of theory and synthesizing more research (Potter, 2014; Potter & Riddle, 2007;Valkenburg & Peter, 2013), using psychologically relevant attributes (Lang, 2013), considering conditional effects (Valkenburg & Peters, 2013), and combining and comparing methods of measurement (e.g., passive, experimental, and self-report) (Andersen, de Vrees & Albaek, 2016; de Vreese & Neijens, 2016; Jerit et al., 2016; Kobayashi & Boase, 2012; Lang, 2013; Moy & Murphy, 2016; Potter, 2018). Yet, these critiques have not examined the common themes and issues shared by each of these. It is essential to reveal the mechanisms through which the predominant theories of measurement and validation currently guide the construction, validation, and use of measures in empirical research and the lack of its use in theoretical development. The lack of coherence within the media effect's literature is contended to be a consequence of the disconnection between methodology, epistemology, and ontology. Measurement and validation practices have become isolated from theory development. In combination, these schisms have inspired the contemporary absence of a shared framework for interpreting concepts, measures and empirical findings.

Due to the abundant variance in approaches to quantitative methodology and the underestimated value of measurement and validity within published research, the next section aims to provide an overview of theories that have guided standards of measurement and validation in order:

1) demonstrate how they inform current practices

2) provide examples of the harms of these practices

3) and propose new practices

Definitions and Theories of Measurement

Theories of measurement hypothesize or provide definitions of the qualities of an instrument of observation (i.e., a tool used to collect data) that are necessary and sufficient for allotting measurement status. They consequently define the appropriate interpretation, uses, and outcomes of instruments of observations that attain measurement status. Within quantitative Communication scholarship and media effects literature, measurement is most frequently defined as, "the assignment of numerals to objects or events according to a rule," (Stevens, 1946, p. 677). Thus, measures are methods of using numbers to *represent* empirical relations or observational comparisons (Campbell, 1920). This measurement paradigm is influenced by the philosophies of representationalism and operationalism, and supports the existing ambiguity in the definition of measurement and validity, as well as their role in theory.

Bridgman (1927), considered the founder of operationalism, proposed that a construct is equivalent to the set of operations for observing that concept. Therefore, the status of measurement can be awarded to an instrument of observation if a target attribute has been *conceptualized* such that it has been defined, and *operationalized*, such that a set of rules for observing that attribute was identified. Building from an operationalist perspective, Campbell (1920) proposed a tradition in which numbers were applied to represent these rules for observing that attribute (McGrane, 2015).

14

Stevens' (1946) in the same vein proposed a few "scales" or rules of operations for assigning numbers. These scales constrained the valid interpretations of numerical assignment or types of rules for observing an attribute. While these philosophies standardized approaches to measurement development, they did not provide a mechanism for evaluating the quality or accuracy of conceptualizations or operationalizations. Within both philosophies, it was possible for a wide range of tools to constitute measures even if they did not provide meaningful interpretations of a construct (Maul, 2017). The only requirement for measurement status becomes definition of a construct and a tool through which observations can be assigned a numeric value based on one of these scales (Chang, 2010). In this perspective, observations are assumed to be adequate proxies for theoretical attributes (Borsboom, 2006). This encourages the practice of allowing operationalizations or the chosen methodology for observing this attribute to function as a definition. This creates a tautological conundrum. The current standards also allow researcher's biases and assumptions to impact operationalization, recognizing that there is a hypothesis or at least model of the concept/attribute as it relates to other concepts and the mechanisms of its metaphysical existence. The inability to trace these hypotheses, examine when or why they were falsified, creates a possibility for researchers to use each other's' measures and generate empirical findings that are inconsistent with fragmented explanations and no solutions. It becomes impossible to identify what if any portion of the theory motivating a measure was accurate, and what requires revising.

Problems Exemplified

The application of operationalist and representationalist philosophy has contributed to a lack of interpretability and cohesiveness within the theories, measures, and findings of media effects research. Operationalist and representationalist philosophy are unconcerned with ontology (the metaphysical nature of a construct), which is assumed to be unknowable, and justify quantitative labeling of observations as measures of a construct (Chang, 2010; McGrane, 2015). Applied in practice, operationalism and representationalism endorse equating epistemology (i.e., how knowledge is created or what claims can be made) with ontology (i.e., metaphysical existence). Every conceptualization and operationalization in essence identifies a different construct that is assumed to exist in reality. For instance, there would be an assumption that there is a *real* distinction between media exposure, use, attention, interaction, and experience that never requires any justification. The creation of new constructs (or terminology) and new "measures" *without* theory or without comparison to other concepts is justified within this paradigm, reflected in conceptual fragmentation.

This has also manifested via little conceptualization of, or agreement about, the quantitative nature of media and the *processing, presence,* or *perception* of media. Many measures are conceptualized and operationalized to observe the frequency of *processing, presence,* or *perception* of media. Without explicating a theory of the ontology (i.e., metaphysical nature) of media or media *processing, presence,* and

16

perception, these concepts have been operationalized as continuous because they are observed via frequency (de Vreese, & Neijens, 2016; Slater, 2004). Scholars casually make claims that conflate epistemology and ontology, such as "more media exposure leads to …", or "less media consumption of violence creates…." Is it the nature of the media, exposure, or consumption that is continuous, are these equated to frequency because their nature is unidimensional, and which is the causal mechanism? Can there be separate variation in the frequency of media and exposure? These questions appear strange within an operationalist and representationalist framework, yet reveal conceptual gaps. Without having answers to the existing questions, it appears media, as a concept, is manifest and categorical variable, and human-media interactions, exposure, attention or experiences are continuous. These assumptions about the ontology of media have not been developed into testable hypotheses, and are challenged within the existing media environment where the boundaries between media are diminishing. The field has also not justified, but has rather actively debated, whether these frequencies, also metaphorically referred to as "doses," are equivalent within each unit of time (Lang, 2013). Scholars have suggested that each medium has affordances that change the ontology end epistemology of media (Evans, Pearce, Vitak, & Treem, 2016; Rice, Evans, Pearce, Sivunen, Vitak, & Treem, 2017; Treem & Leonardi, 2013). First, media are conceptualized multidimensional, and second, their experiences and therefore observations may vary depending on these other attributes or dimensions. Third, the transformation in media and the growing number of attributes that could be

17

subscribed to media suggest that the concept of media is not universal (i.e., generalizable across time and space). Thus, it may be inappropriate to use the extensive (i.e., additive) nature of time (Campbell, 1920) to contend that media *processing, presence,* or *perception* as attributes of media are unidimensional, continuous, and observable.

Simultaneously the theories of media effects have grown more complex (Lang, 2013) in ways that reflect operationalism and representationalism. In the absence of conceptualizing and operationalizing multidimensional constructs and variables, the models of the effects of media have become more complex including conditional models or the increasing presence of mediated-moderated models (Valkenburg & Peters, 2013). Mediated-moderation models identify attributes of: 1) the medium, 2) the *processing, presence,* or *perception* of the medium, and 3) the individuals who are in the presence of or processing and perceiving this medium as mediators and moderators of the effects of media (Bucy & Tao, 2007; Slater, Peter & Valkenburg, 2016). These models intended to model the causal processes through which media impact individuals also reflect a multidimensional construct because they involve more than one quality of media. The existing lack of expressed hypotheses about the categorical ontology of media, and the use of frequency epistemically create measures of variation in media have justified conceptualizing dimensions as separate constructs. For instance, Social Cognitive Theory involves the mediated-moderation of the effect of media exposure to character's behaviors on imitation of or learned behavior of that character via the

reward-valence of characters. Reward-valence is defined as the ratio of rewards or positive resources that a character obtains, as well as their experiences of punishments or negative consequences. Yet, reward-valence can be a dimension of media *processing, presence,* or *perception,* such that there can be variance in the *processing, presence,* or *perception* of reward-valence within media or it can be modeled as an outcome of media *processing, presence,* or *perception*. When dimensions of a construct are ignored, measures and the empirical findings they support acquire error and bias that impact the validity of claims. In essence, these claims are less secure (i.e., involve more uncertainty) regardless of their p-value because the error exists at the ontological-level which impacts interpretation and thus theory. Without conceptualizing the ontology of media and then the appropriate epistemologies, it is not possible to have informed debates about dimensionality.

Proposed Alternative Theories

Philosophers of measurement have been debating suitable theories of measurement to replace operationalism and representationalism (see McGrane, 2015). Three alternative approaches to measurement will be reviewed and evaluated for their potential value for improving the status of measurement and theory development in the field.

Classical theory. One alternative to the current paradigm is to employ the classical theory of measurement in the social sciences, which defines measurement as, "the estimation or discovery of the ratio of some magnitude of quantitative theoretical

attribute to a unit of the same theoretical attribute," (Michell, 2008). While a complete review of this theory is beyond the current scope, there are a few key theoretical attributes. Thus, in comparison to operationalist and representationalist hypotheses that some measures can use a ratio-scale, the classical theory of measurement argues that instruments of observation must have the properties of being measured on a ratio or interval-level. However, the classical theory of measurement specifies more qualifications that must be met in order for an instrument to be acknowledged as a measure (Humphry, 2013; Michell, 1997).

Thus, it is important to understand the principals of measurement identified in the classical theory of measurement and to differentiate it from the existing paradigm. First, the classical theory of measurement hypothesizes that attributes are ontologically either quantitative or qualitative. Operationalist and representationalists who disregard Truth, therefore begin hypothesizing about a construct from epistemic and methodological levels. In contrast, the classical theory of measurement argues that concepts fundamentally, in reality or in their metaphysical existence, either vary quantitatively or do not. They argue that the status or categorization of an instrument as a measure depends on its quantitative structure, but this is only appropriate for concepts that are also ontologically quantitative. Stated differently, ontology and epistemology must both be quantitative and parallel one another. Thus, classical theorists would question whether media or the *processing, presence,* or *perception* of media have a fundamental unit. Third, because an additive, infinitely divisible set of values, unit, and

thus ratio within the concept must be identifiable, instruments becomes measures if the concepts they observe are unidimensional. A full discussion of unidimensionality is beyond the scope of this dissertation.

The classical theory of measurement challenges the existing operationalist and representationalist paradigm which divorced measurement from reality (Michell, 2004; 2008). The lack of explicated hypotheses about the nature of the concept are critiqued as problematic. Yet, even within the representationalist paradigm, there are few to no Communication constructs that employ a ratio scale unless the involve the unit of time or frequency. This may be an issue of current theories of Communication which have been challenged with defining the scope of communication – i.e., is it intentional, is it linear, is it an experience or an artifact of human meaning-making? Michell (2008) has argued that the use of and assumptions propelling measurement or quantitative structures in the social sciences are pathological. However, there are scholars which have critiqued classical theory as too strict and inappropriate for social sciences in which constructs are *mind-dependent*, or derived from or only observable in relation to socioemotional cognition (Mari, Maul, Ibarra, & Wilson, 2016; Markus & Borsboom, 2013; Maul, 2013). Rather, these scholars have headed the warning of classical theorists, and embraced the approach of hypothesizing an ontology and its relationship with epistemology before and within measurement construction. One particular model that has suited modern social scientists who aim to employ quantitative measures for

constructs that are mind-dependent has been the latent model approaches, concepts are differentiated as manifest and latent (Mari, Maul, Ibarra & Wilson, 2017; Tal, 2013).

The case for the latent model approach. This section will describe developing psychometric-based approaches to measurement known as latent variable models. Latent variable model approaches do not have a particular unifying theory of measure beyond a foundational distinction between a latent and an observable attribute. This approach to measurement was motivated by theories of statistical inference and psychometrics which argued that some concepts are not observable directly (i.e., are not manifest) but are indirectly observable (i.e., are latent) via other indicators (Borsboom, 2008; Borsboom, Mellenbergh & Heerden, 2003). This indirect observation suggests that there should be statistically modeled error at both the level of indicators and the latent construct itself. However, these models do not provide a definition of measurement nor do they provide a clear standard for measurement status.

Rather, latent models provide a potential solution to Michell and the classical theorists challenges that social scientists do not have testable hypotheses about the ontology of their constructs. Latent models allow scholars to statistically model and therefore to hypothesize the ontology of the observations and their relationship to the ontological variation in the attribute. In other words, observation often bounded by epistemology and methodology can be theorized as being a distinct entity from the target attribute (i.e., main concept – e.g., media). The latent variable model also allows scholars to embrace the ontology of social scientific phenomenon as unobservable

22

(Markus & Borsboom, 2013) and mind-dependent (Maul, 2013). The target attribute (which is latent), as well as the observed indicators, may be unordered, ordered categorically, or continuously structured (Markus & Borsboom, 2013). Using statistical models (probability distributions) and parameter estimations, observations can be evaluated on the degree to which they are the best predictors of an existing unobserved theoretical attribute (Tal, 2013; 2016). Thus, it requires attention to the relationship between ontology and epistemology, and could help motivate changes.

As opposed to latent attributes, observable attributes have the qualities of determination, causal isolation, and equivalent cardinality (Borsboom, 2008; Borsboom et al., 2003). Underlying these three qualities is the hypothesis that there are patterns in the data that should reflect a causal *structure* that is giving rise to observations. Within this framework, *observable variables* require determination or the epistemological certainty that their data structures (observed variation) are deterministically caused by the structure and variation of the attribute. There should be no sources of error or noise in observing this data. The second requirement is that the "position" (i.e., ontological observation) of a given object within the variation of the attribute has only one given score – causal isolation. Every outcome of the measure has only one clear interpretable spot (score) in the variation of the attribute (e.g., two people with the score of 3 demonstrated the exact same response patterns or variation within the attribute). This is also referred to as local independence. Causal isolation excludes cases of multidimensionality where it is possible that there are multiple ways to obtain the same

23

"score." Finally, equivalent cardinality refers to equivalent variation in observation and number of "positions" (or values) that a person or object could have on that theoretical attribute.

Latent variables are those that do not fit these three assumptions. Applying latent variable theory, it is possible to model which assumptions are not met and use appropriate statistical models. If there is a lack of determination (1:1) in the causal process, then a probability model such as the Rasch model may be useful. A violation of causal isolation may suggest a need for a multidimensional deterministic model. If cardinal equivalence is violated such that there are fewer positions or values on the latent trait than the variation in the observations of the latent trait, then a unidimensional deterministic IRT model applying Guttman scales might be useful (Borsboom, 2005). For instance, if scholars can only obtain an ordinal observation or data structure, but believe the attribute is continuous, Item-Response Theory uses logistical regressions to locate a participant within a continuum or space even without the continuous observation. Latent variable models include four relationships based on the two dimensions of categorical/continuous, and observed/latent attribute such as: factor models (continuous observed and continuous attribute), item response theory (categorical observed and continuous attribute), latent class analyses (categorical observed and categorical attribute), and latent profile analyses (continuous observed and categorical attribute).

Because latent variable models, similar to all statistical models, are blind to the quality of data they are estimated upon, they are only as valuable as the methodology or measures and used to provide them with observations. While this approach allows scholars to identify and evaluate the hypotheses about a construct, because it is possible to compare various measurement models' fit to the data, it still likes a useful definition of measurement.

Intersubjectivity of measurement. Mari, Carbone, Giordani, and Petri (2017) and Maul, Mari, and Wilson (2019) have begun to identify a unifying definition of measurement that could apply across disciplines. Importantly, measurement is treated as a process that was purposefully developed to obtain addition information about an object (Mari et al., 2017). Thus, measurement involves theories of the property, as well as theories of the object and the theory of the values of the property (i.e., a theory of the resulting information that is an outcome of the measure. This approach aims to explicate the black box of measurement which transform observations of the properties of an object into information or values of that property. Thus, the process and result of measurement entails two dimensions: objectivity and intersubjectivity. Objectivity is defined as the *extent* to which the resulting value provides information about the variation in the property of the object of measurement exclusively. Intersubjectivity, on the other hand, is the *extent* to which the resulting values have a shared reference points such that they can be interpreted consistently across contexts (i.e., places and times).

By comparing the existing theories, the current section with explicate how
redefining measurement and revealing its inner mechanisms can modify the value or
role of measures. In operationalist and representationalist paradigms where the property
of the object (the ontology) is undefined, such that the only important value of a
measure is that it transforms observation into a numeric output that could be
mathematically and statistically examined. The purpose of the measure is to
functionally represent observations but there is a lack of necessary justification via
evidence nor a clarification of the theory or hypotheses that are embedded in these
measures. Thus, they define measurement as the process of property evaluation whose
results are credibly documented. A measure can vary in the degree to which it can
traceably relate (i.e., create testable chains of causal relationships) properties of an
object with values of a property.

The next section will review the prevailing theories of validity and validation
practices to identify the ways these have manifested into problems within the literature.
In the final section of this chapter, solutions for measurement and validity will be
combined and explicated.

Theories of Validity

Validity has been a major concern within media effects literature (Fishbein &
Hornick, 2008; Niederdeppe, 2014; Neiderdeppe, 2016; Potter, 2018; Slater 2004) due
to the fragmented application and the dearth of validation processes. The most
commonly employed justifications for validity include reasoning via face-validity or

consensus-based validity, reliability as validity, and methodological mind-independence as accuracy (Potter, 2018). These justifications indirectly, within empirical investigations, value norms amongst experts, psychometrics, and mind-independence. However, the primary theory of validity is construct validity, which is reserved for special empirical investigations: validation efforts. Within the current scope, it is necessary to only briefly address the commonly applied approaches and then focus on essential critiques of the theory of construct validity.

Common Justifications of Validity

Face-validity. Face-validity justifies the validity of a measure based on self-evaluation (often by expert) of its set of operations as rationally and intuitively being appropriate for the construct. This is principally associated with consensus in which the face-validity of a measure, or its "face value," in a study leads others to use it without a validation process, because another scholar has published an empirical investigation using it. The previous significant findings become enough evidence. This approach, however, hides the assumptions and theories about both the ontology of the construct and its epistemology – the process of observing the construct. It also perpetuates a validation process that does not test hypotheses about the appropriate epistemology and methodology for observing the construct. Face-validity values the experts' opinions and harkens to operationalist thinking (Borsboom, 2005); it is justifiable enough to say that a measure is valid because it operationalizes a conceptualized construct.

Reliability. Another common justification of validity is via reliability or an estimate of internal consistency as the main indicator of validity – this approach prizes psychometrics rather than expertise (Cronbach & Meehl, 1955). Reliability is often misunderstood and offered as evidence of validity (Campbell, 1960) and calculated by using a Cronbach's (1951) alpha coefficient. This includes a more recently pattern of fitting a measurement model (i.e., a latent model) and then a subsequent composite reliability. The use of Cronbach's (1951) alpha has remained the standard for measurement validation in empirical investigations even after critiques of alpha have flourished across fields (Cortina, 1993; Green, Lissitz, & Mulaik, 1977; Raykov, 2001; Rodriguez & Maeda, 2006; Schmitt, 1996; Sijtsma, 2009). High alpha coefficients are frequently interpreted as evidence that the items (in survey research) or indicators (in behavioral or physiological research) in a measure are "measuring the same thing" (Millsapp, 2007). Yet, the statistic of reliability cannot provide information about the test validity because it is population-specific. It also does not provide information about the cause of the variation, only that the items or indicators vary together (Borsboom, 2005). Thus, alpha can only vaguely address the average level of interrelatedness amongst indicators or items, though there may be spurious relationships (Nunally, 1978). Reliable measures are not necessarily accurate but they also do not necessarily reflect variation in the attribute (causal processes).

Methodologically mind-independent. Another common justification for validity involves methodologically removing mind-dependence; measures are justified

28

as more valid if human perception is not involved. This has become particularly popular within media effects research. With the growing criticisms that self-report measures (as traditionally constructed) are problematic due to biases and limitations of human recall, Potter (2018) and de Vreese and Neijens (2016) in addition to other scholars have noted the necessity for passive measurement (i.e., naturalistic behavioral observation). While the goal of removing sources of error associated with human perception is sensible, these methods removes error that was incurred due to methodology, but does not necessarily demonstrate a clear conceptualization or more ontological relationships. As big data and computational social sciences grow in number, it is valuable to acknowledge that methodological mind-independence can improve accuracy, which is a necessary but not sufficient condition of validity (Rice, 1990).

Behavioral measures (in experiments or natural observations) have the advantage of increased accuracy for frequencies, but can also enable avoiding theorizing about the construct itself. For instance, in a content analysis of screen-shots of computer-screen behavior Yeykelis et al. (2014) evaluated task-switching or switches between mediated content on a laptop; the assumption is made that opening a window on a screen was synonymous with media experience and switches include any switch between windows. The accuracy increased in capturing a switch between windows passively does not solve the ontological and theoretical questions about the construct of media multitasking. For instance, this behavioral measure implies that task-switching is a categorical and unidimensional construct which can be measured by counting

29

frequencies of task-switches, or adding up instances of task-switches as the fundamental unit. While the tool increased accuracy, there are additional steps that would be necessary to assess the quality of the measure and the evidence that provides security in validity claims. Otherwise, methodologically mind-independent justifications of validity also harken to operationalist assumptions that operations or methodologies with less error necessarily create valid measures.

Construct Validity: The Predominant Theory

The predominant *validation* practices are grounded in the theory of construct validity, which contends there are types of evidence or arguments that can be made for validity (Cronbach & Meehl, 1955; Messick, 1989). In comparison to the previous approaches, which are used to justify claims of validity, construct validity theory provides a framework for *validating* a measure. Construct validity theory defines validity as the *degree* to which a measure captures what it intends to capture (Borsboom et al., 2009; Sircei, 2009). This evidence-based approach to establishing validity beneficially evaluates validity on a continuum; however, it is difficult to implement and in application has demonstrated complications. The ambiguity of intentions allows for post-hoc alterations to the targeted property of the measure to increase validity – similar to the operationalist and representationalist approach to measurement. In this approach, scholars still do not address the metaphysical properties of the theoretical attribute (Borsboom et al., 2009) and construct validity theory does not provide standards for sufficient evidence. There are generally four categories of validity to evidence:

divergent, *convergent, nomological (criterion related),* and *content validity*. Based on evidence in these four categories, it is possible to "validate" a measure.

Though over time, evidence often accumulates that both supports and contradicts construct validity across these types of validity (Markus & Borsboom, 2013). In these occasions, construct validity has limited explanatory power. These forms of evidence of validity can be conceptualized as types of observations that represent qualities of a measure that would be more securely valid. Yet, the ways to interpret the values obtained within each type of evidence of validity it is unclear. Thus, the value or the results of the obtaining evidence of validity does not have clear interpretations. This manifests such that when conducting meta-analysis in an area of research, it is difficult to ascertain whether the measurement is problematic or the theorized relations between constructs are incorrect. Even in proposing new measures it is often necessary to demonstrate evidence of construct validity by relationship with a previous measure which may have been problematic. Thus, scholars have two options: 1) to still use measures that have at least once evidenced validity or 2) to keep developing new measures. Neither is necessarily conducive to increasing the number of valid measures in the field or the cohesive interpretation of empirical findings.

Construct validity theory values association or correlation between measures as the most essential evidence of validity. While its tenets are not inherently problematic, the ambiguity within its principles make it possible for it to become questionable in practice. Construct validity theory privileges identifying a construct within a network of

31

constructs. This includes nomological networks or theorized relationships between constructs (concepts that represent theoretical attributes targeted in a measure) that function in a law-like manner (Cronbach & Meehl, 1955). Yet, there are few constructs that are clearly related in law-like relationships *a priori* to measurement development. The ambiguity in recognizing nomological networks can suggest that validity is an unobtainable goal or can lead scholars to the conflation of criterion or predictive validity with evidence of a law-like relationship. This issue extends to the practice of correlating measures amongst similar and dissimilar constructs – convergent and divergent validity. This is problematic because it is possible that these are statistical artifacts of the nature of social science research (Maul, 2017). The predictive validity of an attribute also does not provide evidence that the measure's information is equivalent across contexts or populations (Millsap, 2007).

Exemplifying the crux of the problem. Construct validity in practice has become problematic in that the norm has become that the evidence for validity can be gathered within one study, *validating the measure indefinitely.* These measures are not re-examined later; rather they are used to evidence criterion and convergent/divergent validity for other constructs, potentially muddling the evidence. This validation approach ignores that attribute may not exist similarly or universally across cultures and ecological changes. As measures continue to develop within these nomological networks and as scholars conceptualize a measure's *intentions*, it becomes challenging to clarify the scope of the measured attribute and explain the reasons these measures did

or did not evidence a relationship. Is the conceptualization problematic? Are the measures lacking in some form of validity? Or are the theories guiding these nomological networks simply incorrect? Furthermore, construct validity theory often disincentivizes scholars to try to find unifying and coherent theories about theoretical attributes. The fragmented state of constructs and measures creates a need for more valid measures and thus provides more options for criterion and convergent/divergent validity. These are again divorced from theory development. Theories remain the same while more measures are development to investigate the same attribute and its relationships with other attributes. However, these ignore theoretical changes that are implied in these new measures. *It is rare that theories (e.g., social cognitive theory or cultivation theory) are revised because of a discovered lack of validity in measures.* Instead, it is possible that empirical findings continue to inform theories, which involve measures with similar issues or sources of error. These measures may have demonstrated convergent and divergent validity, but lack interpretability. (RETURN TO THIS WHEN LESS TIRED)

The validation process within construct validity has a limited role in identifying and assessing the causal processes in which concepts such as media exposure become observable. Fikkers, Piotrowski and Valkenburg (2017), and Seger and Potts (2017) are recent examples of uses of construct validity that should be lauded for making *validation* efforts. Yet, they also demonstrate how the concept of media remains ignored within the current paradigm of validity.

Validation in Iteration: Securing Evidence and Claims of Validity

Validation is a process of calibrating a tool to observe variation in an attribute in meaningful, interpretable, and accurate ways (Borsboom, Mellenberg, & Heerden, 2004; Kane, 2006; Mislevy, 2009; Tal, 2016). Calibration of an instrument across time, contexts, and populations can allow scholars to evidence the degree of confidence placed upon the claims about the conceptualization, observations or methodology, measurement procedures, and measurement outcomes (i.e., uses and consequences of the measures). Validity is an evaluation of hypotheses about the mechanisms and thus interpretations of observations about the concept/property, object to which the property belongs, methodology, procedures, and measurands (i.e., information obtained). In other words, validation extends beyond the efforts of construct validity which focuses the purpose of validation efforts on the values produced as a result of measurement. Validation efforts evaluate the security of the theory and measure of a construct as useful, interpretable, and accurate by identifying, evaluating, and then reconstructing hypotheses. Evidence of usefulness, interpretability, and traceability (Maul et al., 2019) are necessary and inform *security* – or levels of credibility (Tal, 2016). The validity of an instrument or measure within specific subpopulations, contexts, as well as in it provides credible information about the construct reflects the quality of the measure (i.e., objectivity and intersubjectivity), but also the theory that informs the measure.

Therefore, validation of a measure involves evaluating: 1) the appropriate interpretation of a measure, 2) the value of the measure as it relates to theoretical

development, and 3) testing if the interpretation and value of a measure remains applicable across time, contexts, and populations. Validation is proposed to include iterations between theory, measure construction, response process validity, psychometric validity, and statistical or criterion and nomological validity using evidence as feedback for evaluating all previous steps and phases. Although the process is presented linearly in the text and in Figure 1, in practice these steps co-occur (see Chapter 3).

Validation Steps

The process would begin with (1) identifying a theoretical construct, which involves labeling a phenomenon, or a target attribute. It is necessary to clarify and define the attribute of interest, its overlap with existing attributes, and its unique properties. This step should situate an attribute definitionally, and begin the process of evaluating its usefulness and consequences. It also involves hypothesizing attributes internal and external to the objects of interest that may be antecedents or outcomes of variation of the target attribute. Validation begins by identifying and creating hypotheses about the target attribute. Thus, it links measurement to theory.
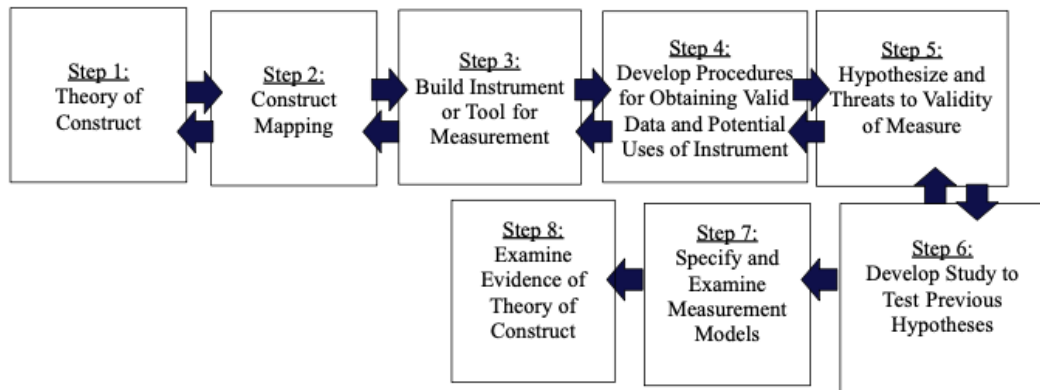
Figure 1. The iterative process of securing evidence of validity claims.

Then, in the second step (2), it is necessary to engage in a process of construct mapping or the process of conceptualizing construct and recognizing limits to observing variation in the construct across the potential methods available. This includes leverage and acknowledging potential limits. This mapping is not a representative mapping like operationalists and representationalists. Rather, construct mapping is essentially a theory yielding step – it requires a theoretical model of the concept. It involves a casual mapping of the relationship between variation in observations of a property of an object in relation to the variation in the property (Mari et al., 2017; Markus & Borsboom, 2013; Maul et al., 2019). Stated differently, hypotheses are developed about the best (i.e., most accurate, interpretable, and meaningful) observable outcomes or predictors of that attribute, and about the strength of relationship or causal relationship between the two (Tal, 2016). The observations that may be the best predictors of the attribute may be easiest to capture via differ methodology.

In step three (3), hypotheses about the attribute and the variation in an object on that attribute, are extended to include potential interactions with epistemological and methodological hypotheses. The observed variation in the attribute within an object which would provide information about the variation of an attribute are examined might interact with epistemic, methodological, and between-concept relationships. Step (3) begins with an assessment of epistemology – what includes the current methods or tools that could be used to observe this construct and the types of claims that are possible given each of these tools' relationships with the ontology of the attribute. In step (3) multiple methods of observation can be assessed as more or less likely to secure claims by evaluating potential threats to observing variation in property values that are causally determined (i.e., traceable) or variation in the property across objects. Finally, in this step, other attributes of the object or of the object's context may be theorized to predict or explain the variation in the attribute or to potentially interact with the method for observing the attribute. In other words, step (3) involves investigating potential threats to intersubjectivity (Mari et al., 2017; Maul et al., 2019).

After hypothesizing relationships between ontology, epistemology, and methodologies for observing the property, it is possible in step (4) to choose a methodology that is appropriate, identify the appropriate claims and begin developing a measure. In step (4), the previous step is reiterated but is specified to the specific methodology and begins development of the physical instrument. It involves designing procedures for obtaining valid data or the creation of the instrument to be used for

obtaining observations as it relates to the construct map original developed. The purpose is to identify procedures for the administration of the measure that protect the relationship between the variation within the observations of a property of an object, and the variation within values of the property.

In the fifth step (5), appropriate hypotheses are derived about the appropriate conditions, structure, or format of the instrument measuring the attribute. This step identifies potential threats to the objectivity and intersubjectivity of the measure by hypothesizing about the relationship between the theory of attribute within objects and its relationship to context. Thus, step (5) involves hypotheses about threats to objectivity such as definitional and instrumental uncertainty (Maul et al., 2017). It may also include identifying other sociocognitive threats to response validity. This step, therefore, guides appropriate conclusions for the uses of scores or outcomes from the instrument.

Thus, the majority of the steps of validation involve hypothesizing about the ontology and the epistemology of the construct, as well as its relationships with other constructs, observations, the instrument for observation, procedures of administration, participation or using the instrument, and interpretations of the instrument. The process identified here does not reject all of the previous logic or justification for validity or validation; rather it embraces them and requires a clearer identification of hypotheses, testing of hypotheses, and iteration in evaluation and development. This could allow all

empirical investigations to contribute to the validation process, and allow the theories that guide measurement development in media effects research to become clearer.

The previous steps facilitate the evaluation of measurement quality to evidence support for a theory about the metaphysics of an attribute, which then become the foundation for investigating the relationship between the attributes and other attributes: theory development. The sixth (6) step involves designing a study or investigation through which it would be possible to test hypotheses about the measurement model, the veracity of the observations, the format of administering the measures, and the theory of the construct across populations or contexts of theoretical value. Thus step (6) provides of evidence for the traceability of the information of a measure and the security with which the validity of the claims based on the outcomes/values of the measure should be asserted (Maul et al., 2019). This can include nuancing claims within a population or context in the situation where there is lower security in the evidence obtained, suggesting that the information quality of the measurement requires further validation and calibration in order to reduce the calibration n uncertainty (Maul et al., 2019). This may involve the evaluation of response process validity, which examines the degree to causal mechanisms that create variation in the attribute within the object (Ercikan & Pellegrino, 2017; Zumbo & Chan, 2014).

The seventh (7) step involves using psychometrics to statistically evaluate models of the relationship between observations and the construct (Almond, Steinberg, & Mislevy, 2002; Borsboom et al., 2004; Mislevy,1994; Mislevy, Steinberg, &

Almond, 2003). The evaluation at the end of the seventh step should influence the

security in claims derived from the findings of step eight. If the previous steps were

taken, even a lower level of *security* (i.e., credibility or confidence) in claims could still

justify interpreting and sharing results. As opposed to current practices, the current

iterative process aims to contextualize empirical findings and information ascertained

from measures within the degree of *security* about the quality of the measure and the

validity of the claims that can be made from it.

Finally, the eighth (8) step involves evaluating evidence for the theory of the

attribute and its relationship to other attributes. In the final step, the security of the

quality and validity of the measure and thus security or confidence placed in the

empirical findings can be reported. In the case that a measure is less secure in quality or

validity, this would not unqualify or invalidate a finding as uninformative. Rather, the

finding of less security in the quality (i.e., objectivity and subjectivity) or validity of

claims based on the measure could inform scholars about theoretically relevant

populations that experience the construct differently. Findings of insecurity can advance

theory or become opportunities for theory development. The value of the iterative

process of validation and securing of evidence of the quality of the measure, validity,

and theory of constructs, would be most apparent when it would be possible to trace the

relationships between theories. Unlike the current status of media effects which

involves fragmented and unclear definitions of constructs, the iterative process of

validation in theory development would ground theory about media effects in theories

of media. Over time, it may be possible to develop standards of security for measures or a "security interval" for the quality and validity of measures through which theories have developed. Thus, much like a meta-analysis of effect sizes it would become possible to engage in meta-security analyses of the evidence of measure quality and validity as well as theory development.

Summary

Validation efforts can secure evidence for the claims about the value and usefulness of a measure. When engaged via *iteratively* validation efforts do not only provide information about the quality of measures and the interpretability of measures, but also provide information about the security of previous empirical findings and theory that guided the development of the measure. The remainder of this dissertation involves a demonstration of the value of the iterative process of *validating* a measure for theoretical development within media multitasking literature.

CHAPTER THREE

ITERATIVE THEORY DEVELOPMENT AND MEASUREMENT

VALIDATION IN MEDIA MULTITASKING

The current chapter will describe the development and investigations of the initial iterations of a new measure of media multitasking (the media multitasking intensity questionnaire, MMTIQ). Collectively, four studies were conducted with various populations of adolescents, young adults, and adults (i.e., parents and those without children) to evaluate the questionnaire's epistemological and methodological interpretability, value, and veracity in reflecting the ontology of media multitasking. Thus, the evolution of the measure was motivated by investigations into the threats to validity within the structure, procedures surrounding, content, and uses of the instrument. These create security in the quality of the measure including objectivity and subjectivity described in Chapter 2. The purpose of Chapter 3 is to review the contribution of each iteration and inquiry to the theoretical, epistemological, and methodological understanding of media multitasking. Each iteration, though revealing the flaws and insufficient validity of each respective version of the measure, provided insights that inform the current questionnaire and theory of media multitasking which are examined in this dissertation.

Overview

Study 1 involved conceptualizing media multitasking and its attributes of interest, item development, selection, and investigating measurement invariance via a within-subjects 2x2 factorial design (difficulty of task (i.e., demanding or easy) and type of context (i.e., work or leisure)). **Study 2** aimed to replicate study 1 and examined measurement invariance when participants answered the selected items for only one context via a between-subjects 2x2 factorial design (difficulty of task (i.e., demanding or easy) and type of context (i.e., work or leisure)). Both study 1 and 2 examined criterion validity and provide evidence of some security in the validity claims developed from the model. Due to the lack of measurement invariance in study 2, **Study 3** involved cognitive interviews within 16 focus groups to examine potentially issues with response process validity and identify comprehension of vernacular or phrasing of items, response options, and formatting that would increase validity. The findings in study 3 led to revisions of problematic items, response options, and formatting. Finally, **Study 4** examined the effects of these changes on the measurement model and response process validity in a theoretically relevant population of parents and adolescents. Its purpose was to evidence whether revisions to the measure were successful. The remainder of the dissertation (**Study 5)** will further develop the concept of media multitasking intensity by explicating a theory, and then evaluate the measurement model over time in a theoretically relevant population of adolescents and parents as well as the significant predictors and effects of media multitasking intensity.

Study 1: Is Media Multitasking *Intensity* a Stable Trait?

Goal

Study 1 involved the initial theory-driven item development and selection, examination of measurement invariance across theory-relevant contexts, and investigation of the share variance between the new instrument with the Media Multitasking Index (MMI; Ophir et al., 2009) as well as the relationships between media multitasking *intensity* and general stress

Theoretical Beginnings

The project began as a search for a theoretically-based and more accurate and valid self-report measure of media multitasking. The first iteration was fundamentally shaped by both a) psychological theories of cognition and multitasking such as limited capacity, cognitive bottlenecking, and threaded cognition models, and b) nascent specific theories of media multitasking. In 2015, existing inconsistencies in the media multitasking literature led scholars to examine heterogeneity in media multitasking. Scholars proposed dimensions that predicted the propensity of various media combinations (Wang et al., 2015) and predicted their subsequent success (Lang & Chrzan, 2015).

These new theoretical approaches to media multitasking predict a relationship between features of tasks, cognitive load, and media multitasking behaviors. Wang, Irwin, Cooper and Srivastava (2015) provided an argument for 11 cognitive dimensions of media multitasking with four higher order categories: task relations, task inputs, task outputs, and user differences. More specifically, these dimensions include: task

hierarchy, task switch, task relevance, shared modality, task contiguity, information modality, information flow, emotional content, behavioral responses, time pressure, and user differences. These dimensions were guided by the "law of less work," a theoretical principle which contends that humans aim to conserve resources (i.e., taking the easy way out). Simultaneously, Lang and Chrzan (2015) reviewed the existing research of media multitasking and proposed that task difficulty and overlap in information processing required by tasks determine the adaptiveness of media multitasking. This included the accuracy and efficiency of completing both media tasks. In other words, task combinations requiring similar types of processing (i.e., translating stimuli into meaningful and actionable information) such as interpreting language or solving problems, that were both difficult, would be less successful when the user was multitasking. In combination, they suggest that human cognitive limits, and law of less work, can be used to develop a model of strategic media multitasking (Ralph & Smilek, 2017).

Though these media multitasking theories had unifying themes, they lacked a theoretically based measure that would allow them to evaluate the various dimensions they proposed. Instead, most of the studies applying these theories relied on the Media Multitasking Index (MMI; Ophir et al., 2009) or similar measures, which are extensive or additive (i.e., there are units that are observable as described in Chapter 2), and therefore unidimensional measures of media multitasking. Therefore, Study 1 aimed to develop a construct and questionnaire that integrated these theories in order to develop

45

and evaluate indicators of the multiple dimensions of media multitasking. In order to do so, it was necessary to also examine indicators and dimensions that were most essential and theoretically valuable. Wang et al. (2015) provide a comprehensive and complex model of media multitasking that was difficult to interpret or apply, while Lang and Chrzan (2016) provided a simple model of media multitasking that lacked some important attributes. Therefore, the goal of developing the new construct and instrument was to parsimoniously but more accurately capture the heterogeneity in media multitasking.

Designing a parsimonious construct and instrument required choosing attributes that would provide theoretically valuable information for the field of media multitasking. Cognitive bottlenecking, limited capacity models, threaded cognition, and theories of multitasking, predict that the rate and frequency that a person switches between stimuli strain cognitive resources (Lang & Chrzan, 2015; Wang et al., 2015; Yeykelis et al., 2014). In these theories, multitasking, either defined as attempting to process two stimuli concurrently or rapidly switching between stimuli, maximally strains and fatigues resources, leading to human error. These theories suggested that *types and rates of multitasking* are fundamental attributes of media multitasking, and are fundamentally theoretically valuable because of their impact on cognitive load (Salvucci, Taatgen, & Borst, 2009). Similarly, media multitasking models suggested that features of the tasks involved are also valuable because they affect cognitive load. *Task difficulty* was identified as one of the fundamental aspects of media multitasking in

the theoretical and empirical work of Wang et al. (2015) and Lang and Chrzan (2016).

Integrating these two perspectives, difficulty of the tasks was theorized to involve the

information flow, behavior inputs, as well as the novelty, skill required, and duration

required for the processing of the information. These indicators were identified within

Wang et al. (2015) and Lang and Chrzan (2016) as contributing to cognitive load.

Finally, both a general and specific theories had one dimension that seemed to connect

divisions of attention and task attributes: relevancy of tasks to one another, and

relevancy of tasks to goals. Threaded cognition models suggest that the inefficiency and

ineffectiveness of multitasking is not only an outcome of difficulty of tasks or the

conflicts between tasks for resources, but also their relationship to one another

(relevance, task contiguity, hierarchy) and a higher-order goal (Wang et al., 2015;

Yeykelis et al., 2014). *Task relevance to a higher-order goal* could diminish the

challenge of integrating multiple sources of information. *Task relevance to each other*

likely causes less competition for cognitive resources and allows for more information

synthesis. These dimensions appeared to be the most fundamental attributes of media

multitasking within both multitasking and media multitasking research.

The commonalities in the theoretical significance of these dimensions suggested

that the most valuable attributes of media multitasking affect cognitive load. The

interaction between these dimensions, reflect that the variation in media multitasking

leads thwarts the possibility of identifying a single and interpretable unit of media

multitasking. Thus, the frequency or the extent of media multitasking would be

47

uninterpretable due to unobserved heterogeneity. Media multitasking *intensity* examines

this possibility for variation in each dimension within itself, but more importantly as it

relates to the other dimensions to change the value of media multitasking. In other

words, *intensity* is a multiplicative *not* additive outcome of measurement because it is

also multiplicative in nature (i.e., ontology and metaphysical properties). The only way

to interpret any instance of media multitasking is to account for all dimensions of it.

This is similar to the nature and measurement of temperature, volume, weight, and

distance which are outcomes of multiple dimensions. Therefore, this multidimensional

construct of media multitasking suggests that each dimension provides information

about the context or the person, but cannot conclusively provide information about the

attribute of *media multitasking* – the same way the knowing the rate at which a car is

driven cannot provide information about the distance it being driven without also

having information about time. Forty items were generated to create a Guttman-like

scale (Zimmerman, Williams, Zumbo, & Ross, 2005) such that the items in addition to

the response options are intended to observe variation in the media multitasking for

each of these dimensions and subdimensions. These items were placed on a construct

map, which expressed the predicted causal relationships between values on these items

and the higher-order dimensions with values or locations on the continuous construct of

media multitasking intensity (See Figure 2).

High Intensity Media Multitasking	High frequency of task switching or overlapping activities, high number of activities, goals, and devices that are high in difficulty and but low in relevancy within goals or activities and low in intentionality.
Moderate Intensity Media Multitasking	Some frequency of task switching or overlapping activities, a high number of activities that are of moderate difficulty and but higher relevancy within goals or activities and moderate to high in intentionality.
Low Intensity Media Multitasking	Low frequency of task switching, low number of activities, goals and devices with higher relevancy between them and less complexity and novelty.
Single Tasking	Distinct media and non-media goals or habits completed serially without any other simultaneous activities

Figure 2. Construct map of media multitasking in relation to items and dimensions

Dimensions. Media multitasking *intensity* was theorized to involve four fundamental dimensions: *divisions of attention, difficulty of tasks, task relevancy to goals,* and *relevancy amongst tasks*. In theoretically modeling these dimensions, it was

49

predicted that some dimensions would be more strongly related to one another than others, yet each provided distinct information. Item generation involved first conceptualizing aspects of each of these dimensions. *Divisions of attention* was defined as the ways in which attention becomes shared or split among multiple tasks particularly it involved two forms of multitasking: *task-switching* (i.e., the shifting of attention) and *dual tasking* (i.e., simultaneous engagement in two or more tasks). This multitasking could occur both within and between devices. *Difficulty of tasks* referred to the degree of perceived challenge created by each task in which an individual engages. Thus, media tasks could range in their difficulty, in addition to the difficulty of the general goal or other tasks in which participants engaged. *Task relevancy to goals* was defined as the perceived degree to which tasks were completed with the intention of accomplishing a goal. In other words, this was a top-down or effortful application of cognitive resources to each task that was purposeful because it related to a relevant goal. Finally, *relevancy amongst tasks* examined the degree to which the media tasks as well other tasks involved in media multitasking shared a purpose.

Media multitasking *intensity* was predicted to depend on *task characteristics and goals*, but these two attributes of experience (as opposed to time) vary greatly depending on context. This created a theoretical and epistemological concern. The existing research using measures such as the MMI (Ophir et al., 2009) had thus far examined media multitasking behaviors (e.g., dual tasking with multiple media) in typical or most recent weeks. This epistemological approach has become criticized for

50

creating noisy data and lacking validity due to limitations to recall (Chein, Wilmer, & Sherman, 2017). By relying on typical media use, these measures are also imbued with the hypothesis that media multitasking behaviors are trait-like and stable (Ophir et al., 2009). Finally, the media combinations observed within the MMI primary include those that are used for leisure such as music and television, instant messaging and talking on a phone call, or surfing the internet and listening to podcasts, but include few potential combinations with work- or school-related because software like Microsoft Word, Excel, or Adobe PDF, all categorized as computer-based applications. Therefore, that measure was not equipped to observe work-relevant media multitasking, and thus observes less relevancy between media in this context. In addition to relevancy, context could have significant interactions with the difficulty of the media tasks. Therefore, Study 1 aimed to investigate media multitasking intensity dimensions in work and leisure contexts.

Cognitive Interviews

Following item generation and construct mapping, the items were investigated for clarity and response process validity (Ercikan & Pellegrino, 2017; Zumbo & Chan, 2014), which is the validation of participants' understanding of survey items as well as cognitive processes leading to their chosen response option, via cognitive interviews. Fifteen (15) think-aloud interviews were used to further refine the items. Participants "thought out loud" digitally. After being presented with each item, they were asked to type out their thought processes (with no editing) and then to provide answers to the

questions about media multitasking intensity. This included he extent to which they typically divide their attention amongst media and other tasks, as well as the typical difficulty and relevance of these media in two contexts: work and leisure.

These participants highlighted a few areas of confusion. First, participants noted that they were unsure what behaviors should be interpreted as engaging in a media activity. The questionnaire was therefore improved by including verbiage which clearly defines media and media activities. Even within this small group of participants, this feedback highlighted a theoretical failure on the part of media multitasking scholars. The literature lacked a clear definition of the unit of analysis of media multitasking: what *is* a media task? In the measure used in Study 1, media were defined as methods of obtaining or sharing information that do not occur via face-to-face contact with another human being. The various ways of consuming this information are different activities – reading an article or status someone posted is different from scrolling through an Instagram feed of images. These activities were ultimately listed for participants (see Table 1) in the investigation of the first iteration of the measure – this feature remains in the final questionnaire validated in this dissertation.

Media Activities
Listening to music offline or online
Listening to non-Music Audio online (e.g., podcasts) or via radio (e.g., news)
Audio/Video Call of any kind with or without Conversation (including voicemails)
Viewing or engaging Internet sites that are not social media (e.g., online shopping, an online school website, Google, banking, yelping, or Expedia)
Viewing/scrolling through social media content
Creating or responding to social media content (e.g., commenting, reacting,or posting)
Taking photos and videos (not on or for social media purposes)
Sending a text or message on any platform (e.g., DMing on Instagram or Snapchat, or Facebook messenger)
Checking a notification on any device (but not viewing the content on an app)
Watching television or movies on a Telvevision
Watching online videos (such as Netflix or Hulu or Amazon Prime or YouTube)
Playing video games on a console or computer
Playing video games on a phone (e.g., bejeweled or candy crush)
Using any software or non-social media applications on a computer or smartphone such as Word, Excel, Photoshop, etc.
Writing emails on any device
Reading a newspaper/magazine(online or digitally)
Wirting in a journal or notebook
Reading a Book (Digitally or Physically)

Table 1. List of media activities provided to participants

Second, participants reported that they often forgot about their specific media

uses throughout the questionnaire. This was a more difficult problem to solve.

However, Study 1 aimed to solve this problem by enforcing a period in which

participants were required to review the options of media activities and their answers first to assess if they completed it correctly and then to provide a total number of media in which they engaged. The purpose of this revision was to at least improve their memory and to communicate to participants that they needed to remember each of the media in which they report they engaged. This highlighted, an often-overlooked assumption within self-report methodology: participants are aware that the information they report will get used in the future. Participants are often aiming to ascertain the purpose for and the uses of the questionnaire and may manage where to allocate their cognitive resource accordingly. They may aim to control conversations or knowledge about the attribute (i.e., media multitasking) and may adjust their response providing invalid responses (Clairmont, 2020; Clairmont, Wolf, & Maul, 2019). Because surveys are tools for communication, it is important for scholars to increase transparency about the role of each question as this informs participants' responses.

Third, after examining the responses of participants to questions about *relevancy of tasks to goals* and *relevancy between tasks* a pattern emerged in which relevancy of tasks to goals and to each other may share more information than was conceptually expected. In other words, it was in people's interpretation of these concepts they may epistemically be similar. Thus, they were theorized to both contribute to a higher order dimensions of relevancy.

Finally, several interviewees noted they found it confusing to report on their "typical" hour of media multitasking across contexts of work and leisure. Even after

adjusting for this issue, they continue to grapple with rating tasks on a continuum from easy to difficulty. This posed an epistemic limitation on the ability to observe difficulty continuously within the current methodology (i.e., self-report). They also suggested that within each work and leisure context, difficulty and ease were relative creating a problem within the dimension of *difficulty of tasks*. Therefore, it was expected that difficulty and context interact in significant manners that currently threaten the potential security of the interpretability (i.e., valid claims) that can be made with the measure of media multitasking intensity.

Study 1 was therefore designed in reaction to these findings. Particularly, Study 1 's 2x2 within-subjects factorial design examined the effect of difficulty (high = demanding, low = easy), and context (work, leisure) on the measurement model (i.e., psychometric evaluation) of media multitasking intensity and its invariance, as well as variation in the remaining dimensions of media multitasking intensity: divisions of attention and relevancy. It aimed to improve response process validity by shortening the period of time for reflection to the most recent hour. The dimension of difficulty was therefore manipulated into high or low demand, and work or leisure, and observed and measured as an aspect of the context.

Study 1 Rationale

Study 1 began validating the initial MMTIQ by testing for measurement invariance, evaluating theoretically-grounded relationships between the dimensions of the model, and exploring the impact of context. Based on the integration of the limited

capacity model, threaded cognition theory, and the cognitive bottlenecking model, it is

deduced that when engaging in difficult goals and activities, people are less likely to

divide their attention and more likely to engage in activities that are relevant to their

goals to save cognitive resources (Lang & Chrzan, 2016; Wang et al., 2015). Though

not previously theorized, it was expected that contexts and types of tasks may moderate

the degree to which people aim to conserve resources. For instance, when goals or tasks

are more extrinsically motivated and involve deadlines with potential consequences for

delays, people are expected to be less likely to split their attention and engage in

irrelevant activities. Thus, work-related that were difficulty tasks were less likely to lead

to divisions of attention and more likely to include relevant tasks than leisure tasks that

were easy. This would align with the law of less work which suggests that people aim to

save their resources, but would suggest the "law of less work" is also context-

dependent. In other words, people do not aim to save or store resources across all

conditions, otherwise they would not prefer stimulation or challenge.

Thus, building from cognitive load models, media multitasking intensity was

informed by the models of flow and stress. The model of the state of flow

(Csikszentmihalyi, 2000) and stress (Lazarus & Folkman, 1984), the allocation of

resources and perceptions of allocating the resources may fundamentally depend on the

difficulty of the tasks. Within ego-depletion models (Baumeister & Vohs, 2017),

resources are conserved as they become perceived and experienced as limited. The

existing model of flow (Csikszentmihalyi, 2000) argues that humans enjoy tasks when

there is synchronicity between the difficulty of the task and the resources they have. Thus, frustration and stress are the outcomes of having a task that is perceived as too difficult because of the ratio of demand to resources (Lazarus, 1993). In contrast, boredom, which is also an undesirable state, involves the perception of too little difficulty because of the ratio of resources to demand. Thus, the perception of limited resources, impacts the degree to which people consciously choose to engage in behavior that would allow them to reserve their resources. In combination with the preference for synchronous or equal ratios of difficulty and resource, this suggests that engaging in an easy task would actually increase the desire to engage in divisions to attention and irrelevant media tasks.

Thus, greater divisions of attention and lesser task relevancy should be expected in easy leisure contexts, while fewer divisions of attention and greater task relevancy should be observed in demanding working contexts. Though dimensions of media multitasking intensity were labelled as divisions of intention and relevancy such that each would indicate the presence of these attributes, in measuring the attributes relevancy was measured in reverse as *irrelevancy* in order to maintain a consistent direction of relationship between dimensions and overall intensity. While conceptually relevancy is not problematic, epistemically the limitations of statistical tools and their consequent interpretations required that the dimensions become measured such that variation in the dimensions have similar relationships with the higher order attribute of intensity. Therefore, it was hypothesized that:

Hypothesis 1: While media multitasking with difficult work-related tasks,

people should report (a) less divisions of attention and (b) less *irrelevancy*

among activities than while media multitasking with easy leisure activities.

Research Question 1: Will difficult leisure and easy work differ in levels of

divisions of attention and irrelevancy?

Additionally, the average relationship between the dimensions of divisions of

attention and relevancy are expected to vary depending on context, but this becomes a

point of contention. Previous research investigating has suggested that the best

explanation of media multitasking frequency is that divisions of attention are a

consequence of deficient abilities to filter irrelevant information (Ophir et al., 2009).

This would suggest that more divisions of attention would be associated with less

relevant media use or that divisions of attention and irrelevancy would be positively

related. However, if the "law of less work", model of flow, and stress are supported,

then, the relationship between divisions of attention and irrelevancy of tasks would be

negative – the more divisions of one's attention would only occur because these tasks

were rather relevant to one another. Thus, Ophir et al. (2009)'s finding would allow be

true of easy leisure contexts, which may be the context for which the Media

Multitasking Index (MMI) is most accurate. Thus, it was hypothesized that:

Hypothesis 2: There will be a a) positive association between divisions of

attention and the irrelevancy among tasks under the condition of easy leisure,

but b) these will be a negative association between divisions of attention and the irrelevancy of tasks under the condition of demanding work.

Research Question 2: What is the relationship between media multitasking intensity in various contexts and self-report media multitasking via the Media Multitasking Index?

Furthermore, based on models of flow and stress, which examine the relationship between perceived demand of one's tasks and environment as well as of one's resources, divisions of attention and relevancy of tasks were predicted to have a relationship with stress. Divisions of attention and less relevancy (or more irrelevancy) increase the demand placed on cognitive resources within a context. However, their effect on the perceptions of the demand-to-resource ratio are dependent on context. In a demanding work context, engaging in more divisions of attention and less relevant tasks should be related to the perception of a ratio where demand surpasses resources, and thus results in stress. In an easy leisure context, engaging in more divisions of attention and less relevant tasks should be related to the perception of a ratio in which resources surpass demand. Therefore, it was predicted that:

Hypothesis 3: Engaging in more a) divisions of attention and b) less irrelevant tasks during demanding work condition would be related to stress in a demanding work condition, but will have no relation under an easy leisure condition.

Study 1 Method

59

Participants. Participants ($n = 500$) were recruited through Amazon's Mechanical Turk (MTurk) and compensated US $5.50 for completing the survey. Participants ranged from 19 to 70 years old and were on average 35.41 years old ($SD = 10.39$), with almost an equal number of females and males (50.4% female). The sample consisted of African Americans (7.8%), Asian Americans (6.4%), European Americans (76.3%), Latina/o Americans (4.6%), Multi-Ethnic Americans (3.0%), Native Americans (1.6%), and Other (0.4%) (unknown: 0.2%). Respondents came from various regions of the country, including the Midwestern United States (e.g., Ohio, 18.8%), Northeastern United States (e.g., New York, 21.8%), Plain States (e.g., Kansas, 2.8%), Southern United States (e.g., Mississippi, 28.6%), Southwestern United States (e.g., Texas, 9.7%), and Western United States (e.g., California, 18.1%) (unknown: 0.2%). Participants classified themselves as working class (29.8%), lower middle class (24.8%), middle class (40.1%), upper middle class (5.0%), and upper class (0.2%) (unknown: 0.2%).

Procedure. Once participants agreed to participating to the study and accepted the HIT on MTurk, they each completed the MMTIQ in four contexts followed by the MMI, a measure of general stress, and a few demographic measures such as occupation and age. First, this initial version of the MMTIQ included the self-reported number of media activities and goals, frequency of task switching and dual tasking, and the degree of relevancy between activities and goals. In a within-subjects 2 x 2 (e.g., demand of task (high, low) and type of context (work, leisure)) factorial design, participants each

answered the MMTIQ for four distinct recent hours in which the participant was engaging in at least one easy work-related, demanding work-related, easy leisure, and demanding leisure activity. These contextual prompts were presented to participants in a randomized order. This allowed for researchers to examine measurement invariance and explore these potential theoretical dimensions of media multitasking behaviors. In order to prompt participants' memory, for each context they were asked to report at least one specific activity that they remember engaging in during this period. Then, they were prompted with a list of 17 other media activities (see Table 1) that they could have engaged in during this hour. The MMI, then measures of stress, and then demographic items including gender, age, location, and occupation were employed for further validation of the measure.

Data cleaning. There was less than 5.0% missing data overall and thus no imputation methods were used. Before analyzing the data, boxplots and histograms were used to search for univariate outliers. Only one outlier appeared across all items, and was thus deleted from the dataset. This case was one of an elementary school teacher who reported extremely high scores across all items. Other outliers were not removed from the data and no transformations were preformed because kurtosis and skewness were not significantly different from normality.

Measures.

Perceived general stress. An adapted 5-item variation on the 10-item scale by Cohen, Kamarack, and Mermelstein (1983) was used for perceived general stress (α =

.86). These items were rated on a 5-point scale from (0) never, (1) almost never, (2) sometimes, (3) fairly often, (4) very often. Items include statements such as, "In the last month, how often have you felt nervous and "stressed"?"

*Frequent media multitasking***.** Frequent media multitasking was measured by the Media Multitasking Index (MMI) as developed by Ophir et al. (2009) includes asking participants to self-report the hours they spent with 12 media: television, music, non-music audio, online videos, computer-based applications, video games, newspaper, books, instant messaging, text messaging, social media, and phone calls. Following this self-report measure, participants are asked to report how often they used each of these media with the other form of media (e.g., video games and music) with response options (0) Never (.33) Sometimes (.66) Often (1) Very Often. This creates as 12 X 11 matrix, which together with the previous questions, amounts to 156 items. Because this is an index of behaviors, it does not have an alpha reliability; this index is used to calculate the proportion of time spent with each of these media as primary (media that was the main focus) and secondary media (background or additional media) out of total amount of time spent on media.

Media Multitasking Intensity. Media multitasking intensity was measured by the Media Multitasking Intensity Questionnaire (MMTIQ) it involved the total number of activities engaged in, how often participants shifted their attention, how often participants open and minimized windows of activities on one device, how often there were two to three media activities occurring simultaneously as the degree to which the

activities they reported engaging in seemed relevant to one another, the degree to which these activities shared a similar topic, the degree to which they were unconnected to goals, and the degree to which the activities were related to goals These items were rated on a Likert-type scale from (1) never to (4) always.

Analyses.

Conceptual model. The conceptual model specified that the items in the Media Multitasking Intensity Questionnaire (MMTIQ) included two dimensions or factors (see Figure 3). The first factor, *divisions of attention*, included the following items: the total number of activities engaged in, how often participants shifted their attention, how often participants open and minimized windows of activities on one device, and how often there were two to three media activities occurring simultaneously. These items were rated on a Likert-type scale from (1) never to (4) always. The second factor, *relevancy*, included items that measured the degree to which the activities they reported engaging in seemed relevant to one another, the degree to which these activities shared a similar topic, the degree to which they were unconnected to goals, and the degree to which the activities were related to goals. These items were rated on a Likert- type scale from (1) never to (4) always for all items except the "unconnected to goals" items because they were reverse coded. By specifying this measurement first, it was possible to evaluate whether the hypothesized relationship between the items and the latent variable was evidenced.
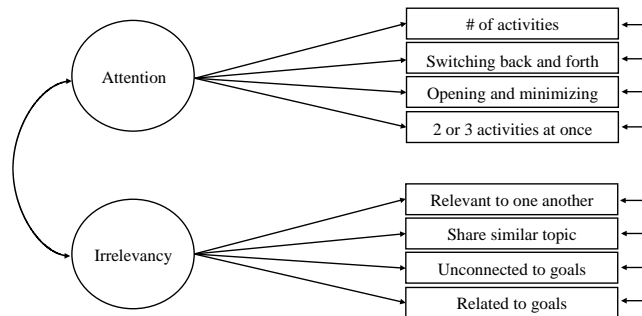
Figure 3. Conceptualized two-factor structure of the MMTIQ

Measurement invariance. Measurement invariance is an attribute of a

questionnaire or instrument, when the instrument obtains the same measurement model

across groups of people or contexts. This demonstrates that the measure has more

universal usage and interpretation. A Multigroup Confirmatory Factor Analysis

(MGCFA) an analysis to examine whether attributes (i.e., means, variance, slopes,

factor loadings) of a measure shifted between multiple groups of observations was

conducted using maximum likelihood (ML) estimation and variance-covariance

matrices within M*plus* 7.4 (Muthén & Muthén, 1998-2017). MGCFA was employed to

establish the measurement invariance of the MMTIQ between contexts. First, the

measurement model was established within each context separately, and then in the two

most distinct contexts (i.e., easy leisure and difficult work), which were essential to the

hypotheses. Second, invariance was established by estimating a series of models that

sequentially increased the number of model constraints between the groups (i.e., factor

loadings, slopes, means, and variance) and then were compared across goodness of fit statistics (Brown, 2014; Cheung & Rensvold, 2002). In line with recommended practices, the model was next estimated between the groups separately. Comparison of model fit indices across the two group models were then evaluated to establish a basis for measurement invariance.

Preliminary analyses (confirming the two-factor solution). Before constraining the model between groups, the conceptual model in Figure 3 was first identified and estimated in the larger population to establish that the two-factor solution demonstrated adequate fit. Thus, as a first step, the psychometric properties of the MMTIQ were investigated to verify the hypothesized two-factor structure (i.e., divisions of attention and irrelevancy among activities) within each of the four combinations and conditions of difficulty and context (e.g., work vs. leisure) and then in the sample combined. In line with recommendations of Brown (2014) and Fabrigar et al. (1999), model fit indices were used to evaluate if the two-factor solution demonstrated adequate fit. Then, the model was also fit separately in the two most dissimilar conditions, easy leisure and demanding work. These are both preliminary analyses that are necessary conditions before beginning multigroup confirmatory analyses because they would then support that the theorized structure to the measure is generally appropriate.

Multiple group confirmatory analysis (model constraints). Following these preliminary analyses, a series of three constrained models were evaluated for

65

appropriate model fit to establish measurement invariance (Brown, 2014; Byrne, 2008; Widamen et al., 2010). First, the model was constrained to hold equal *factor structures* or equal form such that the numbers of factors as well as indicator-factor loadings are alike between easy leisure and demanding work, as well as easy work and demanding leisure. Second, "weak factorial invariance" or *metric invariance* was tested by restricting factor loadings to be equal across easy leisure and demanding work, and then across easy work and demanding leisure. This establishes whether these latent constructs have the same relationship to the items in each of these contexts (i.e., is the traceability of the observations on items to the latent attribute equivalent in all contexts?). Third, "strong factorial invariance" or *scalar invariance* was established by restricting both factor loadings to be equal and intercepts to be equal across conditions. This restriction could potentially identify differential item functioning. Differential item functioning occurs when item's measurement properties change due to either a characteristic of a respondent such as (e.g., native English speaker vs. non-native English speaker) or due to the context in which items are answered. Differential item functioning threatens the validity of a measure by affecting latent means and slopes. Once these three hierarchical steps were taken, *structural invariance* was pursued by constraining factor variance to be equal and co-variance to be equal. Once equivalence of factor variance and covariance were supported, then factor means were compared.

  Model evaluation and fit statistics. Brown (2014) recommends evaluating models across three major indices of fit that can provide different perspectives about the

model's goodness of fit. Absolute fit indices, such as the chi-square of model fit and standardized root-mean-square residual (SRMR) assessed the extent to which the overall model replicated the observed covariance-variance matrix. Though a non-significant chi-square of model fit is considered one of the best metrics of fit, it is sensitive to sample size and thus, is commonly expected to become significant with large samples (Fabrigar, Wegner, MacCallum, & Strahan, 1999). Unlike the chi-square, the SRMR is not sensitive to sample size. Adequate fit values are below .08, but good fit values are equal to or less than .06 (Hu & Bentler, 1999). Comparative fit indices such as the comparative fit index (CFI) and Tucker-Lewis index (TLI) were considered to indicate fit if their values exceeded.95 (Brown, 2014). Finally, a parsimony index, the root-mean-square error of approximation (RMSEA) value, was evaluated and considered acceptable at .08 or lower but is best when the value is .06 or lower (Brown, 2014; Browne and Cudeck, 1989).

Additionally, nested models were then compared using nested significance tests such as changes in chi-square and CFI. Because chi-square differences are sensitive to large samples such as the one in this study, changes in CFI were mainly used to evaluate differences in model fit (Cheung & Rensvold, 2002; Kline, 2015). As recommended by Cheung and Rensvold (2002), differences are considered acceptable such that they are not significantly worsening model fit if they are equivalent or below .01 (ΔCFI < .01).

Results

Descriptive statistics. The covariances, correlations, and means of the observed data in each conditions of the within-subjects experiment (i.e., easy leisure as compared to demanding work) are shown in Tables 2, 3 and 4. The correlations among items and means are consistently larger for demanding work than for easy leisure.

Table 2

Study 1 Covariance-Variance Matrix Among Variables in 2-Factor Model Demanding Work

Variables	1	2	3	4	5	6	7	8
1. Total Activities	9.19							
2. Switching between Activities	1.09	.72						
3. Opening and Minimizing	1.21	.53	.82					
4. 2-3 Media at Once	1.20	.49	.52	.75				
5. Relevant Activities	-.08	.05	-.07	-.03	1.41			
6. Share Topic	-.06	.00	-.10	-.05	1.13	1.25		
7. Unconnected to Goals	.24	.10	.06	.09	.79	.65	1.32	
8. Related to Goals	-.03	.01	-.10	-.04	1.09	.96	.79	1.33

Table 3

Study 1 Covariance-Variance Matrix Among Variables in 2-Factor Model Easy Leisure

Variables	1	2	3	4	5	6	7	8
1. Total Activities	14.94							
2. Switching between Activities	1.65	.85						
3. Opening and Minimizing	1.79	.57	.89					
4. 2-3 Media at Once	1.66	.61	.58	.88				
5. Relevant Activities	.57	.24	.17	.20	.99			
6. Share Topic	.14	.15	.08	.12	.72	.87		
7. Unconnected to Goals	.37	.24	.20	.23	.36	.29	1.12	
8. Related to Goals	.17	.18	.08	.10	.53	.48	.51	.96

Table 4

Study 1 means, Standard Deviations, and Correlations Among Variables in 2-Factor Model Easy Leisure and Demanding Work

Variables	1	2	3	4	5	6	7	8
1. Total Activities		.42	.44	.46	.49	.40	.43	.50
2. Switching between Activities	.46		.69	.66	.05	.00	.11	.01
3. Opening and Minimizing	.49	.66		.66	-.06	-.10	.06	-.09
4. 2-3 Media at Once	.46	.71	.65		-.03	-.05	.09	-.04
5. Relevant Activities	.15	.26	.18	.21		.85	.58	.80
6. Share Topic	.04	.18	.09	.13	.77		.50	.75
7. Unconnected to Goals	.09	.25	.20	.23	.34	.29		.60
8. Related to Goals	.05	.20	.09	.11	.54	.52	.49	
M for Easy Leisure	5.27	2.43	2.23	2.35	3.07	2.93	2.28	2.81
SD for Easy Leisure	3.87	.92	.86	.94	1.00	.93	1.06	.98
M for Demanding Work	3.75	1.91	1.88	1.83	2.55	2.52	2.03	2.47
SD for Demanding Work	3.03	.85	.91	.87	1.19	1.12	1.15	1.15

Note: demanding leisure appears on the top right and easy leisure appears on the bottom left.

Preliminary analyses: Confirming the two-factor solution. The confirmatory factor analysis (CFA) of the model depicted in Figure 3 was estimated in the overall sample, and demonstrated adequate fit across indices, χ^2 (19) = 164.93, $p < .001$, CFI = .96, TLI = .95, RMSEA = .09 [.07 .10], SRMR = .04. However, modification indices indicated that there were focal strains in the solution and areas of ill fit. For instance, the covariance between the error terms of the items "activities seemed relevant" and "activities shared a topic," had a modification index of 87.58 and a standardized expected parameter change (EPC) of above 2.00, which suggested that the model should be specified to allow the residuals between the two items to correlate freely. After including the correlated residual between the items, the two-factor solution for the overall sample produced appropriate and significantly improved model fit, χ^2 (18) =

80.46, $p < .001$, CFI = .98, TLI = .98, RMSEA = .06 [.05, .07], SRMR = .04. It is

presumed that these items correlate above and beyond the shared latent factor because

these two items are worded similarly. This new model can be seen in Figure 4. The two

latent variables of *divisions of attention* and *irrelevancy* demonstrated a moderate

significant correlation, r = .60, $p < .001$.
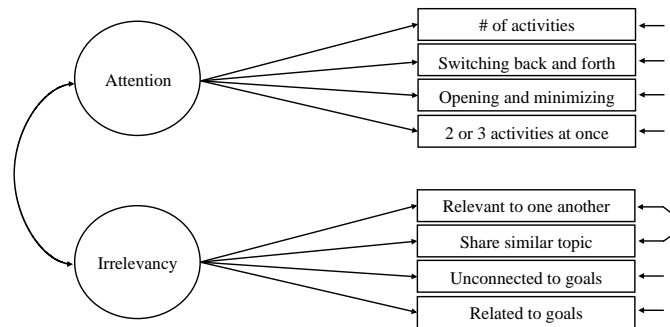
Figure 4. Modified conceptual two-factor model of the MMTIQ including

correlated residuals.

In order to determine whether claim of measurement invariance was possible,

this two-factor solution with a correlated residual was estimated in each of the four sub-

samples. The two-factor solution CFA in the sub-sample of easy leisure demonstrated

adequate fit, χ_2 (13) = 27.04 $p < .001$, CFI = 0.99, TLI = 0.97, RMSEA = 0.05 [.02,

.07], SRMR = .03. Similarly, the CFA in the sub-sample of demanding work exhibited

appropriate model fit, χ_2 (13) = 15.47, $p < .001$, CFI = 0.99, TLI = 0.98, RMSEA = 0.02

[.00, .05], SRMR = .02. Finally, confirmatory factor analysis of the two-factor

conceptual model also demonstrated adequate fit in both the easy work, χ_2 (13) = 22.66,

p < .001, CFI = 0.99 TLI = 0.96, RMSEA = 0.04 [.01, .06], SRMR = .03, and

demanding leisure conditions, χ_2 (13) = 25.06, p < .001, CFI = 0.99, TLI = 0.98,

RMSEA = 0.04 [.02, .07], SRMR = .03. Therefore, these indicated that the model fit

well within each of the conditions separately, and revealed that testing for measurement

invariance was justified.

Multiple group confirmatory factor analysis: Measurement invariance.

After establishing adequate fit in the two most distinct sub-samples (i.e., easy leisure

and demanding work), the MGCFA was conducted sequentially. Please refer to Table 4.

First, the model fit was appropriate across easy leisure and demanding work for

configural invariance. There was no change in CFI (ΔCFI $\leq$ .01). Thus, we continued to

the next level of model constraints. Next, the model was estimated for *metric*

invariance, which also demonstrated adequate model fit across indices and also had a

ΔCFI less than .01. Third, the model was estimated after restricting equal factor loading

and intercepts for *scalar invariance*, and again demonstrated appropriate fit. The ΔCFI

was again below the cutoff value. Thus, the model was considered to exhibit strong

measurement invariance. Similarly, when the *invariance of factor variances and*

covariances was tested, the model ΔCFI did not surpass the permissible amount of

change. Thus, the model demonstrated the strictest or fullest invariance.

Table 5

Model Fit Statistics for Demanding Work and Easy Leisure Conditions Study 1

	χ^2	*df*	$\Delta\chi^2$	Δdf	RMSEA [90% CI]	SRMR	CFI	ΔCFI
Single Group Solutions								
Overall Sample	36.505	13	-		.043 [.027, .059]	.024	.994	-
Easy Leisure	27.043	13	-		.047 [.021, .072]	.030	.993	-
Demanding Work	15.466	13	-		.020 [.000, .051]	.021	.999	-
Measurement Invariance								
Configural	42.509	26	6.004		.036 [.014, .055]	.026	.996	-.002
Metric (equal loadings)	68.172	31	25.663	5	.049 [.033, .065]	.049	.991	-.005
Scalar (equal loadings and intercepts)	88.053	36	19.881	5	.054 [.040, .068]	.050	.987	-.004
Population Heterogeneity								
Equal Factor Variance	119.923	39	31.879	3	.065 [.052, .068]	.112	.980	-.007

Hypothesis 1. In support of the first hypothesis, *latent means* of *divisions of attention and irrelevancy* were significantly higher when media multitasking with easy leisure activities than when media multitasking with demanding work activities by a standardized unit of .65 in divisions of attention and by a standardized unit of .43 for the irrelevancy between activities and goals. When engaging in easy leisure tasks, participants were more likely to engage in more divisions of attention and less relevant

tasks than when engaging in demanding work tasks. Demanding work involved less off-task or unrelated media multitasking than easy leisure.

Research question 1. Using multigroup confirmatory factor analysis (MGCFA), the media multitasking intensity questionnaire was compared between the contexts of easy work and demanding leisure; remember that it was theoretically unclear about how they would differ from one another, and thus stated as a research question. The two-factor model demonstrates strict measurement invariance, including equivalent factor variances (see Table 6). Easy work and demanding leisure demonstrated measurement invariance or adequate fit when constrained for factor loadings and intercepts, as well as, when constrained for equivalent factor variance. The measure functioned similarly in these two conditions. Both conditions demonstrated no relationship between *divisions of attention* and *irrelevancy* of tasks, and both had no significant difference between the latent means on the factor of *irrelevancy* of items ($p < .001$). However, easy work and demanding leisure significantly differed in the latent means of *divisions of attention* by a standardized unit of .61, $p < .001$. On average, there were more divisions of attention during easy work than during demanding leisure. These findings reflect that the difficulty of tasks on average predict the degree to which people divide attention.

Table 6

Model Fit Statistics for Easy Work and Demanding Leisure Conditions Study 1

	χ2	df	Δχ2	Δ df	RMSEA [90% CI]	SRMR	CFI	ΔCFI
Single Group Solutions								
Overall Sample	27.363	13	-		.035 [.018, .053]	.025	.996	-
Easy Leisure	22.664	13	-		.038 [.005, .064]	.032	.995	-
Demanding Work	25.056	13	-		.043 [.016, .068]	.026	.994	-
Measurement Invariance								
Configural	47.721	26	18.358		.041 [.022, .059]	.029	.994	-.002
Metric (equal loadings)	53.381	31	5.660	5	.038 [.020, .055]	.036	.994	-.000
Scalar (equal loadings and intercepts)	58.021	36	4.632	5	.038 [.017, .051]	.038	.994	-.000
Population Heterogeneity								
Equal Factor Variance	69.649	39	11.628	3	.040 [.024, .054]	.069	.992	-.002

Hypothesis 2. Hypothesis three predicted that divisions of attention would be driven by distraction and therefore the factor would be corelated with irrelevancy. However, there was only partial support for hypothesis three. For easy leisure, divisions of attention and irrelevancy were significantly correlated at, $r = .27, p < .001$. However, divisions of attention and irrelevancy had no significant relationship for the remaining three contexts of demanding work, easy work, or demanding leisure. Multitasking was only related to engaging in unrelated tasks when people work on engaging in easy and

leisure tasks. Otherwise, the division of attention was not predictive of engaging in more or less relevant tasks.

Research question 2. Research question two explored which of the scores on the media multitasking intensity questionnaire in any of the four contexts would correlate with the Media Multitasking Index (MMI). There was no relationship in any of the four conditions between scores on the MMTIQ and response on the MMI. Neither divisions of attention nor relevancy of tasks in these four contexts predicted or related to self-reported dual tasking with multiple in the last week. This may suggest there are other forms of error or heuristics that lead to responses on the MMI. The value of the MMI or the construct it measures may be related to the presence or absence of these errors, but this finding in combination with the inconsistencies in existing literature undermine the current interpretation of the MMI.

Hypothesis 3. It was predicted that increased divisions of attention and irrelevancy would have the most significant relationship with stress during demanding work. After confirming measurement invariance across all four contexts, each context was correlated with the 5-item factor of general stress. There was evidence of partial support of this hypothesis. Increased divisions of attention ($b = 0.11$, $p < .001$) and irrelevancy ($b = .15$, $p < .01$) during demanding work contexts were significantly predictive of general stress. Irrelevancy only significantly predicted general stress under demanding work, but when engaging in easy work, easy leisure or demanding leisure; thus, supporting the second hypothesis. However, divisions of attention were more

75

predictive of general stress during demanding leisure ($b = .20, p < .001$) and easy work ($b = .19, p < .01$). This was not predicted. Divisions of attention and irrelevancy had no significant relationships with general stress under easy leisure contexts.

Discussion

Study 1 indicated that the media multitasking intensity instrument (MMTIQ) had a two-factor structure that includes division of attention and irrelevancy of activities. This version of the measure demonstrated measurement invariance across the four contexts: demanding work, easy leisure, easy work, and demanding leisure. Divisions of attention were measured by the number of activities participants engaged in, the frequency in which they switched between activities, the frequency of opening and minimizing engagement in activities, and the frequency of overlap between two or three activities. Irrelevancy of activities and goals were measured by the degree to which activities were relevant to one another, shared a topic, were unconnected to goals, and were related to goals. Findings suggested that the measure is invariant across the four contexts.

Moreover, the measure provided unique insights into media multitasking across behaviors that support the integration of general theories of multitasking and specific theories of media multitasking. Findings from Study 1 supported Wang et al.'s (2015) law of less work and Lang and Chrzan's (2015) dimensions, which placed cognitive load as an essential factor in predicting the frequency and success of media multitasking. People reported more frequently splitting attention between more

irrelevant media during easy leisure than during demanding work. When people are doing a simple task for self-fulfillment or relaxation in their own free time, they report engaging in more intense media multitasking. While relevancy did not differ between easy work and demanding leisure, there were also more attention divisions during easy work as opposed to demanding leisure. Therefore, dividing one's attention with media was more common under easy conditions, but the likelihood that those media are irrelevant depends on context.

While easy conditions in general involved more multitasking, easy leisure was the only condition in which divisions of attention were significantly related with irrelevant media use. This finding suggests that under the easy leisure condition, people reported more frequently engaging in distracted media multitasking – unintentional and irrelevant. However, there was no relationship between divisions of attention and irrelevancy of activities during demanding work, easy work, or demanding leisure. It is particularly interesting that there is a lack of relationship between divisions of attention and irrelevant tasks in easy work, where there is a relationship in easy leisure. This suggests that divisions of attention during easy work involve a greater range of tasks some of which may be relevant, such as switching between emails and phone calls. While the ease of a task may impact the degree to which people divide their attention, it is the type of context whether leisure or work that impacts whether these divisions are with irrelevant media.

Finally, the stress or well-being impacts of media multitasking depends on contexts. For instance, engaging in unrelated tasks was only predictive of greater general stress during demanding work. In other words, distracting engaging in unrelated tasks when one's task is challenging and has consequences if left incomplete, leads to greater nervousness, fatigue, and frustration. The same is true for greater divisions of attention during demanding work. The increased cognitive load which likely leads to more error or less efficiency increases daily and general stress. Divisions of attention are also predictive of stress during demanding leisure or challenging tasks with consequences for one's personal pleasure as well as easy work or simple tasks with consequentiality for one's job or position. In these cases, greater cognitive load was also stressful. Interestingly, the relevancy of these tasks had no impact on stress. This potentially reflects the use of irrelevant media to reduce stress and the potential stress of engaging in multiple relevant tasks. However, neither relevancy nor divisions of attention predicted stress during easy leisure. In this context, intense media multitasking was unrelated with general stress. This suggests that the types of tasks and goals involved in media multitasking impact it's likelihood to contribute to general stress. Future investigations should consider context in examining media multitasking behaviors and considering their impact on general well-being, academic or work performance, and other metrics of success.

Study 2: The Stability of MMTIQ

Goal

Study 2's purpose was to re-evaluate the measurement model found in Study 1 by only presenting the selected items that demonstrated good fit and measurement invariance. In addition, it sought to understand if measurement invariance was achievable when participants only completed the scale for one context as opposed to all four. Therefore, it involved a 2x2 difficulty (easy vs. demanding) x context (leisure vs. work) between-subjects factorial design investigating measurement invariance and further examining whether the MMTIQ demonstrated theorized relationships with other constructs.

Study 2 Rationale

The impetus for study 2 grew from two concerns about 1) the added shared variance because of a within-subjects design within study 1 that may have inflated measurement invariance and 2) the practical implications of using the questionnaire. Therefore, the second iteration involved only the items that were involved in the factor structures of *divisions of attention* and *relevancy* in study 1 (that is, not the prior MMI). These 9 items were to be investigated again to investigate for model fit, measurement invariance across contexts, and finally theoretical contribution with additional measures. However, this time, these items were not repeated. Therefore, study 2 shared hypotheses with study 1, such as:

H1: While media multitasking with difficult work-related tasks, people should report (a) less divisions of attention and (b) less irrelevancy among activities than while media multitasking with easy leisure activities.

H2: Those who engage in more a) attention divisions and b) less relevant

tasks during demanding work condition will experience more stress than those

in easy leisure, demanding leisure, or easy work.

H3: There will be a positive association between divisions of attention and

the degree of irrelevancy among activities.

Predictors of media multitasking. Theory, in addition to previous research,

reveal that *impulsivity* and *sensation-seeking* are strong predictors of media multitasking

behaviors. These two constructs, though potentially related, reflect distinct mechanisms

that drive media multitasking.

Impulsivity reflects a cognitive deficiency model of media multitasking

behaviors. Trait impulsivity is the propensity to swiftly act or react without thinking and

planning despite potential negative consequences (Barratt & Patton, 1983). The

theoretical model underlying the relationship between media multitasking and

impulsivity implies media multitasking is a risky behavior similar to smoking a

cigarette or unhealthy food consumption. Those who have examined *impulsivity* as a

predictor suggest that people engage in media multitasking, despite costs incurred via

increased cognitive load, due to deficiencies in executive functioning. Bottom-up

processes such as emotional reactions overwhelm and redirect attention, because of a

lack of ability for impulse inhibition. These individuals who have *insufficient* inhibitory

capacities in general should demonstrate greater behavioral impulsivity and engage

more frequently in media multitasking. Minear, Brasher, McCurdy, Lewis, and

Younggren (2013), Sanbonmatsu, Strayer, Medeiros-Ward and Watson (2013), Wilmer and Chein (2016), and Magen (2017) have all fond relationships between media multitasking frequency and increased impulsivity traits. In their review, Levine, Waite and Bowman (2012) argue that media multitasking across multiple contexts (e.g., academic performance, driving, waking and working) is related to trait impulsivity. Therefore, study 2 examined the relationship between media multitasking intensity (i.e., divisions of attention and relevance) and impulsivity across the aforementioned four contexts: easy leisure, easy work, demanding leisure, and demanding work. Due to a lack of consistent evidence relating the breadth-bias and relevancy of tasks within media multitasking study 2 considered this relationship exploratory, previous research had only examined the relationship between media multitsaking frequency and impulsivity.

Hypothesis 4: Greater trait impulsivity is positively associated with more divisions of attention across all four contexts: easy leisure, easy work, demanding leisure, and demanding work.

Research Question 1: Is there a relationship between trait impulsivity and task relevancy across contexts?

Sensation-Seeking is also related to risky behavior, but it models media multitasking behaviors as an outcome of preference for stimulation rather than a cognitive deficiency. Sensation seeking is defined as the need for novel experiences and the preference for taking risks (Zuckerman, 1994). Sensation seeking or the desire for

stimulation is predicted to motivate media multitasking such that those who are high sensation seekers engage in media multitasking more frequently in order to reach a more optimal state of arousal or situation. Previous research has found support for this relationship between sensation seeking and media multitasking frequency (Chang, 2016; Duff, Yoon, Wang, & Anghelcev, 2014; Jeon & Fishbein, 2007; Kononova, 2013; Lim & Shim, 2016). Unlike the research on impulsivity, previous research has not supported a relationship between sensation seeking and media multitasking across contexts. If sensation-seeking led to divisions of attention out of preference for stimulation, then the law for less work (Wang et al., 2015) should still apply. Those with high sensation-seeking should rather prefer greater demand when tasks were easy, regardless of context. However, with the dearth of literature relating task relevancy and sensation seeking, the examination of this relationship is exploratory.

Hypothesis 5: Greater sensation seeking predict more divisions of attention only in easy contexts: a) easy leisure and b) easy work, but does not predict divisions of attention in c) demanding leisure and d) demanding work.

Research Question 2: Is there a relationship between sensation seeking and task relevancy across contexts?

Study 2 Method

Participants. Participants ($n = 500$) were recruited through Amazon's Mechanical Turk (MTurk) via Turk Prime (beta) and compensated US $3.00 for completing the survey. Participants ranged from 18 to 73 years old and were on average

36 years old ($SD = 11$), with almost an equal number of females and males (50.7% male). The sample consisted of African Americans (7.9%), Asian Americans (8.3%), European Americans (74.6%), Latina/o Americans (5.7%), Multi-Ethnic Americans (2.6%), Native Americans (0.4%), and Other (0.6%). Participants reported their household annual income: almost half of them reported earning less than the country poverty line of $40,000 (41.6%), the majority of the sample (88.4%) reported earning less than $100,000 (46.8% earned between $40,000 - $99,999), a smaller portion of the sample (11.6%) reported earning more than $100,000.

Procedure. Once participants agreed to participating to the study and accepted the HIT on MTurk, they were each randomly assigned to complete the MMTIQ for one of four contexts (i.e., easy leisure, easy work, demanding leisure, and demanding work), followed by the measures of general stress, impulsivity, sensation seeking, and a few demographic questions such as occupation, gender, annual income, ethnicity, and age. The MMTIQ included the nine selected items from Study 1, which asked participants to self-report on divisions of attention and relevancy of tasks. In order to prompt participants' memory, for each context they were asked to report at least one specific activity that they remember engaging in during this period.

Measures.

Perceived general stress. The shortened 3-item scale by Cohen et al. (1983) was used for perceived general stress ($\alpha = .91$). These items were rated on a 5-point scale from (0) never, (1) almost never, (2) sometimes, (3) fairly often, (4) very often. Items

include statements such as "I feel overwhelmed with all the things I am doing in my life" and "What I am doing in life is demanding too much of me."

Trait impulsivity. The shortened 8-item Barratt's impulsivity test (S-BIT, Barratt, Stanford, Kent & Alan, 1997; Fields et al., 2015; Patton et al., 1995) was used to measure trait impulsivity ($\alpha = .82$). These items were rated on a 4-point scale from (1) rarely/never, (2) occasionally, (3) often, to (4) almost always/always. Items include statements such as, "I do things without thinking" and "I act on the spur of the moment."

Sensation seeking. The shortened 8-item sensation seeking scale (SSS; Zuckerman, 1994) was used to measure sensation seeking ($\alpha = .81$). These items were rated on a 5-point scale from (1) strongly disagree, (2) disagree, (3) neither disagree or agree, (4) agree, to (5) strongly agree. Items included statements such as "I like to do frightening things", and "I get restless when I spend too much time at home."

Media multitasking intensity: Conceptual model. The conceptual model included two factors (see Figure 3). The first factor, *divisions of attention*, included the following items: the total number of activities engaged, how often participants shifted their attention, how often participants open and minimized windows of activities on one device, and how often there were two to three media activities occurring simultaneously. These items were rated on a Likert-type scale from (1) never to (4) always. The second factor, *relevancy*, included items that measured the degree to which the activities they reported engaging in seemed relevant to one another, the degree to

which these activities shared a similar topic, the degree to which they were unconnected to goals, and the degree to which the activities were related to goals. These items were rated on a Likert- type scale from (1) never to (4) always with all items except the "unconnected to goals" items because they were reverse coded.

Study 2 Results

Multiple group confirmatory factor analysis: Measurement invariance.

After establishing adequate fit in the two most distinct sub-samples (i.e., easy leisure and demanding work), the MGCFA was conducted sequentially. Please refer to Table 7. First, the model fit was appropriate across easy leisure and demanding work for *configural invariance*. There was no change in CFI (ΔCFI $\leq$ .01). Thus, we continued to the next level of model constraints. Next, the model was estimated for *metric invariance*, which also demonstrated adequate model fit across indices and also had a ΔCFI less than .01. Third, the model was estimated after restricting equal factor loading and intercepts for *scalar invariance*, but this it did not demonstrate invariance. The ΔCFI was above cutoff value of .01. Thus, the model was considered to exhibit moderate measurement invariance. After scalar invariance was not established, when *invariance of factor variances and covariances* was tested, the model diminished in its goodness of fit at ΔCFI was below the cutoff value of .01. Therefore, the MMTIQ demonstrated moderate measurement invariance (See Table 7).

Table 7

Model Fit Statistics for Demanding Work and Easy Leisure Conditions Study 2

	χ^2	*df*	$\Delta\chi^2$	Δdf	RMSEA [90% CI]	SRMR	CFI	ΔCFI
Single Group Solutions								
Overall Sample	26.759	13	-		.065 [.029, .100]	.059	.985	-
Easy Leisure	17.353	13	-		.051 [.000, .107]	.043	.991	-
Demanding Work	21.309	13	-		.073 [.000, .127]	.070	.979	-
Measurement Invariance								
Configural	38.662	26	12.974		.063 [.000, .102]	.058	.985	0
Metric (equal loadings)	45.223	31	6.561	5	.061 [.007, .097]	.071	.984	-.001
Scalar (equal loadings and intercepts)	59.449	36	14.226	5	.072 [.037, .104]	.075	.973	-.011
Population Heterogeneity								
Equal Factor Variance	64.755	39	5.332	3	.074 [.039, .103]	.105	.970	-.003

Hypothesis 1. In support of the first hypothesis, *latent means* were significantly higher when media multitasking with easy leisure activities than when media multitasking with demanding work activities by a standardized unit of .55 in divisions of attention, and standardized unit of .54 for the irrelevancy between activities and goals. When engaging in easy leisure tasks, participants were more likely to engage in more divisions of attention and less relevant tasks than when engaging in demanding work tasks. Demanding work involved less divisions of attention, but more related media tasks than easy leisure.

Table 8

Study 2		General Stress		Impulsivity		Sensation Seeking	
Context	Factor	β	p	β	p	β	p
DL	Divisions of Attention	.28**	.003	-.19	.06	.19	.07
	Relevency	-.04	.66	.08	>.05	.04	>.05
EL	Divisions of Attention	.17	.07	-.11	.26	.32**	.001
	Relevency	-.21*	.02	-.13	.20	.20*	.04
DW	Divisions of Attention	.28**	.004	-.34***	<.001	.33**	.001
	Relevency	-.07	.46	.10	>.05	-.13	>.05
EW	Divisions of Attention	.03	.79	-.08	>.05	.12	>.05
	Relevency	.24*	.03	-.17	.08	.25*	.01

Hypothesis 2. It was predicted that more divisions of attention and less relevant

tasks would generate more general stress during demanding work than any other

context. This hypothesis was partially supported. Greater divisions of attention in

demanding work was statistically significantly related with greater general stress ($r =$

$.28, p < .01$), but there was not a significant association between the relevancy of tasks

in this context and stress. See Table 8 for all comparisons. Divisions of attention during

demanding leisure was also significantly positively related with stress ($r = .28, p < .01$),

but not relevancy.

Contrastingly, the relevancy of tasks only impacted general stress during easy

contexts: easy leisure and easy work. However, this trend of relationship reverses such

that during easy work more relevant tasks were significantly associated with more

general stress, and likewise less relevant tasks were associated with less general stress.

In contrast, relevancy was significantly related with less general stress in easy leisure

contexts (i.e., the more relevant media tasks were to one another and to goals, the less

stress one reported), though similarly, divisions of attention were not significantly

related to stress in this context. In summary, while greater divisions of attention within demanding work was significantly associated with more general stress, this relationship was not stronger than the relationship between divisions of attention and general stress within demand leisure. Also, within demanding work, relevancy was unrelated to general stress.

Hypothesis 3. It was predicted that divisions of attention and relevancy would be

positively related across contexts. This was partially supported; See Table 8. While divisions of attention and relevancy were significantly positively related in demand leisure ($r = .24$, $p < .05$) almost significantly related within easy work ($r = .19$, $p = .05$), there was no significant relationship between the two dimensions with demand work or easy leisure.

Hypothesis 4. It was predicted that greater trait impulsivity would predict more divisions of attention across all four contexts: easy leisure, easy work, demanding leisure, and demanding work. This was partially supported; only in the context of demanding work. Impulsivity did not produce significant relationships with divisions of attention in easy leisure, demanding leisure, or easy work. However, trait impulsivity was significantly and positivity related with divisions of attention in demanding work conditions ($r = .33$, $p = .001$). Trait impulsivity did not predict divisions of attention in any condition other than demand working suggesting that on average, media multitaskers are strategic (Ralph & Smilek, 2017).

Research question 1. In investigating the relationship between task relevancy and impulsivity, no statistically significant relationships were found.

Hypothesis 5. It was predicted that greater sensation seeking would predict more divisions of attention across all four contexts: easy leisure, easy work, demanding leisure, and demanding work. This was partially supported. Sensation seeking was only predictive of divisions of attention in the context of demanding work ($r = .33$, $p = .001$) and easy leisure ($r = .78$, $p < .05$). Divisions of attention were not related to sensation seeking under demanding leisure and easy work conditions (i.e., the two middle contexts).

Research question 2. In investigating the association between task relevancy and sensation seeking, only one statistically significant association was revealed. Sensation seeking was positively related with relevancy in easy work context ($r = .25$, $p = .01$).

Discussion

Study 2 replicated the two-factor structure within the media multitasking intensity instrument (MMTIQ): divisions of attention and task relevancy. However, in this second iteration the measure did not demonstrate measurement invariance across contexts (particularly between demanding work and easy leisure). Despite indications that the measure functioned differentially across contexts, hypotheses were supported. Media multitasking involved more divisions of attention and less relevant tasks in easy leisure conditions than in demanding work. However, divisions of attention and task

89

relevancy were only significantly related with one another in easy work and demanding leisure. This relationship was positive suggesting that this situation involved more frequent on-topic media multitasking.

In addition, study 2 provided theoretically valuable insights into the relationships between media multitasking intensity and general stress, impulsivity, and sensation seeking across contexts. Particularly, divisions of attention were correlated with stress under demanding contexts, while relevancy predicted stress under easy contexts. While attentional division was related to greater stress in both demand conditions, relevancy predicted greater stress in easy work conditions and less stress in easy leisure conditions. The splitting of attention amongst relevant, easy, but potentially consequential tasks (i.e., work) was significantly associated with greater stress. On the other hand, when attention is split with relevant, easy, and potentially less consequential tasks (i.e., leisure tasks) it was negatively associated with stress. These findings are likely driven by differences in the experience and creation of goals during work and leisure.

Moreover, impulsivity was only predictive of greater divisions of attention during demanding work, and sensation seeking was predictive of divisions of attention only during demanding work and easy leisure. Impulsivity had no relationship with relevancy while sensation-seeking predicted engagement in more relevant tasks during easy work. Impulsivity perhaps drive more frequent media multitasking behaviors within contexts in which in which attention divisions are correlated stress: high

90

difficulty work. Sensation seeking was correlated with more relevant tasks during easy work, which was correlated with more stress. This suggests that sensation seeking and impulsivity both lead to media multitasking and subsequent stress, but differ in the contexts in which they influence behavior. For instance, sensation seeking impacts media multitasking within easy conditions whereas impulsivity impacts media multitasking within difficult conditions. This finding is consistent with the conceptualizations of these construct; sensation seeking is preferential and therefore may be controllable as opposed to impulsivity implies poorer executive functioning or ability to control.

Study 3: Focus Groups, Cognitive Interviews, and Response Process Validity Goal

Because study 2 was unable to reproduce measurement invariance, study 3's purpose was to examine the formatting, structure, and content of the media multitasking intensity questionnaire (MMTIQ) for issues with response process validity: recall, comprehension, response options, and invalid responding. This involved re-examining the theory of media multitasking and re-evaluating threats to the validity on epistemological and methodological levels. Therefore, sixteen (16) 90-minute focus groups with five to eight people each were held in which participants completed the questionnaire, engaged in cognitive interviews, and participated in an unstructured focus group about media multitasking.

Study 3 Rationale

The lack of measurement invariance in study 2 suggested that there may be unobserved sources of heterogeneity or error in the current measure. Therefore, study 3 aimed to investigate sources of error by examining response processes. The measure had previously secured some evidence of validity within traditionally categorized types of validity such as construct validity and criterion validity (MacKenzie, Podsakoff, & Podsakoff, 2011), including measurement variance. However, the evidence suggests that there were test-retest effects. Measurement invariance was possible if participants were asked the questions multiple times as opposed to only once. Thus, the lack of measurement invariance led to the hypothesis that the existing error was a consequence of undetected threats to response process validity. Though response process validity is a fundamental source of validity evidence within the Education discipline (Kane, 2006; Messick, 1995), it also has important applications to the broader scope of social science and the interpretability of measures. Response process validity's importance has been corroborated by research that demonstrates that solely relying on statistical models as evidence of validity can creating 'false positives' even when survey items are meaningless (Maul, 2017). Therefore, study 3 examined this often-ignored aspect of validity.

Response process validity evidence remains rare and is still in the process of gaining clear conceptualizations, techniques, and standards (Ercikan & Pellegrino, 2017; Zumbo & Chan, 2014). Ercikan and Pellegrino (2017) define response processes as, "thoughts, strategies, approaches, and behaviors of respondents who read, interpret,

and formulate solutions/ answers," to the items and tasks in an instrument (p. 2). Evaluating the evidence of response process validity involves assessing the "fit between the construct and the detailed nature of response processes actually engaged in by its examinee" (AERA et al., 1999, p. 12; Kane, 2006). Stated differently, examining response process validity includes proposing and testing hypotheses about the causal mechanisms that generate observations. It also involves isolating the features of the items, formatting (i.e., language, space between items, instructions), and structure (i.e., order of items) within an instrument that disrupt these causal mechanisms necessary for accurate and reliable responses/observations. These causal mechanisms involve the cognitive, emotional, and social forces involved in interpreting and responding to an instrument that create unintended variation in responses/ observation (Gorin, 2007; Mislevy, 2003). Unintended sources of error may explain why specific populations or contexts exhibit challenges to valid measurement.

Cognitive interviews are valuable tools for assessing dysfunction in the process or the causal mechanisms that lead to a valid observation. Though cognitive interviews primarily provide information to improve response process validity, they may also reveal issues construct validity that were previously unobserved (Messick, 1995). Cognitive interviews are interviews in which participants either with (via verbal probing) or without (via a think-aloud) a moderator's guidance share their thoughts while completing the instrument, and in this case, the questionnaire (Beauchamp & McEwan, 2017; Oremus, Cosby, & Wolfson, 2005; Willis, 2005). The participants'

thoughts are used to assess the alignment between the predicted response process and the actual response process. They also assess the alignment between the information the participants choose to provide and the information that is purported to be observed. There are two common techniques for cognitive interviews: verbal probing and think-a-louds. Verbal probing involves develop a set of questions that are of interest to the researcher, which are asked after the completion of the instrument. Think-alouds are cognitive interviews in which participants are asked to state all their thoughts out loud as they complete the questionnaire. The two approaches each have different advantages and disadvantages. Willis (2005) found that verbal-probing is useful for focusing on specific issues, while think-alouds may be more open-ended. Beatty and Willis (2007) warn that verbal probing may lead to artificial problems, via artificially creating opinions (Converse, 1964); others have suggested this is because of spontaneous probes (Conrad & Blair, 2009) but the threat is perhaps overstated (Schuman & Presser, 1996). However, Preide and Farrall (2011) randomly sampled people to test a set of items via verbal probing and think-alouds. They found that verbal probing produces deeper thinking and longer responses, and could allow researchers multiple chances to understand the participants' response process. They argue that though this increases the chance for researcher effects than think-alouds, researcher effect is impossible to remove, and verbal probing provides a deeper understanding of the issues with items.

Verbal probing was the chosen method of cognitive interviewing for the current investigation (study 3) because the researcher wanted to test specific hypotheses about

the types of response process error occurring within examinees. In survey or self-report instruments, participants must interpret the question or item, recall and then assess their own experiences, interpret the response options, evaluate the most accurate response option, and then accurately indicate that for the researcher (Wolf, Ihm, Maul, & Taves, 2019). While it may not be possible for each participant to complete these steps in the process of formulating a response or observation identically, the goal of obtaining more valid observations and thus more valid measures requires some systematic behaviors or similarity. Gorin (2007) contends that education assessments require cognitive models of their items, including a model of the cognitive processes involved in interpreting and responding to items. Wolf et al. (2019) demonstrate that these models are also necessary for social science constructs beyond education. It is vital to model the response process and the difficulty of each of these response processes across contexts and populations of interest, and conceptualize whether and how they may interact with the construct-related information. Therefore, the following research question was investigated:

RQ1: Do participants experience issues with a) recalling information, b) comprehending

items, c) interpreting response options as aligning with their experiences, and d) providing accurate responses? If so, how frequently do each of these occur?

Because the construct of media multitasking intensity was also developed from theory, the other source of potential error was that the current conceptualization of the

construct was invalid because it lacked ecological validity. Stated differently, the items were hypothesized as the best predictors of the construct derived from language within the literature. While this increases the theoretical value of the construct, it could potentially contribute to reduced construct interpretability and thus poorer measurement quality (Maul et al., 2019). Essential observations of the construct may currently lack theoretical reasoning. Therefore, the cognitive interviewing process occurred within focus groups, and involved a more open-ended discussion about media multitasking experiences. The open-ended discussion was also considered valuable because the questionnaire and thus cognitive interview only observe media multitasking within a specific and short period of time (30 minute to 60 minutes). This may not be generalizable to all the respondents' experiences, but may in contrast reveal aspects of their media multitasking experiences that the questionnaire does not currently capture. Therefore, a second research question was posed:

> **RQ2a:** How do participants' experiences of media multitasking vary?

> **RQ2b**: Which instances of media multitasking are associated with stress as opposed to success?

Study 3 Method

Study design. Sixteen (16) focus groups, each involving between five and eight college students, were recruited from a pool of participants, from a Communication department in a university in the southwest. Each focus group was held for a duration of 90 to 120 minutes (1.5 – 2 hours). Due to the composition of the Communication

department pool, the majority of participants were female. The first hour of the focus group involved verbal probes about each of the items of the media multitasking intensity questionnaire (MMTIQ), while the remaining half-hour to hour was spent on an open-ended discussion of media multitasking. After the focus groups were conducted, audio recordings were transcribed by the Landmark Associates (https://www.thelai.com/). These transcripts were then divided into transcripts of the cognitive interview and transcripts of the open-ended discussion. A content analysis was employed to investigate the cognitive interview transcripts for instances of issues with recall, comprehension, response-options, and in-valid responses. A thematic analysis was conducted to investigate the open-ended discussion transcripts used to evaluate the essential concepts and themes.

Focus group procedures. Each focus group was moderated by two researchers. One researcher was a teaching assistant and graduate student within the Communication department and the other was an undergraduate student who was also a Communication major. The undergraduate researcher was involved in the focus groups in attempt to increase the comfort of and relatability to focus group participants. One researcher was the observer, note-taker and time-tracker, tasked with noting ideas, noting which participants had not spoken enough or did not have the chance to complete a thought, and keeping the group on time. The other researcher was the group facilitator, tasked with becoming immersed within the focus group while developing and maintaining a sense of rapport and flow amongst the participants of the group. The role of

moderator/facilitator and timer/note-taker were switched between researchers. At the beginning of the focus group, participants were told to begin by completing the questionnaire in silence, and that 10 minutes were reserved for questionnaire completion. Following completion, participants were told the remainder of the time would be spent talking. While participants completed the questionnaire, moderators took notes of puzzled faces or people who were pausing while completing the questionnaire. If participants became tempted to ask questions, they were asked to write down any questions they had about the questionnaire, which would be addressed later. Following completing the questionnaire, the participants in each focus group introduced themselves to each other, shared something unique about themselves, and finally developed a focus group name to which they connected. Once this connection was developed, the moderator began with the verbal probing. Following this verbal probing, participants were given a break to stand, use the restroom, and consume refreshments the researcher provided (i.e., water, bars, and chocolates). Then, the participants were told that the rest of the conversation would be more informal and would mainly be about exploring their experiences of media in daily life. The moderators at this point asked a few questions systematically, but otherwise allowed for spontaneous question formation. At the end of every two focus groups, the moderators reviewed notes and decided if they should revise items, response options, or formatting.

Verbal probing. Each focus group's discussion began with a set of ground rules, including that the moderator would aim to get as many opinions as possible and that the

goal was to hear experiences that were shared amongst everyone as well as those that were most unique. The first question for each of the focus groups was to ask if they felt they recalled their time-period well, and why. Following this question, each item of the questionnaire was reviewed with respect to how they interpreted it, and how they chose their response. At the end of the questionnaire, they were asked if they thought any of the items were redundant or still unclear. Participants were also asked if they had a problem with the response options (i.e., Never, a little, very often, etc.).

Open-ended discussion. Only a few questions were asked within each open-ended discussion portion of the focus groups. The open-ended discussion began with asking participants to share their perceptions of the purpose of the questionnaire. They were then asked to define media multitasking, share the locations where they most frequently media multitask, as well the situations in which media multitasking is most problematic as opposed to those that felt most beneficial. All other questions were spontaneously developed.

MMTIQ adjustments. This section will refer to the modifications of the items, response options, or formatting that the researchers chose to make across the focus groups.

Time. At the beginning of each focus group, participants were provided a consent form, a physical copy of the questionnaire, and they were assigned a specific time (e.g., 5:00 pm or 10:00 am) in the last 24 hours which they were to recall and assess while completing the questionnaire. The time at which the focus groups were

held was based on both the availability of the researchers and the availability of the space for holding focus groups. However, they were purposefully varied in order to reflect various portions of the time: early morning, mid-late morning, noon, afternoon and evening. In addition, the times on which they were asked to report was varied throughout the study, between 9 am – 10 pm. Although, after focus group two, it became clear that participants reported more recall issues if they were asked to report on a time more than a few hours ago. Thus, researchers limited the scope to three to four hours ago, and then further limited it to within the hour before the focus group began.

Duration of reflection. Focus groups before number eight were asked to reflect upon a 60-minute period, and focus group eight through 16 were asked to reflect upon a 30-minute period. Moderators noticed that participants divided their hour often into portions of 20 or 30 minutes. Participants then also reported challenges in averaging their experiences across these subsections of the hour. This posed a threat to response process validities, and therefore a 30-minute was investigated as a solution to this issue.

Item revisions. The researchers transcribed notes after each focus group, and items which demonstrated shared experiences of confusion or problems were adjusted.

Response options. Response options were altered based on participants' comments that they preferred six options or five options. Middle responses were removed originally because of research that suggested that the interpretation of this chose was unclear. However, some focus groups had mentioned preferring having a middle option, and was therefore compared.

Content analysis and coding procedures of verbal probing. After obtaining

the transcripts of the 16 focus groups, the moderators in addition to three other

undergraduate researchers were tasked with coding these focus groups for indicators of

issues with 1) recall, 2) comprehension, 3) response options, and 4) invalid responses.

These categories were conceptualized as mutually exclusive but any given comment

could demonstrate evidence for multiple of these issues. Coding was first completed

using markers on physical copies of each transcript through which lines were

highlighted, and then later entered into NVivo for further analysis.

Coders and formative reliability. Four undergraduate students majoring in

Communication were given up to five units course credit in the major in exchange for

coding this material. Of the total 16 focus groups, all four coders first coded one

randomly selected focus group, to assess formative reliability. They conducted iterative

content analyses until intercoder reliability exceeded the recommended standard

(Krippendorf's α = .80). The remaining 15 focus group transcripts were divided

amongst four coders and coded independently. Each coder coded at least three focus

groups units; three of these coders coded four focus groups. These were balanced for

length of focus group transcript.

Coding measures. The following variables were coded by reading the

transcripts, listening to the audio recordings, and highlighting the line(s) on the physical

transcripts. Lines were highlighted if respondents' communication indicated one of the

following response processing issues. The content analysis involved an event-based

sampling unit. Lines did not have a consistent length; therefore, the instance of response process issues was selected as the unit. Each instance could span multiple lines. All of the following measures were coded dichotomously as present or absent. In addition, a rule was applied to all construct measures: when other participants responses revealed agreement (e.g., "Uh uh", "Yes") with an issue another participant addressed, they were coded as another instance within the same category.

Recall. Recall was defined as the ability to easily and comfortably retrieve the information that is required for the questionnaire. Coders were told to code recall issues in three conditions: 1) if the participant explicitly expressed that they lacked confidence of their memory, 2) if a combination of the participant's vocalics or paralanguage and language reflected uncertainty about their memory, or 3) if participants reported finding inconsistencies in their responses due to a lack of accurate recall.

Comprehension. Comprehension was defined as the accuracy in understanding the questions and statements as intended as well as the clarity of the items to the participants. Comprehension issues occurred in two ways: 1) a lack of confidence in one's understanding of the item, or 2) a lack of accuracy in understanding the item as intended. In order to accurately code for comprehension issues, coders were provided a list of the intended definitions of the main concepts in each item (See Table 9).

Table 9

Table of Definitions for Codebook

Item	Definition
Shifting Attention	any perceived attentional shift between two tasks (i.e., could be both media or media and non-media)
Simultaneously or Overlapping	exposure or any level of attention provided to multiple media or media in addition to other tasks
Full attention	necessitating entire focus or attentional capacity
Simple	ease of task completion
Energy	general perceived effort exerted (both physical and mental)
Complement	activities are related to each other because they support or are useful for one another
Relevant/Similar Intentions	acitvities that share a smilar goal
Related/Share Topic	activites have share a common context but not purpose
Distraction	a task that is off topic or that is unrelated to goal
Planned	tasks in which people intend to engage
Habitual/Automatic/Mindless	tasks that are engaged effortlessly often without one's notice

Response option reasoning. Response option reasoning referred to the alignment between the logic participants provided for their choice of response option with the intended interpretation of the option. Importantly, response options were intended to refer to the frequency or proportion of time in which they engaged in the behavior. Response option reasoning issues occurred if: 1) people interpreted response options as the degree to which they agree with the statement, 2) lacked reasoning (e.g., "I don't know"), 3) expressed uncertainty about their reasoning (e.g., "I think…but maybe…"), or 4) provided other or unclear reasoning (i.e., miscellaneous reasoning).

Invalid responses. Invalid responses were defined as responses to the questionnaire that were consciously or intentionally incorrect. Coders were instructed to identify instances of invalid responses as those in which: 1) participants chose an answer because they were avoiding the extremes, 2) participants chose an answer

because they did not want to acknowledge their own behaviors, or 3) participants chose an answer for social desirability reasons.

Thematic Analyses

Five researchers read through the 16 transcripts of the open-ended discussions. After reading through these transcripts and listening to their audio, researchers began noting interesting concepts. After creating their own lists of concepts, researchers read through the transcripts again to identify quotes that examined the range of ways the concept appeared in participants' responses. Researchers then met, discussed, and compared quotes attributed within each concept. After agreeing upon a selected set of concepts, researchers began searching for relationships between concepts, and created themes that best reflect the value of the concepts and their relationships with one another. After generating themes independently, researchers evaluated their themes, combining those that were similar. Themes were compared with one another in this fashion until distinct themes appeared. Finally, researchers read through the transcripts once again in search of any unrepresented themes. Those themes and their quotes were discussed, leading to a finalized list of themes (provided below).

Results

Research question 1. In total, there were 690 errors in participants' response processing. The majority of errors occurred within comprehension ($n = 297$; 43%), followed by invalid responses ($n = 233$; 33.8%), response option reasoning ($n = 121$; 17.5%), and finally the smallest category of error was recall ($n = 39$; 5.6%). See Table

10 for a full review. The terms for which participants most commonly experienced *comprehension problems* were media activities, required full attention, simultaneous media use, and habitual media use. The items or concepts that had the most invalid responses included simultaneous media uses and shifted attention between media. *Response option reasoning issues* occurred most frequently for the question about mental and physical energy, the simplicity of primary activities, and whether media complemented one's primary goals. The concept that demonstrated the most *recall issues* was that of media activities. In total, the most problematic items were: required full attention and simultaneous media use. Comprehension issues accounted for the greatest number of invalid responses, followed by expressions of edge aversion (avoiding the response options placed at the extreme ends). In qualitatively reviewing the issues within each concept, distraction was also identified as a problematic item. Despite having a lower frequency of issues, the concept of distraction demonstrated a large disparity between desired and interpreted definition. These findings are the basis and justification for revisions to the MMTIQ items, which are examined in Study 4 below.

Table 10

Response Process Validity Issue Frequencies and Percentages Per Item

	Comprehension	Recall	Response Option Reason	Invalid Response	Frequency	Percentage
1. Energy	6	0	14	9	29	4%
2. Required Full Attention	43	3	10	19	75	11%
3. Simple	8	5	11	15	39	6%
4. Simultaneous and Overlap - Main Activity	11	0	0	6	17	2%
5. Simultaneous Media	35	6	8	37	86	12%
6. Simultaneous Other Activity	0	0	0	0	0	0%
7. Distraction	13	1	5	13	32	5%
8. Habitual	35	0	5	13	53	8%
9. Planned	4	2	7	4	17	2%
10. Main Intention	16	2	6	3	27	4%
11. Media Activities	45	9	5	16	75	11%
12. Complement	14	5	11	15	45	7%
13. Related	8	2	10	16	36	5%
14. Relevant	7	1	7	15	30	4%
15. Similar Intentions	4	0	2	8	14	2%
16. Similar Topic	8	1	8	14	31	4%
17. Shift Attention Main Activity	9	0	3	6	18	3%
18. Shift Attention Media	31	2	9	24	66	10%
Total	297	39	121	233	690	100%
Percentage	43%	6%	18%	34%		

Research question 2. The analysis of the open-ended discussion resulted in eight themes. These eight themes can be categorized into umbrella themes: contextual effects, necessity, emotional regulation, and maturity. These umbrella themes help synthesize the key insights gleaned from the focus groups. Contextual effects included the conditional effects of media multitasking based on its relational level communication and the contrast between the rarity and difficulty of media versus other tasks. Media multitasking emerges as a function of necessity both for relational enhancement and management as well as for impression management and productivity. Media multitasking is motivated by and often serves as a method of emotional regulation such that it can provide temporary relief or create a distraction from existing

discomfort. Finally, media multitasking is adaptive, it evolves over the life span. Quotes for each theme can be found in Table 11 below.

Theme 1: Contextual Effects.

1.1 Media multitasking's conditional effects on relationships: Communicative valence. In the focus groups, participants demonstrated that media multitasking in face-to-face interactions communicates information at the relational level. Though the content of the communication does not change when one is media multitasking, participants reported that they interpret it as a reflection of the closeness and connection between interactants. Because media multitasking is perceived as communicating relationship status, participants also reported expecting reciprocity from their face-to-face interactants. This reciprocity is entirely norm-based. In close relationships, norms could be established in which media multitasking is either entirely unacceptable or acceptable under conditions (e.g., information is shared when it creates an emotional response (i.e., laughter)). Regardless, participants reported disliking when others violated their expectations for reciprocity, and evaluated this behavior as disrespectful and rude. For these reasons, many reported engaging in social policing such that they make rules with others or directly inform people to stop engaging in their media. They also, report that this leads to self-monitoring because they experience guilt when they find that they are not adhering to their own expectations or are not reciprocating someone else's behavior. This self-monitoring was expressed as vital when others are not media multitasking (i.e., in class).

107

1.2 Perceived contrast effects on perceptions of media multitasking. Participants shared that media multitasking was more stressful or frustrating in situations when media was perceived to differ starkly in *effort, investment, urgency, or limitations* from the other tasks (including other media tasks) with which it co-occurs. Participants reported the most frustration when the potential time available for an activity, such as spending physical time with a friend, was limited and rare. They reflected that this was disappointing because they know the potential time that can be spent on media is unlimited and common. Similarly, participants reported experiencing stress when there was a greater perceived difference in the urgency of the other task then that of the media, which co-occurred with or interrupted it. For instance, completing an important work-related media task has a greater immediacy than scrolling through a social media newsfeed. Participants also reported that their stress and annoyances were driven by perceived differences in investments and effort-required. Engaging in media is perceived as so easy that it is more difficult to not engage in it. When others media multitask in face-to-face settings, this was especially frustrating in-part because they felt they were effortfully controlling their own desires for and habits of media use.

Theme 2: Necessity.

2.1 Self-reinforcing media multitasking: Constant connection. Media multitasking is perceived as a necessary way to enhance, maintain, and sometimes even supplant relationships. Similar to previous research (Rice et al., 2019), participants in the focus groups shared that being constantly connected to friends and family digitally

has created new norms and practices for these relationships. As digital communication becomes a habit or norm for a relationship, they create expectations which reinforce more media multitasking. Thus, the act of media multitasking, interrupting tasks or simultaneously engaging in media to share one's experiences with loved ones was frequently used to share information or communicate with others, but also created a positive feedback loop. In some instances, participants reported that this can lead to unclear relational statuses. Some participants discussed having relationships that were only maintained digitally, created a new category of friendships: "media friends,". These are people with whom a person might exchange memes or photos on a frequent basis, but with whom a person rarely or never connects with in face-to-face settings. Participants report the lack of indicators for closeness or motivations for these media-based relationships create uncertainty.

   2.2 Social media multitasking: The paradox of digital and physical impression management. Participants reported that a common situation in which they media multitasking is when they use social media to capture or share a current experience. The uses of social media while engaging in other tasks is often motivated by a desire to maintain social inclusion or social status. Thus, social media multitasking becomes a necessary way through which participants manage their digital impressions or self-representation. However, paradoxically, participants report that the absorption with social media increases the frequency of media multitasking in ways that can create poor in-person or physically made impressions. Some reported that this absorption has

created a new norm such that it impedes upon the ability to create physically new relationships. Others reported frustrations that people evaluate their social media or digital impressions as more important than those they make in-person.

2.3 Media multitasking as productivity and efficiency. Participants also reported that media multitasking has become a necessary skill at school and at work. Many argue that it is not possible to stay up-to-date or aware of their tasks within their role as a student or employee without frequently checking their emails or other websites. They argue that the incorporation of media into most of their goals at work or school has created both a need to media multitask. Stated differently, completing school or work-related tasks can involve media multitasking: answering phone calls in the middle of sending emails or taking notes while sitting in lecture.

Theme 3: Emotional Regulation.

3.1 Media multitasking as relief. Participants also stated that easier media tasks or simply the act of media multitasking was often actively used to relieve the mind when exhausted or when they realized their attention was no longer fully placed onto their primary tasks (e.g., listening to lecture or writing a paper). They described these media uses or this form of media multitasking as an intentional to reset their minds. Participants reported the value of media multitasking for mood repair when work or school-related tasks led to too much fatigue.

3.2 Media multitasking to avoid discomfort/seeking comfort. Another common experience of media multitasking involved the engagement of media in daily life in

110

order to avoid discomfort or to seek comfort. As opposed to mood relief, participants described this media multitasking as being motivated by the desire to feel safe or to manage feelings of discomfort. Specifically, they shared that they check their notifications to obtain a sense of "rushing endorphins", use media as background noise when silence feels uncomfortable, and most frequently use media in novel and uncomfortable physical settings (e.g., a party where they may not know anyone).

Theme 4: Maturity.

4.1 Media multitasking as a function of maturity and mastery. Participants reported that their media multitasking habits have fluctuated over time, they argued that it matured and evolved as they have gained experience and awareness of its effects. They also projected that these personal changes would one day be reflected in social changes such that society will mature alongside devices, one group compared media multitasking to cigarettes, contending that its negative effects and potential benefits will be fully understood and regulated in the future.

Table 11

Quotes exemplifying themes

Themes	Exemplary Quote
1. Contextual Effects	
1.1 Media multitasking's conditional effects on relationships: Communicative valence.	"When I'm at a meal or hanging out with someone, and it's like, if I'm making a conscious effort to make sure that I'm not [on my phone] to spend time with them, I would hope for the same from them." (Last Girls, 431-444)
1.2 Perceived contrast effects on perceptions of media multitasking.	"I think it depends on the instance, on the context cuz, if I'm not doing anything and I've done the reading and I've done homework, then I'm okay with it, and I'm fine with it, and I'm more than happy to see that people wanna talk to me. Yeah, if I'm already stressed and I know I haven't read or haven't done homework yet, then, yeah, I don't like it." (Fruity Pebbles, 402-407)
2. Necessity	
2.1 Self-reinforcing media multitasking: Constant connection.	"I know that I'm keeping up with all the story in the group chat so I know what's going on, but at the same time, I know that I'm good at keeping up or replying pretty fast. Now it's becoming a stress for me like, oh, I have to do this. I always have to do this." (Freebirds, 337-341)
2.2 Social media multitasking: The paradox of digital and physical impression management.	"That's so true, yeah, cuz everybody looks at who looks at their story. Even though my friend had a private one for my birthday, we were still—there were certain things like nobody could see, but then there were other things I'm like, "Oh, I'm in Vegas. Let me show everybody."" (Fruity Pebbles, 703-707)
2.3 Media multitasking as productivity and efficiency.	"I feel like also, what I was saying, it's a necessity. In order to keep up with the fast pace of everything around us, we have to do it. It's like, especially for business or for school, it's like you have to multitask in order to do multiple things for classes or put attention to multiple things at once." (Gladiators, 132-135)
3. Emotional Regulation	
3.1 Media multitasking as relief	"I feel like sometimes it could like—it's like a temporary stress reliever. If I have a paper to do, and I kinda just go on my phone and then stay there for an hour, I don't think about my paper; but, then, later, it's like, "Oh! Yeah!" [They all laugh]" (Girl Squad, 241-244)
3.2 Media multitasking to avoid discomfort/seeking comfort.	"Yeah, kind of like, throughout this whole study, I've just been thinking about how much I do it now. It made me more conscientious of—like you said—just the different instances in my day. I kinda have come to not prefer silence anymore. Kind of like what you said, I need music in the car. I need TV while I eat. I don't know. Just thinking about how that affects my life, if it does, and maybe experimenting with taking it away." (Stress Team, 667-785)
4. Maturity	
4.1 Media multitasking as a function of maturity and mastery.	"I feel like I, maybe when I was younger, I was stressed doing this. I've learned how to—you learn what to balance and balance your distractions and your media uses at the same time. I even see the younger generation. My little brother, he's fine media multitasking. He can do that. Maybe us and older people who haven't grown or a little bit later growing up with that, smartphones and constantly using it, at first it was a little bit overwhelming. Now I don't ever feel stressed by that. " (Gladiators, 239-249)

Study 4: Pilot of Revised MMTIQ Items

Goal

After identifying problematic concepts in study 3 and identifying themes about the experiences of media multitasking, study 4 investigated whether revisions to the formatting and verbiage of the items within the questionnaire would improve comprehension. Based on the findings from study 3, study 4 evaluated the media multitasking intensity questionnaire within a theoretically significant population: parents and adolescents. The necessity for Likert-scale or polytomous response options, and the measurement model of the media multitasking intensity questionnaire were also examined.

Rationale

Study 3 revealed issues in response processes that could potentially explain the problems with study 2. Conceptually, study 3 in conjunction with study 1 and 2 suggested that the construct of media multitasking was moderately secure, there was evidence that type of multitasking, intentionality, difficulty of tasks, and relevance of tasks predicted a diverse range of experiences of media multitasking in the analysis of open-ended discussions and in statistical analyses. However, the security of validity claims for the questionnaire as a measure for media multitasking intensity was challenged by the heterogeneity (revealed by study 3) in the interpretation of items and response options as well as the propensity for invalid responding. It seemed that issues with recall mainly occurred within focus groups in which participants reflected upon a

60-minute period which had not most recently occurred. However, the issue of recall was rare in comparison to those of comprehension, invalid responding, and response option reasoning. Improving comprehension of focal concepts required revising items, but also finding more influential way of the definition of media, which challenged interpretations of multiple items that are essential to the construct. Therefore, study 4 involved revising the most problematic items, and investigating comprehension via meta-questions as explained in Wolf et al. (2019), and then coding responses in aims to reach at least 70% accuracy. This cut-off value was determined by allowing for some error in both participants' abilities to express their comprehension, and researchers' abilities to ascertain comprehension.

RQ1: Will revised items of a) simultaneous media use, b) required full attention, c) habitual media use, and d) distraction be comprehended accurately in at least 70% of the sample?

In addition to comprehension, one third (33%) of response process issues arose from invalid responses. This was troubling. Issues in invalid responses suggested the necessity for revising the format of response options in order to deter edge-aversion (i.e., avoiding extreme responses). This new format should also aim to solve the main issue with response processing, by interpreting responses as the extent to which they agree with statements. While this edge adverse behavior was sometimes described as construct irrelevant, edge aversion also revealed that media multitasking was a culturally sensitive concept, especially for young adults or older adolescents. Young

114

adults were aware that society and adults believe media multitasking is destructive and that they should not engage in it, but do so more frequently than they realized. This motivated some invalid responding partially because it inspired a sense that "never or rarely" could not be accurate and because "always or almost all the time" would be socially disparaged.

Yet, they also shared that their invalid responses were motivated by the intrapersonal communication consequences of choosing a response option. While the questionnaire served to communicate their observations to the researcher, the questionnaire also inspired new insights that were sometimes uncomfortable for the participants to reveal to themselves. For instance, one participant noted that he/she/they had never noticed that they frequently begin to media multitasking when on the phone with their romantic partner and began to explore why this was the case. Many shared that they had not reflected on these aspects of media uses before the questionnaire. Thus, providing a certain response on the questionnaire increased self-awareness and required acknowledgement of their own behaviors of which they may previously have been blissfully unaware. Their own judgment of their behavior prevented them from providing accurate responses. The instances of social desirability, edge aversion, and sensitivity to self-awareness, suggested that participants may benefit from a small introduction or disclaimer that the researcher will not judge their behaviors because they do not necessarily think media multitasking is damaging and that media multitasking is

something everyone does more frequently than they recognize. Thus, study 4

investigated the following research questions:

RQ2: Will revising the response option process such questions are posed

first dichotomously and then polytomously increase the frequency at which

people select response options at the edge?

RQ3: Will providing some myth-busters about media, a disclaimer about the

researchers' evaluation of media multitasking, and changes in the formatting of

the response options lead to low levels of invalid responding?

Theoretical Advances

In alignment with the iterative process of validity described in Chapter 2, these

methodological changes in the MMTIQ prompted changes at the theoretical level.

Given the response process validity issues and themes identified in Study 3, adjustments

were made not only to the design and items of the questionnaire but the

conceptualizations dimensions of *media multitasking intensity*. After the focus groups in

which the contrast effect emerged as a theme, task difficulty was reintroduced as a

dimension of the construct rather than a dimension of the context. It appeared that

experiences of media multitasking transformed through the various perceptions of

difficulty of the general tasks and intentions that were experienced by a person as well

as the role of media in exacerbating or competing with these tasks. Similarly, intentions

were re-conceptualized as the conscious experience of choosing or planning to prepare

in tasks. In Study 3, it became apparent that intentions were an experience that also

116

varied across context and people. Some people reported they did not know what their intentions were, others shared that it was difficult to think of an "intention" for a time in which they were relaxing. Thus, in *Figure 5,* is the resulting two dimensional model. After reassessing the theory of media multitasking it appeared that the demand creating or cognitive load created by the multiple tasks and the demand created by the difficulty of tasks would interact with one another, and become nested within a higher-order dimension of cognitive demand. On the other hand, intentions and task relevancy appeared to guide information synthesis by linking stimuli and creating sensory integration. In Study 3, many participants reported that media tasks like listening to music, did not share a goal or topic with their daily tasks (i.e., exercising, eating dinner, talking to a friend, or completing homework), but was relevant in that it helped them enjoy and engage more successfully in their others tasks. Thus, study 4 examined conceptual changes as well as changes to the MMTIQ's quality and validity.
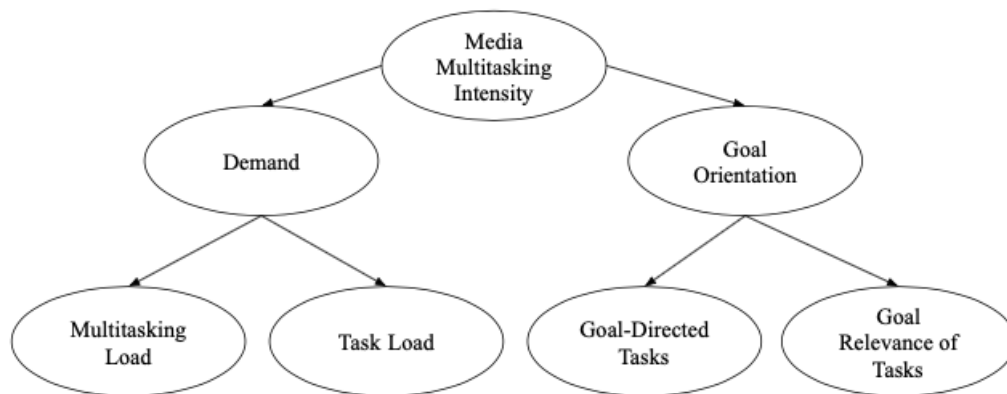
Figure 5. The re-conceptualized two-dimensional model of media multitasking intensity.

The Role of Theory in Item Revisions

Media multitasking had been conceptualized as primary and secondary activities and primary goals, which were key attributes derived from limited capacity models (Wang et al., 2015; Yeykelis et al., 2014). However, it became apparent that though these psychological theories investigate the cognitive causal mechanisms of multitasking such as goal-oriented behavior. The theories of cognition may not reflect conscious experiences of cognition. Participants in study 3 sometimes reported that they did not even have goals or priorities. Thus, though all effortful behavior likely had an intention, participants may vary in their consciousness or awareness of their intentions, or may have less intentional behavior in general. Those who were not aware of their own intentions would not be able to accurately respond to the questionnaire because there would be no response option for them. Moderators noted that this was especially true in easy or leisure conditions, which aligned with study 2's measurement invariance issues, in which easy leisure and demanding work no longer demonstrated similar intercepts. In other words, the items were experienced differently for those in an easy leisure and demanding work contexts.

Moreover, participants expressed that they did not always have "goals." Sometimes intentions were to make progress or to begin to engage in task as opposed to accomplish a goal or task. Thus, the language in the questionnaire was transformed. As

118

opposed to the modifications made based on comprehension, this change reflected reconceptualizing the epistemology of the construct in relation to ontology. Though perhaps cognitive neuroscientists and psychologists could find mechanisms that reflect goal creation and conflicts (Salvucci et al., 2009), these were not always conscious experiences for participants. However, rather than regard this as simply an issue of human limitations and error, this suggested that the variation in consciousness about goal creation and goal conflicts was theoretically significant. Media multitasking intensity could vary based on how consciously participants experienced, realized, and thus could evaluate their own intentions. Moreover, in the open-ended discussions of media multitasking intensity participants mentioned that their media multitasking intensity had evolved and fluctuated as they gained insights into their own experiences. This theoretically relevant variation inspired investigations into a new construct: *self-regulation*.

Scholars have investigated the role of mindfulness and self-regulation on media multitasking behaviors. Schilhab (2017) recently argued that the increasingly media-saturated environment requires greater self-regulation, attentional regulation, and self-control. The use of awareness to adjust behavior over time has similarly been posited to protect one from excessive habitual media multitasking behavior. In support of this prediction, Zhang (2015) found that self-regulation predicted the amount of media multitasking participants engaged in on their laptops within classrooms. Thus, relating self-regulation to the likelihood of focusing in class. Schutten, Stokes, and Arnell

119

(2017) revealed that heavy media multitaskers (as measured by the MMI) report less self-control but more impulsivity (as discussed earlier in this chapter), while demonstrating a reduced capacity to delay gratification. Simultaneously, Levy, Wobbrock, Kaszniak, and Ostergren (2012) found that individuals who engaged in an 8-week mindfulness meditation program were less likely to engage in task switching and were able to sustain their attention on one task for longer However, this effect was stronger for those who were heavy media multitaskers than those who were light media multitaskers. Fan, Gong, Wang, and Wang's (2017) also found that meta-cognition as defined as thoughts about thoughts or thoughts about feelings (Flavell, 1979; Mayer & Gaschke, 1988) moderated the volume of media multitasking behaviors during learning tasks with greater difficulty. These studies in conjunction with findings from study 3 and study 2, suggested that media multitasking intensity must fundamentally involve awareness of one's intentions. This intentionality could impact the degree to which one divides their attention, engages in difficult tasks and engages in relevant tasks. As opposed to impulsivity, which mostly impacts motor control and behavioral inhibition, self-regulation involves the ability to observe behavior, evaluate behavior, and strategize when any behaviors require change.

Due to these findings and theoretical developments, study 4 evaluated media multitasking behaviors amongst a key population that most varies in self-regulation capacity: adolescents. Adolescents' cognition experiences a key transformation in meta-cognition, self-regulation, and introspection (Steinberg et al., 2017). Thus, validating

120

the MMTIQ and the developing theory about the causes of variation in media

multitasking and its consequences required evaluating its ability to demonstrate validity

in this population. However, adolescents are traditionally difficult to recruit in studies.

Their participation requires not only their own consent but consent from their parents

and potentially educators depending on how they recruited. Therefore, study 4 found a

feasible way to recruit adolescents to complete the questionnaire. In addition, it was

necessary to ascertain whether the revisions to the MMTIQ would also benefit older

adults. This revealed an important component of the ontology of media multitasking

intensity that required explicating and evaluating: its universality. Therefore, study 4

ultimately used Turk Prime, a private research company which has built extensions for

Amazon's Mechanical Turk, the crowd sourcing platform. The study obtained a sample

of parents and their adolescents through Turk Prime to investigate the fit of the

measurement model, in addition to the comprehension of items, and invalid responses.

RQ4: Will the revised set of items demonstrate good fit to the

conceptualized measurement model?

Study 4 Method

Participants. Parents of adolescents (*n=150)* were recruited through Turk

Prime and compensated US $3.00 for completing the survey. Parents were then asked to

recruit their own adolescents to participate in the study as well. Parents provided their

own consent and then consent for the adolescents. Participants ranged from 13 to 62

years old and were on average 31.43 years old (*SD* = 13.71). Parents' age ranged from

121

26 to 62 years, and adolescents ranged from 13 to 18 years old. The total sample included only slightly more females than males (53.8% female), however, the majority of parents were females or mothers ($n = 107$; 71.3%) and the majority of adolescents were males ($n = 71$; 73.2%). The sample consisted of African Americans (13.8%), Asian Americans (2.8%), European Americans (75.7%), Latina/o Americans (4.5%), Multi-Ethnic Americans (1.6%), Native Americans (1.2%), and Other (0.4%). Participants classified themselves as working class (18.1%), lower middle class (24.2%), middle class (46.3%), upper middle class (10.7%), and upper class (0.7%). The median income in sample was 60-79,000 US dollars.

  Procedure. Parents and adolescents took the questionnaire one week apart. Once parents agreed to participate to the study and accepted the HIT on Turk Prime, they each began a block of narrative questions (e.g., those meant to bolster participants' memory). They began by reporting the time 30 minutes ago, as well as whether they had an intention, or multiple intentions, at the beginning of that time. For those who reported not having an intention, they were asked to share how they understood the word "intention." The rest of the questionnaire was adjusted based on their answers to this question: those who did not have intention were only asked to report on a specific activity that was closest to having a priority over others. Those who had only one intention named an activity that was most aligned with this intention (i.e., primary activity). Finally, those with multiple intentions were asked to choose one primary intention and then a primary activity for that intention. The remainder of the

questionnaire was designed to populate the time participants originally reported to appear as a remainder for the time period they were recalling. The purpose was to continue to continue to support accuracy in recall and reduce the cognitive load of answering the questionnaire. All participants completed the MMTIQ, followed by questions about whether they were uncomfortable in the process of completing the questionnaire and if that ever lead them to change their answers. This was motivated by the experiences of participants in Study 3 who shared that the questionnaire often increased their awareness about their behaviors in ways that were previously unknown to them and once revealed face-threatening.

Measures.

Media multitasking intensity. A revised version of the Media Multitasking Intensity Questionnaire (MMTIQ; Guttman $\lambda 6 = .66$) was employed, in which adjustments were made to the items, formatting, and response options of the MMTIQ, as identified through Study 3.

The first factor, *attentional demand*, multitasking demand and task demand. The sub-dimension of *multitasking demand* included items about the total number of media activities in which they engaged (during the half-hour period), as well as if participants shifted their attention between media, shifted their attention between media and their main intention, used multiple media, used media while engaging in other tasks, and used media for background noise. If participants reported that any of these statements were true, they were then asked to report on the proportion of that half hour for which

123

this statement as true on Likert-type scale of (1) a little of the time, (2) some of the time) and (3) most of the time. The sub-dimension of *task demand* included questions about whether tasks felt they deserved full attention, they gave their full attention to tasks, the task required a lot of energy, and their media tasks were difficult. If participants said yes to any of the above questions, they were then asked to report on the proportion of the half hour for which this statement was true on Likert-type scale of (1) a little of the time, (2) some of the time) and (3) most of the time.

The second factor, *goal-synthesis*, which is the degree to which information from the tasks in which participants engaged could easily be synthesized, also included two sub-dimensions. The first sub-dimension was *goal-directed behavior*, which included items that first measured if and then how frequently participants engaged in media tasks that were relevant to their primary activity, that complemented primary intentions or activities, or that were engaged in without even thinking. If participants reported any of these items to be true of their half-hour, they then reported the proportion of time they spent with these media activities on a Likert- type scale of (1) a little of the time, (2) some of the time) and (3) most of the time. The second sub-dimension was *task-relevance*, which included items such as whether their media tasks seemed related to one another, and whether their media tasks were unwanted distractions. Participants were then asked to rate the frequency of these behaviors on a Likert- type scale of (1) a little of the time, (2) some of the time) and (3) most of the time only if they reported they were true.

Item comprehension. In order to investigate item-comprehension, participants were asked meta-survey questions only for those items that were most problematic in Study 3 and which had therefore been revised. Item comprehension was defined as the accuracy in understanding the questions and statements as intended as well as the clarity of the items to the participants. Item comprehension was appraised for: multiple media use, gave full attention, needed full attention, media use without even thinking, and unwanted distraction (See Table 12 for definitions). Two coders coded 60 randomly selected answers (30 of which were from parents and 30 of which were from adolescents) to each item (240 responses) to evaluate formative reliability. Similar to study 3, the codebook and definitions of each concept were revised and clarified, and these 60 answers were recoded three times before reaching the recommended standard (Krippendorf's α = .80). After reliability was established, one coder completed the remaining coding. The items were coded ordinally, where (0) indicated there was no alignment, the answer was completely incorrect or provided nor relevant information, (.5) the answer is partially aligned or there was insufficient information to be confidence of the alignment, and (1) the answer provided was completely aligned.

Table 12

Table of Definitions for Codebook for Study 4

Item	Definition
Multiple media at the same time	exposure or any level of attention provided to multiple media or media in addition to other tasks
Full attention	providing entire focus or attentional capacity
Needed Full Attention	requiring or needing entire focus or attentional capacity
Without even thinking	tasks that are engaged effortlessly often without one's notice
Unwanted Distraction	activities that were undesirable, off-topic, and unplanned

Results

Research question 1. The study explored whether the revised items within the questionnaire would demonstrate acceptable levels of comprehension (i.e., 70% precision). This involved analyzing the items' comprehension scores (See Table 13 for frequencies). In assessing the five items -- multiple media use, gave full attention, needed full attention, used media without even thinking, and unwanted distractions quantitatively -- every item except for those regarding full attention met the 70% accuracy requirement. However, in assessing the five items qualitatively, multiple media use still demonstrated some uncertainty about the distinction between media and devices (i.e., laptops, smartphones, and tablets).

Table 13

Item	No accuracy (0)	Partial accuracy (.5)	Complete accuracy (1)
Multiple media use	43	23	215
Full attention	67	75	139
Need for full attention	75	51	153
Without even thinking	27	43	210
Unwanted Distraction	24	58	199

Item	No accuracy (0)	Partial accuracy (.5)	Complete accuracy (1)
Multiple media use	15.3%	8.2%	76.5%
Full attention	23.8%	26.7%	49.5%
Need for full attention	26.9%	18.3%	54.8%
Without even thinking	9.6%	15.4%	75.0%
Unwanted Distraction	8.5%	20.6%	70.8%

Research question 2. This study also explored the degree to which changes in formatting the response options such that they are first presented dichotomously and then polytomously were investigated by examining the frequencies of never and most of

126

the time, as well as in fitting the measurement model both with dichotomous (0) no and (1) yes or (0) absent and (1) present indicators and polytomous indicators from (0) never to (3) most of the time. Frequencies of never and always were at either 50% likelihood or more frequent than the middle options of (little of the time or some of the time). The combination of the measurement models, examined below in research question four and the observed frequencies, provided evidence for less invalid responses via edge aversion.

Research question 3. In order to explore whether the added disclaimers reduced invalid responding due to discomfort, the frequency of responding yes to the item that asked about discomfort and response changes due to discomfort were analyzed via a Chi-Square analysis. These frequencies were first explored by conducting a chi-square analysis within the entire sample, and then examined for parent and teens independently. The proportion of people who admitted that they felt uncomfortable was extremely small; only one person admitted to being uncomfortable and only two people admitted to changing their answers due this discomfort. While this could suggest that there was less invalid responding, this could have also been a consequence of the compensation mechanisms on Turk Prime in which participants who provide invalid responses are not compensated for their time.

Additional Analyses. After examining frequencies for the discomfort measures, the correlations between attitudes toward media and multitasking, discomfort, and response changes were assessed to investigate if the explanation for these invalid

127

responses was that they viewed media or multitasking as more harmful. This correlation was first analyzed within the whole sample, and then within parents and adolescents independently. Unfortunately, the frequency of admitting to discomfort and answer changes was so low that it was not possible to examine this correlation. Therefore, the researcher examined the relationship between believing multitasking is harmful and believing that media was harmful. The two demonstrated a weak correlation ($r = .20$, $p < .01$), and were both also correlated with age (harmful media $r = -.22$, $p < .001$; harmful multitasking $r = -.14$, $p < .05$), as well as whether one was a teen or not (harmful media $r = .26$, $p < .001$; harmful multitasking $r = .13$, $p < .05$). Thus, teens were more likely to interpret media and multitasking as more harmful than were parents or adults.

Research question 4. In answering research question 4, four measurement models were fit and compared to each other for demonstrating adequate fit. As opposed to study 1 and 2, which used factor analysis to analyze the measurement model, study four investigated a probabilistic relationship between either dichotomous or categorical response options and a latent variable of media multitasking that is continuous (See Chapter 2 and 5 for more detail) Thus, it was appropriate to employ item response theory models.

Comparing dichotomous and polytomous models. In comparing dichotomous and polytomous models, there is a comparison between whether the observations are best observed ordinally or nominally. In other words, the current study investigated

whether observed ordinal variation in frequency as opposed to presence or absence of

attributes intensity was more valuable. The partial credit model, which is within the

Item Response Theory family, was used to evaluate the value of each response option.

The value of adding additional response options (i.e., polytomous instead of

dichotomous) can be evaluated by examining if each response option has a value on the

latent attributes or theta at which their probability of being selected is greater than that

of the other options, these probabilities are visualized within response option curves.

These options included, "rarely or never" (P1), "little of the time" (P2), "some of the

time" (P3), and "most of the time or always" (P4). The response options curves revealed

that the options "little of the time" (P2) and "some of the time" (P3) frequently did not

have unique points on the latent attribute at which they were most probable. This is

demonstrated Figures 6 and 7 which show the category response curves for the

polytomous models of parents and teens. In these figures, lines P1 and P4 often cross

one another, which depicts that if participants did not report "never or rarely" they most

likely or had the highest probability of choosing "most of the time or always". In other

words, the other categories (P2 and P3) did not provide more information about the

latent attribute of media multitasking intensity. Rather, these response curves

demonstrate that the options none of the time, and most of the time provided the most

information about variation within media multitasking intensity in the data.

Item trace lines
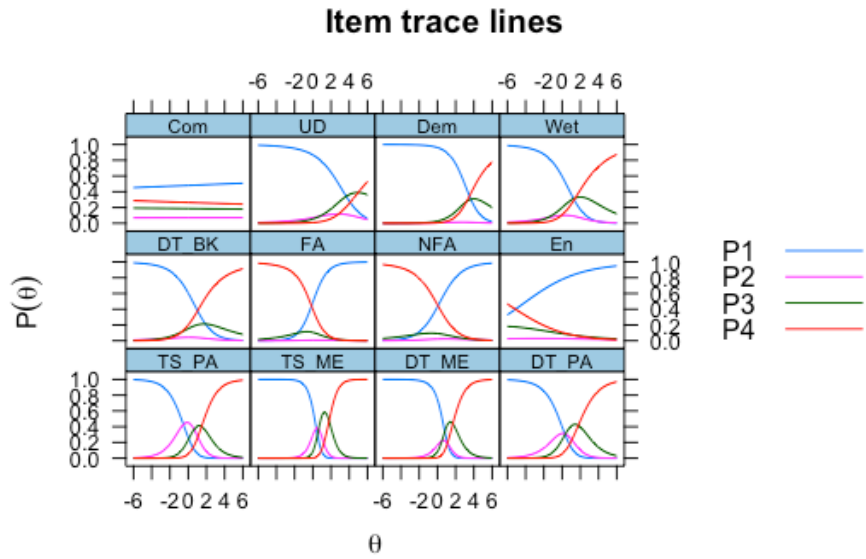
Figure 6. Parent response category curves for each item and response option.

Item trace lines
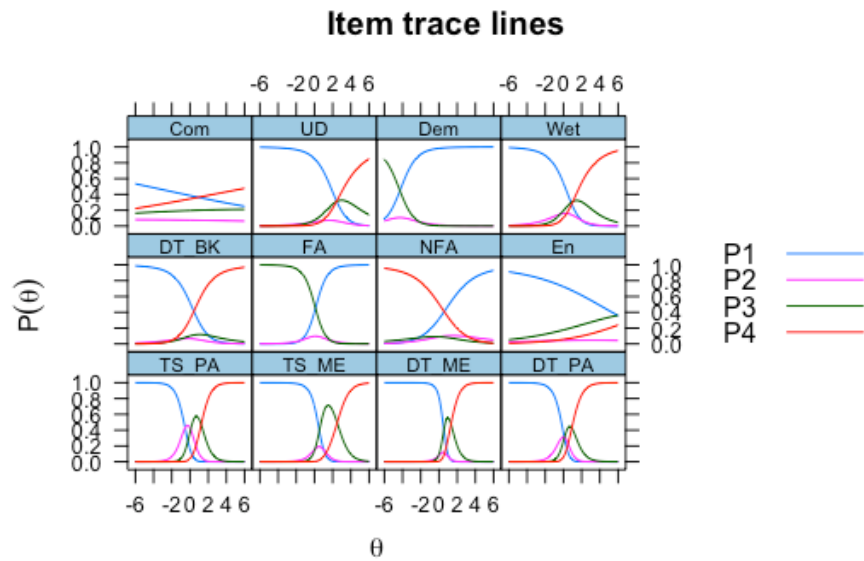
Figure 7. Teen response category curves for each item and response option.

Comparing unidimensional and multidimensional model specification. The

statistics reflecting goodness of fit between the data and the conceptualized two-factor

(parents and adolescents) multidimensional model was compared with that of a unidimensional model (combined). Model fit indices for parents and adolescents independently as well as collectively appear in Table 14. Comparison of these fit indices suggested that the unidimensional model fit best, as evidenced by its noticeably lower deviance (G2), log likelihood, AIC, and BIC (in absolute value) compared to the fit indices of the multidimensional model.

Table 14

Model Fit Indices for Teens

	Log Likelihood	AIC	BIC	Deviance	Parameters
Undimensional Rasch	-494.05	1008.00	1036.00	988.11	10
Two Dimensional Model	-492.52	1013.04	1051.83	135.63	14
Unidimensional Partial Credit Scale Model	-1138.20	2360.40	2476.76	1186.93	42
Two Dimensional Model	-1137.88	2365.77	2490.45	1186.30	45

Model Fit Indices for Parents

	Log Likelihood	AIC	BIC	Deviance	Parameters
Undimensional Rasch	-1192.92	2406.00	2442.00	2385.83	10
Two Dimensional Model	-1186.57	2401.14	2451.57	239.74	14
Unidimensional Partial Credit Scale Model	-1528.20	3146.41	3282.78	1551.41	45
Two Dimensional Model	-1526.27	3148.54	3294.01	1547.55	48

However, a portion of the reason why the unidimensional model may have demonstrated better model fit is because the *item-fit* statistics demonstrated problems. The MMTIQ 12 items demonstrated on average (Teen outfit $M = .99$, $SD = .22$; Teen infit $M = 1.00$, $SD = .12$; Parent outfit $M = .96$, $SD = .23$; Parent infit $M = 1.00$, $SD =$

.08) appropriate infit and outfit statistics with mean squares that fall between 0.79 and 1.21 calculated by the recommendations of Wu and Adams (2013).

Item misfit. Within the IRT framework described in detail in Chapter 6, items are fit to the model by identifying the extent to which the match the expected pattern specified within the model. This pattern refers to the probability of choosing to either "yes" instead of "no" or of choosing a higher or lower category on a Likert-scale. These patterns are connected to variation within the latent trait, which in this investigation is media multitasking intensity. Misfit of an item suggests that it is not functioning according to the model, and is contributing to poorer model fit statistics.

Two items demonstrated the most problematic fit-statistics: the need for full attention and the given full attention. The items including the vocabulary "full attention" demonstrated mean squares above these cutoff scores, or underfit of the model, suggesting there was more variation in responses than expected. Because these items were also demonstrated to have poor response process validity – they were not comprehended by 70% or more of participants (see Table 13), they were removed from the model. Thus, the MMTIQ with 14-items was re-evaluated.

In the 14-item model, it appeared that the items of complemented tasks, difficult media tasks, energy required by a task, relevance of media task, and relationship between media tasks were similarly not performing as expected. The model fit statistics demonstrated that there was more noise in these answers than predicted. Ultimately a 9-item model was estimated which included total media (1), media use with non-media

132

(2), multiple media use (3), media as background noise (4), switching between media (5), switching between media and non-media (6), whether or not there was an intention (7), whether media complemented other tasks (8), and whether media were used without even thinking (9). Note, all the items related to the dimensions of task difficulty, and many of those related to goal-synthesis, were removed, which may explain the fit of the unidimensional construct.

Person-to-item fit. Once the nine-item model dichotomously demonstrated better fit for parents and teens, the person-to-item fit was examined. Within IRT models as described in Chapter 6, people and items are estimated to be on the same scale, which are referred to as logits (see Figure 6 and 7). These logits are units of the latent variable, which in this investigation is media multitasking intensity. The ability for a measure to provide ample information about people across the scale of potential values on a latent trait depends on the ability for items to capture variance across the scale of a latent trait. Thus, Wright Maps (Markus & Borsboom, 2013), which compare variance in items and people on the continuum of the latent variable are used to assess the quality of the measure. Wright Maps for the unidimensional models of parents and adolescents are available in Figures 8 and 9. On Wright Maps as depicted in these figures, participants' distribution of values of media multitasking intensity are visualized by aggregate participant scores on the left, and the continuum of media multitasking intensity in units or logits placed on the right. Thus, using these Wright Maps, one can easily visualize the distribution of parents and teens across the continuum of media multitasking

intensity. Similarly, one can visualize the variation of media multitasking intensity

captured by the items or observations. On the bottom, items are labeled and indicated by

dots on the same logits scale on the right denoting the value of the attribute necessary

for the probably of a specific quality observed by the item to be present (i.e., endorsed

as present by saying yes) is at 50%. Each item also obtains a logit scale score.

Comparing parents and teens. While the items' values of media multitasking

intensity spanned from 4 (very high intensity) to -5 (very low intensity), the range of

adolescent values on the media multitasking intensity spanned from 4 to -5. Meanwhile

the range of values for parents spanned from 2 to -4. Therefore, the items demonstrated

a better

Figure 8. Parent MMTIQ Wright Map.

Wright Map

Wright Map
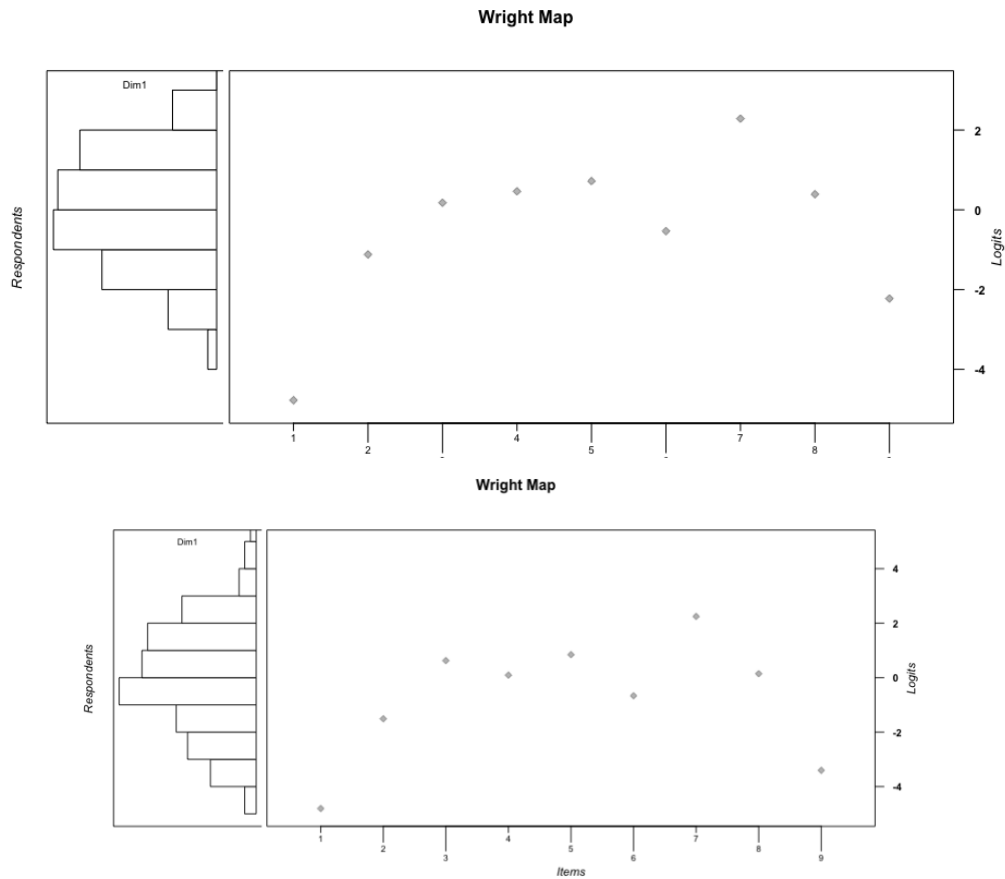
match amongst parents than adolescents. Stated differently, the items were able to

provide more information about parents because the items better captured the range of

variation in their media multitasking intensity. Thus, adolescents (teens) have a greater

range of scores that were not observed by the existing items. There are other attributes

that should and could be observed to gain a more accurate and complete understanding

of their media multitasking intensity. However, the existing items provide a more

accurate understanding of the parents

Figure 9. Teen MMTIQ Wright Map.

Discussion

Study 4's results provided mixed evidence for the security of validity claims based on the MMTIQ. While item comprehension and invalid response issues found in study 3 diminished, the measurement model in study 4 suggested that these some of the items were still problematic. Though beyond the scope of the current study, researchers investigated if these problems would have been uncovered had the current measurement models been specified in study 1 and study 2. In these additional analyses, study 1 did not demonstrate these issues, but the data in study 2 similarly best fit a unidimensional and similarly reflected issues with relevance items. However, in study 2, there had been evidence that polytomous items, or the partial credit model, was more suited for the data than the Rasch model. The potential error in observing scores on the questionnaire at only one time point in combination with the various issues with items obscured the ability to securely claim that the dichotomous model was preferential or provided substantial information. This may also be a consequence of the new population investigated. While studies 1 and 2 shared similar samples of adults on MTurk, the samples of study 3 and 4 differed by including college students, parents, and adolescents. Specifically, the reduced reading comprehension or range of interpretation of the key concepts among adolescents may explain the increased amount of error which the models could not explain in research question four. Therefore, the validation or the attempts to secure more evidence of validity for MMTIQ demonstrated that the current questionnaire was not universally capable of measuring media multitasking intensity.

Despite the weaker security of the MMTIQ for universal assessments, the four studies highlighted that media multitasking intensity as a construct and the MMTIQ as a measure benefitted from greater conceptualization compared to earlier measures of media multitasking. Thus, the four studies were successful in guiding theory development as well as measurement development for media multitasking intensity. They demonstrate the ways in which iterative measurement validation efforts can guide researchers to the fundamental questions that are not currently asked within current theories and those that need to be asked about the relationship between the ontology and epistemology of the construct (as explained in Chapter 2). Within the MMTIQ, the iterative process revealed issues with the conceptualizations of the construct of goal synthesis and difficulty. The interactions between these dimensions of media multitasking intensity provided theoretically valuable insights in studies 1-3. Yet, their interactions also appeared to be central to the challenges of securing evidence of a more valid measure of media multitasking intensity. Therefore, these studies in conjunction were interpreted to reflect the necessity for revisiting a cohesive and universal theory of media multitasking intensity, hypothesized relationships between items, response-option, and high-order dimensional levels, and epistemological assumptions, such as the method through which item-language is chosen. The theory that was developed, and the revised measure provided at the conclusion of these four studies, are the central foci of the remaining portions of the dissertation.

CHAPTER FOUR

A THEORY OF MEDIA MULTITASKING AND MEDIA

MULTITASKING INTENSITY

The purpose of this chapter is to first situate media multitasking in a larger theory of the changing media ecology. Then, it will explicate a theory of media multitasking intensity, and finally propose hypotheses about the relationship between media multitasking intensity and well-being, both to validate a measure of media multitasking intensity and to secure evidence for the theory.

Rationale: The Impetus for Theory Development

The development of new media and the evolution of devices facilitating experiences of media have transformed the role of media and media devices in society (Lang, 2013; Valkenburg, Peter, & Walther, 2016). Today, the economic growth of various cities and countries have depended on the proliferation of information and communication technologies that are imbedded and related to a multitude of facets of human goals and daily life. Digital divides are narrowing in terms of access and the global society has become dependent economically, politically, and interpersonally on these devices that are connected to the Internet (Pew Research Center, Internet & Technology, 2018; Statista, 2018). Media have become assimilated into even the most basic daily goals: eating, sleeping, breathing, working, studying, creating, connecting with others, relaxing, and exercising. These societal advances, which have produced a high volume of social, political, health, work, and entertainment information, are

potentially troubling: *can humans cope with and adapt to a society driven by devices with information processing, storing, and creating capabilities that far surpass them? How do people integrate media into daily activities and goals and can they leverage a media-saturated environment to their benefit*?

As the physical environment including homes, classrooms, and restaurants are filled with more media, there is a growing and unavoidable demand to process more mediated information while learning, speaking with family members, and eating. Simultaneously, these multifaceted devices facilitate the creation of goals that were previously unimaginable. These goals, which increasingly depend on media, increase interference or overlap between the digital and the physical (See Chapter 3; Rice et al., 2018; Walsh, 2016).

The role of media, and more frequently multiple media, within daily goals has been associated with an increased frequency in which media coincide, interrupt, or interfere with other tasks. Media, which are stimuli that represent or communicate information in ways that extend beyond the capacities of the human body, have become an undeniable presence that appear to compete with one another and other physical experiences. The ways in which media co-occur with or interrupt daily life have become known as media multitasking. As media transform daily experiences, this has increasingly become a topic of concern (Reinecke & Eden, 2017).

Though media multitasking is not directly identified as a pathological media behavior, nonetheless, few articles address benefits of media multitasking (Lang &

Chrzan, 2015). The majority of scholarship has concentrated on its implications for cognitive functioning (Chein et al., 2017; Van der Schurr et al., 2015). Yet, as Fisher and Keene (2019) have also indicated, the exciting research often implicitly considers the brain as a black box, and yet implies that increased involvement of media in daily life is fundamentally transforming cognitive functioning (i.e., through neuroplasticity). Media multitasking research has already assumed, though without asserting it, that media multitasking is a human adaptation to a changing environment. Without declaring this assumption, scholars have lacked lucid conceptualizations of media multitasking or have not defined it within the scope of their research, leading them to uncertainty in the interpretations of measures and thus empirical findings. For instance, perhaps the changes in media systems and thus human environment also need to change the definitions of success (Lang & Chrzan, 2015).

The remainder of this chapter will explicate a theory of media multitasking *intensity* to fill existing gaps in the conceptualization, measurement, and theoretical development of the phenomenon of media multitasking. This explication will begin by contextualizing the theory of media multitasking within theories of stress and flow in addition to those of cognitive load. These existing models become the assumptions on which the causal processes of media multitasking intensity develop. Then, assertions about media multitasking intensity will be articulated and linked to measurement and validation processes as defined in Chapter 2 and the theoretical developments that emerged from the iterative process beginning in Chapter 3.

Stress and Cognitive Load as a Framework

The intersection of models of flow and stress is the foundation for the theory of media multitasking intensity. These are more generalizable or universal models of *human-environment* interactions that describe and explain the mechanisms as well as the interpretations and value of media multitasking intensity. The theory of *media multitasking intensity* specifically explicates hypotheses about the functional (i.e., valuable) variation within experiences of media as they overlap with and interrupt experiences of the physical world. Thus, it is fundamentally a theory about the ways in which humans adapt and react to their environment.

Models of stress (Lazarus & Folkman, 1984; Lazarus, 1993) and flow (Csikszentmihalyi, 2000) evaluate individuals' resource-to-environment-created-demand ratios, while accounting for and demonstrating the critical role of human perception. These have secured more validity both in evidence of the quality of their measures and of the theoretical claims derived from their empirical investigations. Based in theories of both evolution and psychology, these theories state that stress or flow are outcomes of human *fitness* or (in)ability to adapt to the challenges and demands of their environment.

Stress occurs due to the presence of a set of demands that challenge, exceed, and/or are at least perceived to exceed the existing set of resources (Ellis & Del Giudice, 2013; Folkman, 2013; Lazarus, 1993; Lazarus & Folkman, 1984). These resources may be inherited or developed, internal or external. All of the following are

considered resources that allow humans to adapt to their environment: skills, abilities, knowledge, physical health, wealth, and social capital. These resources, however, are limited and can become strained, creating cognitive and physiological distress as is explicated in the limited capacity, bottlenecking, and threaded cognition models (Wang et al., 2015; Yeykelis et al., 2014).

Flow is a state in which demands and resources are synchronized. Flow reflects experiences of enjoyment, fulfillment, and full attention, which are derived from the perceptions of the environmental demand creating a challenge that is potentially surmountable (50/50 chance). The interaction between social, psychological, and biological systems create the manifestation and impacts, thus, the value of examining stress and flow. The mechanisms of the construct of both expand beyond a stimulus-response model. Exposure to a situation or stimuli does not cause the same quantity of stress or flow across all individuals. The context or situation is only one dimension or mechanism of stress; the other is the existing availability and accessibility of resources.

Beyond physical demand and resources, these theories of stress and flow prize human perception, revealing that the manifestation of these phenomenon and their subsequent effects are consequences of *subjectivity*. The state of stress and flow are traceable or have causal relationships to *subjective* experiences or those that are mind-dependent (Maul, 2017). Thus, it logically follows that stress and flow are interpretable or meaningful in relation to individuals' perceptions of their own experiences and capacities. Lazarus and Folkman (1984) theorize that the cognitive appraisal of

demands as exceeding resources not only creates but also increases the psychological and physiological experience of stress. Csikszentmihalyi (2000) similarly contends that flow is a subjective state, which may be reflective of both the "actual" relationship between task demand and cognitive resource, and perceptions of that relationship. Briefly connecting the constructs, media multitasking is conceptualized as a perception of interactions with and adaptations to a media-saturated environment: it entails applying resources to meet existing demands, as well as perceptions of existing resources and demands. Media multitasking can be framed as the resource allocation and demand created by the concurrence and competition amongst mediated stimuli or between mediated and physical stimuli.

Based on the theories of stress and flow, physical (i.e., physiological and cognitive) consequences of the demand-resource ratio have a reciprocal relationship with the perception of the demand-resource ratio and consequent behavior. Perceptions of challenges can lead to coping, which aims to increase the resource available to meet the demand. This coping can occur cognitively via reframing perceptions and behaviorally via applying resources. When coping or obtaining and applying resources, individuals may find that they reach a state of flow in which their resources meet the demands of the environment - in stress literature this is often referred to as resilience (Bonano, 2004; Carver, 1989). They may even feel that they gained resources or coping capacities (i.e., strategies for obtaining or applying resources) such that they have learned and now no longer find this previously stressful situation as challenging.

Experiences of flow, therefore, are associated with learning. Though flow often refers to a short-term state, within the stress literature the experience of obtaining new resources in reaction to a stressor and then growing is referred to as thriving. In relating this to media multitasking, the challenges or demands that are derived from a media-saturated environment may invoke coping behavior like those predicted by the "law of less work" (Wang et al., 2015). People may find these challenges surmountable if they have sufficient existing resources or perceive having sufficient resources. The availability of and perceptions of resources may be dynamic. These resources may develop as a result of coping skills, and experiences of flow such that they reflect growth from their experiences of challenges.

In summary, the models of stress and flow provide a fundamental set of assumptions: 1) experiences of tasks are consequences of the reciprocal relationships between behavior and cognition, 2) experiences of one's environment depend on demand-to-resource ratios as well as perception of this ratio and consequent coping behaviors, 3) challenges are not inherently harmful but are a natural and fundamental aspect of human growth and satisfaction, and 4) humans adjust their behaviors and cognition to adapt to challenges. These assumptions guide the theory of media multitasking intensity.

The Theory of Media Multitasking Intensity

The theory of media multitasking intensity hypothesizes that media multitasking is a multidimensional construct involving the interaction between cognition and

behavior, which varies meaningfully on both dimensions *intensively* (i.e., multiplicatively). In addition to providing hypotheses about the causal mechanisms of media multitasking and its effects, the theory aims to remedy existing ambiguity about the role of behavior, cognition, and the perception of cognition and behavior in the manifestation, measurement, and effects of the phenomenon. Therefore, this section will first define media multitasking and propose and define media multitasking intensity. Then, this section will explicate the value of media multitasking intensity and the measure of media multitasking intensity (i.e., the MMTIQ).

Defining Media Multitasking

Media multitasking is the perception of the co-occurrence of or interference between two or more tasks, when at least one of these tasks' stimuli is a form of mediated information (i.e., media). *Tasks* are the allocation of senses such as vision, audition, tactile reception, vestibular sensation, and proprioception to stimuli. They can vary in duration but fundamentally require time for orienting to a stimulus and then applying sensory capacities to process it. This definition diverges from colloquial definitions of a task in that tasks may be neither conscious nor goal-directed. *Media* are defined as tools through which information can be communicated, which extend communication beyond human's unaided physical capacity. *Media tasks* therefore are the allocation of senses to mediated information. *Physical tasks* are the allocation of senses to physical stimuli in the environment. Media multitasking can involve multiple media tasks alone as well as media tasks and physical tasks. Media, therefore, are also

146

mind-dependent or subjective. They exist in the physical environment but are identifiable by comparing perceptions of human capacity for information sharing and absorption and tools that extend beyond human capacity. Thus, the proliferation of media and devices that facilitate media increase the number of stimuli that involve information otherwise not communicable without these tools, creating an influx of potential media tasks. Still, media tasks like physical tasks, vary in their experiences because of attributes that impact the task-demand-to-resource ratio and perceptions of the task-demand-to-resource ratio.

Defining Media Multitasking Intensity

Media multitasking *intensity* is derived from the conceptualization of media multitasking as a latent, intensive construct that is multidimensional. *Intensity,* therefore, is an attribute of the multiple dimensions of media multitasking: the behaviors and cognition that correspond with various sensory resource allocation and the behaviors and cognition that correspond with various *perceptions of* sensory resource allocation. Media multitasking is contended to meaningfully vary in its *intensity. Intensity* varies continuously and involves four dimensions.

There are four theorized dimensions of media multitasking intensity (refer to Figure 10), which capture variation in attributes of media and physical tasks and collectively explain variance in media multitasking: *co-occurrence and interruption*, *difficulty (or demand), intentionality,* and *relevance. Task co-occurrence and interruption* involves the perceived temporal overlap or interference between multiple

147

media tasks or media tasks and physical tasks. This overlap or interference between multiple media tasks can occur within (Ralph & Smilek, 2017; Yeykelis et al., 2014; Yeykelis et al., 2018) and across devices (Brasel & Gips, 2011; Pool, Koolstra, & Voort, 2003; Segijn et al., 2017; Van Cawenberge, Schaap & van Roy, 2017). *Difficulty* involves the perceived degree of challenge or demand imposed by media and physical tasks. *Intentionality* is the perceived degree of purpose, motivation, and value of engaging in media and physical tasks. *Relevance* examines the perceived shared purpose amongst tasks as well as the degree to which the tasks complemented or enhanced one another. The interactions amongst these dimensions are theorized to be fundamental components and potent predictors of variance in media multitasking. These dimensions are also informed by theories of limited cognitive resources or information processing capacities in addition to theories of flow and stress, which contend that media multitasking varies meaningfully as it demands and exhausts resources (Lang & Chrzan, 2015; Wang et al., 2015; Yeykelis et al., 2014). Notably, these dimensions were first theorized as divisions of attention, difficulty, goal-relevancy, and task-relevancy in Chapter 3. They have transformed through the iterative process of measurement validation and theory development, which revealed insights including that sensory allocation does not necessitate divisions of attention, and goal-relevancy assumed that each experience involves a goal or intention.
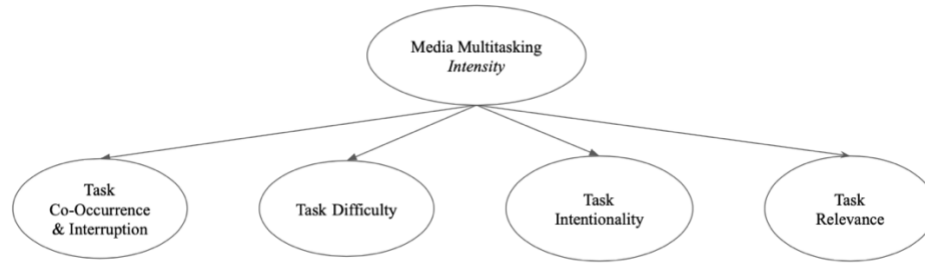
Figure 10. The four-dimensions conceptual model of media multitasking intensity.

The *co-occurrence and interruption, demand, intentionality,* and *relevance* of media and physical tasks explain variation in media multitasking because they impact perceptual load (i.e., sensory processing) and cognitive load (i.e., effortful processing) of media multitasking. While each dimension individually has theorized impacts on demand-resource ratios and the perception of demand-resource ratios, they are theorized to become valuable and more interpretable in relation to each other. Thus, these four dimensions' impact on the demands and resources are interdependent. Perceptual load refers to the quantity of perceptual cues or stimuli that must be processed and the relevancy amongst these cues (Lavie, 1995). These cues are at a level of information processing that occurs before cognition. Thus, perceptual load is predicted to have stronger relationships with task co-occurrence and interruption, intentionality, and relevancy than with difficulty. Contrastingly, cognitive load refers to the attention and working-memory demand of the information that becomes processed (Wang et al., 2015). Cognitive load is more consciously experienced. Cognitive load is predicted to

have stronger relationships with task co-occurrence and interruptions, difficulty, and relevance, than with intentionality. These relationships are demonstrated in Figure 11. Perceptual load and cognitive load overlap in their effects on task co-
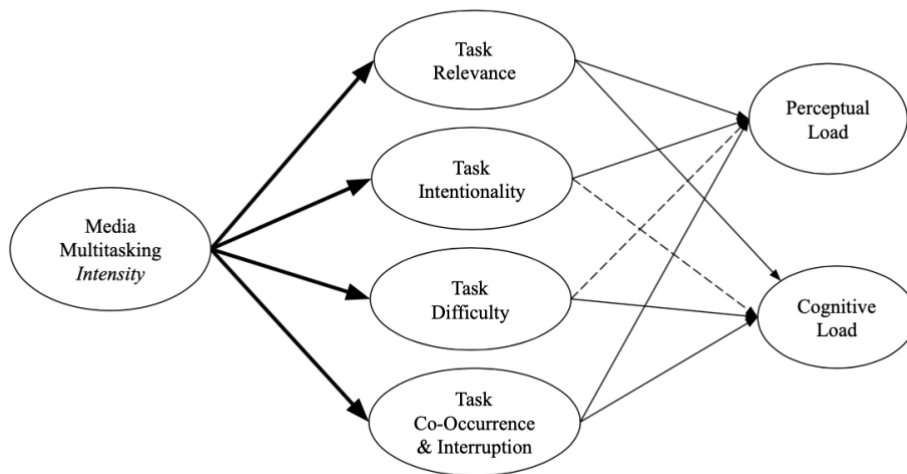
Figure 11. The relationships between media multitasking intensity dimensions and perceptual and cognitive load.

occurrence and interruptions and relevancy, but are predicted to diverge in the strengths of their relationships with difficulty and intentionality. They also have disparate relationships with perceptions of the demand-resource ratio.

150

The interactions between these dimensions on the continuum of *intensity* occur as shown in the construct map of media multitasking intensity (pictured in Figure 12). For instance, task difficulty and intentionality interact such that difficulty can transpose the effects of intentionality. When task difficulty decreases and task intentionality increases, intensity increases. Imagine the difference between breathing and observing the environment automatically as opposed to intentionally. The challenge of intentionally breathing or focusing on one's environment is often involved in mindfulness and meditation. In other words, perceived task-demand increases because of the cognition involved in the behavior, not because the behavior itself is difficult. On the other hand, as task difficulty increases and task intentionality decreases, intensity also increases because the behavior is difficult and yet the cognition involved does not support meaningful processing of the environment. For instance, imagine being an amateur baker who is baking a cake. As the engagement in the essential tasks such as observing the rise of the cake in the oven becomes less intentional and attention given to the oven is only provided randomly rather than as planned, the intensity of the experience increases. This is because the perception of the demand of the task is more likely to exceed the resources one is applying. In other words, the probability that an individual will check the oven too late and burns his cake increases. Specifying every permutation of the interactions amongst these dimensions is beyond the purpose of this section and scope of this paper, they are visualized in Figure 12.

Co-Occurrence and Interruption	Difficulty	Intentionality	Relevance
High CI	High D	Low I	Low R
High CI	High D	High I	Low R
High CI	High D	Low I	High R
High CI	High D	High I	High R
High CI	Low D	Low I	Low R
High CI	Low D	High I	Low R
High CI	Low D	Low I	High R
High CI	Low D	High I	High R
Low CI	High D	Low I	Low R
Low CI	High D	Low I	High R
Low CI	High D	High I	Low R
Low CI	High D	High I	High R
Low CL	Low D	High I	Low R
Low CL	Low D	Low I	Low R
Low CL	Low D	High I	High R
Low CL	Low D	Low I	High R

Higher Intensity → **Lower Intensity** (vertical axis)

Figure 12. Construct map of media multitasking intensity.

It is vital to address the interpretations of the variance in the media multitasking *intensity* across the continuum from *high* intensity to *low* intensity. *Higher* media multitasking intensity involves increasing the difficulty of information processing and/or sensory demand by increasing the volume and variety of input that cannot become integrated (i.e., are irrelevant) and thus competes for processing resources. *Lower* media multitasking intensity involves decreasing the difficulty of information process or the sensory demand by decreasing the volume and variety of input that become integrated (i.e., are irrelevant) and thus do not compete for processing resources

Measuring Media Multitasking Intensity

Media multitasking *intensity* is a subjective attribute of experience and an outcome of a person-by-context interaction. Therefore, any measure of media

multitasking intensity would include information about the *context* as well as the

individual. Because the four dimensions of media multitasking *intensity* depend on

perception of tasks, it is recommended that scholars employ self-report methods.

However, there are two critical challenges to accurate self-report measurement. The

media-saturated environment has made media experiences mundane (Potter, 2018), and

the cultural shifts in the definitions of media more colloquial, collectively posing

challenges to non-experts' abilities to report on their media tasks (see Chapter 3

subsection Study 3). Moreover, self-report requires self-awareness, working-memory,

and attention, all of which have been predicted in the past to relate to media

multitasking frequency and may have relationships with dimensions of intensity such as

task intentionality and relevance.

   Media Multitasking Intensity Questionnaire (MMTIQ). This dissertation

develops, employs, and validates an ecological momentary assessment (EMA;

Hedstrom & Irwin, 2017; Hektner, Schmidt, Csikszentmihalyi, 2007) of media

multitasking intensity called the Media Multitasking Intensity Questionnaire (MMTIQ).

The development of the items, response process options, and formatting of this

instrument is described in Chapter 3, and more specific details are provided in Chapter

5. The Media Multitasking Intensity Questionnaire (MMTIQ) is a self-report instrument

of *task co-occurrence and interruption*, *task difficulty, task intentionality,* and *task

relevancy.* The MMTIQ is designed to focus on specific contexts by constraining the

duration of time on which the participant reflects, and asking participants to assess their

most recent engagement in tasks, hence it is categorized within self-report methodology

as an ecological momentary assessment (EMA).

If this questionnaire is applied within other experience sampling methods (ESM)

such as diary logs or event-based reporting, this would invite different sources of error

that should be examined in order to secure evidence of claims of validity. Given the

secured evidence of validity claims obtained in previous research employing the

MMTIQ (see Chapter 3), administering the questionnaire only cross-sectionally would

not provide enough information to make some claims about an individuals' media

multitasking intensity. Thus, the current dissertation investigates whether this four-

dimensional model of media multitasking *intensity* theorized here fits the observations

obtained from ecological momentary assessments of behavior and cognition better than

a three-dimensional, two-dimensional, or unidimensional model. A unidimensional

model may suggest that these dimensions do not interact with one another and do not

provide unique information. On the other hand, a two-dimensional model may

demonstrate that task co-occurrence and interruption and task difficulty can be

combined into a dimension of task demand which is more related to cognitive load, and

that task intentionality and task relevancy are more related to one another (i.e., share

similar information) such that they can be combined into one dimension of goal-

synthesis. Finally, the three-dimensional model examines if co-occurring and

interrupting tasks and task difficulty are distinct, but if intentionality and relevance of

tasks are truly one dimension. The unidimensional, two-dimensional, and three-

dimensional models are motivated by findings in previous iterations of the media multitasking intensity questionnaire described in Chapter 3, and are examined to secure evidence that the four-dimensional model is the best explanation for variation in media multitasking intensity.

Hypothesis 1: Data obtained from the MMTIQ will demonstrate a better fit to a four-dimensional model (Figure 10) than a three-dimensional, two-dimensional, or unidimensional model.

Although the current theory of media multitasking intensity contends that the four-dimensions of co-occurrence/interruption, difficulty, intentionality, and relevance are fundamental attributes that explain variation in media multitasking, previous research has primarily examined that the frequency of media multitasking demonstrated the most valuable variation in its effects. While previous investigations of media multitasking extensity or frequency have resulted in inconsistent and inconclusive findings (Chein, Wilmer, & Sherman, 2017; Jeong & Hwang, 2017; Lang & Chrzan, 2015; Van de Schuur et al., 2015; Wiradhany & Nieuwenstein, 2017), it is still possible that including the frequency in which a person engages in higher or lower media multitasking intensity would provide valuable information. Therefore, it was necessary to examine whether including frequency of the indicators, as opposed to the presence of the indicators, of each of these four dimensions (*i.e., co-occurrence/interruption, task difficulty, intentional allocation, and task relevancy)* improved the measure or explained valuable variation.

155

Research Question 1: Does the presence and absence of indicators (i.e., binary), or does the frequency of these indicators (i.e., polytomous), provide a better fit to the data (self-reported observations)?

Due to the context-specific attributes observed in the MMTIQ, longitudinal investigations of individuals across a variety of contexts would be warranted to evaluate the security of the validity of claims, which apply media multitasking intensity as an attribute of an individual alone. Longitudinal designs could identify the portion of variance explained by the individual as evidence of stability. Yet, this would depend on the variation of contexts within the longitudinal investigation. Previous research has demonstrated that media multitasking can be motivated by context because people aim to conserve their cognitive resources when they perceive them as limited, and aim to expend more resources when they experience the demands in their environment more easily (see Chapter 3; Csikszentmihalyi, 2000; Lang & Chrzan, 2015; Ralph & Smilek, 2018; Wang et al., 2015; Yeykelis et al., 2014). Thus, longitudinal or repeated observations of an individual's media multitasking intensity in a specific context may allow for early investigations to secure some evidence of validity for media multitasking intensity as an attribute of an individual though within that context. For instance, using the MMTIQ within the context of bed time longitudinally can allow for generalizations about people who are *bed-time* media multitaskers. This may not generalize to morning media multitasking behaviors. Therefore, the current dissertation examined the MMTIQ longitudinally to examine if there was stability in an individual's

media multitasking intensity within the context of weeknights over the course of a week.

Hypothesis 2: A significant portion of the variation in MMTIQ observations across time will be attributable to the individuals.

Individual Differences in Media Multitasking Intensity

This section will review the two predominant individual differences that are theorized to contribute to variation in the behavioral and cognitive dimensions of media multitasking: executive functioning and self-regulation. Then, these two individual differences will then be related to development and thus differences between adolescents and their parents.

Executive functioning (EF)

Executive functioning refers to attention, working memory, and inhibitory capacities that are involved in goal-directed cognition and behavior (Nigg, 2017). In a recent review, Nigg (2017) summarized the definitions of executive functioning as functions that support rule-following or top-down processes. This includes selective attention, shifting attention (task switching), filtering information, response inhibition, and sustained attention. These capacities allow people to anticipate, plan for, and solve problems (Diamond, 2013), and thus have been theorized as predictive of media multitasking.

Executive functioning has been the predominant individual difference related to media multitasking since the seminal work of Ophir et al. (2009) found a diminished

filtering of irrelevant information occurred amongst the most frequent media

multitaskers, termed the breadth-bias. The multitude of studies that aimed to replicate

these findings have failed to produce consistent results (Chein et al., 2017; Wiradhany

& Nieuwenstein, 2017). Within the existing literature, some studies have associated

heavy media multitaskers with advanced executive functioning (Alzahabi & Becker,

2013; Cordoso-Leite et al., 2016; Kononova, Joo, & Yuan, 2016; Minear, Brasher,

McCurdy, Lewis, & Younggren, 2013), while others found it was associated with

diminished executive functioning (Baumgartner, Weeda, & Van der Heijda, 2014;

Jeong & Hwang, 2016; Ophir et. al., 2009; Uncapher, Theiu & Wagner, 2016). In their

meta-analysis, Wiradhany and Nieuwenstein (2017) conclude that when existing

findings are adjusted for sample size, the estimated association between executive

functioning and media multitasking across studies neared zero. The methodological

explanations for these inconsistencies, including error in self-report of media task

engagement over a week and comprehension of concurrent engaged in media tasks,

have been described in Chapter 3. However, in this section, it is contended that the

unobserved heterogeneity in media multitasking *intensity* within these examinations

may explain the lack of consistent relationship. In other words, these inconsistent

findings are explained by the relationship between *co-occurring or interrupting tasks,

task difficulty, task intentionality, and task relevancy*.

The extensive model of media multitasking assumes that all media task

combinations are equivalently weighted. However, this overlooks cognitive and

behavioral variation in the ways in which tasks co-occur or interfere with one another, differ in difficulty, intentionality, and relevance, and thus fails to consider the differences in the cognitive and perceptual effort they require. It is predicted that poorer executive functioning, or less attentional and working memory capacity, selective attention, behavioral inhibition, and filtering of irrelevant information, would increase the likelihood of engaging in tasks that overlap and interfere with one another. In other words, poorer executive functioning is expected to predict a greater degree of overlap and interference between tasks or more shifts between tasks because of lacking capacities to direct selective attention. Those with poorer executive functioning are also more likely to engage in difficult co-occurring or interfering tasks, due to their inability control their attention and inhibit behavior. They may even be more likely to perceive more tasks as difficult because of their lower attentional and working-memory capacities. Both the increased difficulty and co-occurrence/interruption amongst tasks are likely to occur with intentionality because it is an outcome of a lack of attentional and behavioral control. *Finally, e*xecutive functioning is necessary for identifying and filtering out irrelevant information. Thus, it is predicted that those with poorer executive functioning will report engaging in tasks with less relevancy.

In summary, deficient executive functioning is likely to predict higher media multitasking intensity demonstrated by increased task occurrence and interruption and task difficulty, but decreased task intentionality and relevancy.

159

Hypothesis 3a: Executive functioning will be negatively associated with a) task co-occurrence and interruption and b) task difficulty.

Hypothesis 3b: Executive functioning will be positively associated with c) task intentionality and d) task relevancy.

Self-regulation

Though only several studies have considered the role of self-regulation in media multitasking (Reinecke et al., 2018; Xu et al., 2016; Zhang, 2015), self-regulation has been associated with increased likelihood to engage in healthy behaviors (e.g., eating healthy) and to avoid unhealthy behaviors (e.g., binge drinking) (Hagger, Wood, Stiff, & Chatzisarantis, 2009; John & Gross, 2004). Resilience (i.e., the ability to overcome stress) to stressors either psychological or physiologically has been theorized as an outcome of human capacity to allocate internal and external resources in order to adapt to one's environment (Afifi, Merill, & Davis, 2016; Campbell, 2010; Eisenberg, 2017; Folkman, Lazarus, Gruen, & DeLongis, 1986; Floyd, Pauley, & Hesse, 2010; LaRose, 2015; Lang, 2015; Lavee, McCubbin, & Patterson, 1985). Schilab (2017) thus argues that a media-saturated environment requires self-regulation to adapt to its best uses. The ability to allocate resources toward the goal of adaptation includes the capacity to regulate one's cognition, emotion, and behavior.

Self-regulation, thus, is defined as the ongoing intrinsic process of managing mental and physiological states via altering cognitive, emotional, and behavioral responses to achieve personal goals (Nigg, 2017). It includes deliberate or top-down

processes, bottom-up or reactive/automatized processes, and the interplay between the two, that are used to monitor and adjust one's behavior in order to achieve explicit or implicit goals or goal states (Baumeister & Vohs, 2017; Nigg, 2017). Thus, there are three components of self-regulation: the ability to monitor behavior, the cognition about and emotional appraisal of behavior, and the following adjustment or adaptation of behavior (Baumeister & Vohs, 2017; LaRose, 2010, 2015; Nigg, 2017). All three can occur both via conscious thought and automatic processes, which function together and react to one another to drive goal-oriented behavior. Self-regulation is a capacity that depends on executive functioning capacities. Yet, it is distinguished from executive functioning because it includes bottom-up processes or automatic reactions to stimuli and it refers to applying attentional and working memory capacities to the self (Nigg, 2017).

Conscious and unconscious forms of self-regulation may both be limited. The top-down aspects of self-regulation or effortful self-monitoring, evaluation, and behavioral control capacities are especially limited resources that require energy and glucose expenditure (Baumeister & Vohs, 2017; Lang, 2009). Due to the limited resources of energy and thus executive functions such as attention, working-memory, and behavioral inhibition (Barkley, 2012), it is possible to experience depletion after instances of self-regulation. According to ego-depletion models, as self-regulatory resources are expended more frequently, people are more likely to experience failures of self-regulation. People are motivated to allocate self-regulatory resources to maintain

their energy and conserve resources and thus to strategically use their attentional resources with media or reframe their media use (Baumeister & Vohs, 2017; Gross, 2014; Lang, 2009; Wang et al., 2015).

Previous research, though sparse, does support the relationship between media multitasking and self-regulation. Xu et al. (2016) revealed that self-control, often defined as a stable trait-like capacity for self-regulation, not only predicts the frequency of media multitasking, but also the types of media multitasking in which people tend to engage. Those with lower self-control reported engaging in more cognitively demanding forms of media multitasking. In addition to this self-report evidence, Szumowska, Polawska-Boruc, Kis, Osoweicka, and Karamarczyk (2018) conducted an experiment where participants were asked to multitask in the lab. Only heavy media multitaskers who had low self-regulation capacities a) task-switched more often and b) performed worse on their tasks. These findings extend to the relationship between self-regulatory capacities and media multitasking in consequential contexts such as lectures or within classrooms (Rosen, Lim, Carrier, & Cheever, 2011; Zhang, 2014). Schutten, Stokes, and Arnell's (2017) findings echoed these results. Their survey found that frequent media multitasking was associated with less reported self-control and greater impulsivity or less task intentionality. These studies in conjunction suggest that deficient self-regulation would predict higher degrees of *co-occurrence and interferences between tasks* and *task difficulty,* but lower degrees of *intentional*

allocation and *relevance*. Therefore, it is hypothesized that insufficiencies in self-regulation will also predict more frequent engagement in intense media multitasking.

Hypothesis 4a: Self-regulation will be negatively associated with a) task co-occurrence and interruption and b) task difficulty.

Hypothesis 4b: Self-regulation will be positively associated with c) intentional allocation and d) task relevance.

In addition to testing these hypotheses about the relationship between executive

functioning and self-regulation separately as predictors of media multitasking

intensity, the current dissertation aims to investigate if one is a greater explanatory

variable than the other. Due to the specificity of self-regulation, its role in allowing

people to master their media uses, and to learn from adversity and adjust (Rice et al.,

2018), it is predicted that self-regulation will be a stronger predictor of media

multitasking than will executive functioning.

Hypothesis 5: Self-regulation will explain more variance in all four

dimensions of media multitasking intensity than will executive functioning.

Adolescents: A Vulnerable Population

Self-regulation capacities develop across the lifespan. While adolescents begin

to develop their self-regulatory capacities, their developmental stage poses new

challenges. Adolescents' insufficient self-regulatory capacities are juxtaposed with their

need for autonomy. Adolescents are also generally motivated to discover their identity

or role within their social environment. Given today's adolescents have perpetual access to social networks and social information, their developmental goals can serve as internal impetuses for more intense media multitasking. The increased likelihood of depleted self-regulatory resources and the preoccupation with social networks (especially their online social media networks) can prompt adolescents to engage in more irrelevant media use when multitasking. For instance, their concerns about their peers may motivate internal interruptions or self-interruptions to their completion of homework, family dinners, and social events (Turkle, 2011). Additionally, members of their increasing social networks, as adolescents interact with others in new contexts (boyd, 2014), may be equally likely to have lesser self-regulatory capacities. Their behaviors could serve as external interruptions and motivations for media multitasking with demanding tasks. Due to their growing peer networks, adolescents can experience significant demands on their emotional regulation capacities and information processing capacities that are still underdeveloped (Carrier, Black, Vasquez, Miller, & Rosen, 2015; Konijn, Veldhuis, Plaisier, Spekman, & den Hamer, 2015). Thus, the interaction between adolescents' self-regulation capacities, developmental stage, and the media-saturated environment increases their likelihood of engaging in more intense media multitasking and experiencing the most detrimental effects.

Scholarship on adolescents' well-being and media use provides some evidence about how their motivations for media use can lead to problematic media multitasking. Reinecke and colleagues' (2017) survey of media multitasking across the lifespan

reports that youth (14-25) engage in more frequent media multitasking and experience more digital stress than adults of any age group. In their study, youth reported engaging in Internet multitasking more often due to social pressure and fears of missing out (FOMO). They also reported feeling more overwhelmed by technology than did adults. Afifi et al. (2017) similarly find that adolescents demonstrate more physiological and psychological stress due to their technology use than do their parents. Corresponding with previous research (Lee, Son & Kim, 2015; Steinfield, Ellison & Lampe, 2008), this stress was not associated with the frequency of their social media use but rather with the size of their social networks on social media. This suggests that their extended social network and social pressures may be a fundamental motivator for harmful media multitasking. While the effect of developmental stage has been previously investigated, the current dissertation is the first (to my knowledge) to provide self-regulation as an explanation for why and how these motivations lead to problematic media multitasking. Therefore, it is predicted that the relationship between self-regulation and media multitasking *intensity* will be moderated by whether one is an adolescent (Valkenburg & Peter, 2013).

Hypothesis 6: The relationships between self-regulation and intense media multitasking will be moderated by developmental stage, such that it is stronger for adolescents than for adults (parents).

Media Multitasking Intensity and Stress

165

Within the existing literature, the propensity to media multitask has been related to diminished relational satisfaction between partners (McDaniel, 2015; Wajcman, Bittman, & Brown, 2008), lower relational satisfaction between parents and children (McDaniel & Radensky, 2017; Radensky et al., 2016), diminished sleep, and increased unhealthy eating. In these studies, people can experience conflict, loneliness, and frustration due to either their own or their loved ones' media multitasking habits. In a few studies, increased media multitasking has also been associated with diminished health such as experiences of digital information overload (Reinecke et al., 2016), social stress (Pea et al., 2012; Xu et al., 2016), and even depression and anxiety (Becker, Alzahabi, & Hopwood, 2013; Rosen, Whaling, Carrier, & Cheever, 2013). Some of this scholarship has revealed that the context in which one media multitasks (McDaniel & Radensky, 2017; for example, frequently engaging in media tasks at dinner) and the types of media used while media multitasking (Lang & Chrzan, 2015; Lau, 2017; Wang et al., 2015; for example, irrelevant and difficult task combinations) predict these negative outcomes. These contexts, in summary, are those in which the concurrence and interference between media tasks and physical tasks are irrelevant and difficult (i.e., answering a work-email when on a romantic date) or irrelevant and less intentional (i.e. such as watching TV before and scrolling through Instagram while aiming to sleep).

Media multitasking intensity increases as tasks co-occur and interrupt one another, are more difficult, less intentional, and less relevant, increasing goal conflicts and requiring more perceptual and cognitive resources. Thus, high media multitasking

intensity is predicted to result in demand-resource ratios in which demand far exceeds resources, resulting in fatigue, frustration, and stress. It is predicted that those who engage in *high media multitasking intensity* will perceive greater amounts of general stress (Lee, Son and Kim, 2015; Misras & Stokols, 2012; Reinecke et al., 2017; Wang et al., 2015).

Hypothesis 7a: Task co-occurrence and interruption as well as task difficulty during the week will predict more stress.

Hypothesis 7b: Task intentionality and task relevancy will predict less stress.

CHAPTER FIVE

METHOD

Participants

Six hundred and ninety participants were recruited via Qualtrics panels, most of whom Qualtrics reported as dwelling in the Midwest or Eastern states of the United States of America. The majority of parents identified as females or mothers (n=585; 84.8%). To qualify for the study, parent-adolescent dyads were required to live together full-time, to have access to at least one technological device at home, and to be available to check their email and answer questions at a random time between six and nine PM on weeknights. Of the adolescents, 41.6% reported being males, 48.6% reported being females, and 10% said these options were not applicable to them. Of the parent-child dyads 44% were mothers and daughters, 32.6% were mothers and sons, 9% were fathers and sons, 4.5% were fathers and daughters. The majority (56.7%) of parent-adolescent dyads reported their ethnicities to be White or European, 18.5% identified as Black or African-American, 6% identified as Latinx or Hispanic, followed by 6% who involved a White parent with a multi-ethnic child, and 3% who identified as White or European with a Latinx or Hispanic child. The median income of these homes was between $40,000 – $59,999, but one fourth of the homes have an annual income of less than $40,000 and one fourth of the homes have an annual income grossing larger than $100,000.

Of the 690 dyads who completed the screener, 648 parents and teens at least completed one following survey of the study. These 648 parent-adolescent dyads participated in a one-week longitudinal intensive study. The average adolescent age was about 15 ($M = 14.92$), and the average parent age was about 41 years old ($M = 41.42$).

Procedures

Screening. In an initial screening survey, a short questionnaire asked parents about the number of children for which they are currently a primary caretaker and/or live with, the age of their children, whether they live with their children full-time, as well as questions about whether their adolescent owns a smartphone, laptop or an iPad, the number of devices they have in their home, and demographics such as ethnicity, sex, age, income, and education level. Parents were screened out of the study if they did not have a child between the ages of 13-18 with whom they currently lived full-time. Parents were also screened out of the study if their adolescent did not have primary or secondary ownership of a technological device (e.g., smartphone, computer, laptop or tablet).

In the screener, first parents answered the questions; if they were deemed eligible, they were provided a description of the study and asked to let their adolescent answer the remainder of the questions. After both parents and adolescents completed questions about device ownership and demographics, they were provided a description of the one-week study and then were prompted to select one of up to three upcoming weeks to participate in the study. Data collection and screening began in mid-November

2018 and continued until late January 2019. Recruiting and screening occurred for a week prior to the participants' selected week. Data collection occurred on a total of seven weeks; no weeks in which there was a holiday such as Thanksgiving, Christmas, or New Year were used for data collection.

Pre-survey. On the Monday of the given week on which parent-adolescent dyads had selected to participate in the study, they received an email to complete the pre-survey. They were paid a dollar at this point to incentivize participation and build greater trust. The pre-survey questionnaire included self-report measures of executive functioning, self-regulation, and general stress. At the end of the pre-survey, participants were reminded to expect emails every evening between six to nine pm for the next four days: Tuesday, Wednesday, Thursday, and Friday.

Experience sampling surveys. On the Tuesday of the given week on which participants had selected to participate in the study, they received an email reminding them that later that night between 6 – 9 pm they would receive an email with a survey to complete immediately. Qualtrics managed the participant list and contacted participants via Product Report Card. On Tuesday, Wednesday, Thursday, and Friday night of that week, Qualtrics assigned each parent and teen to a randomized time between six and nine pm at which they received the link to a questionnaire. This questionnaire on average took 10 minutes and measured intense media multitasking behaviors during the last half-hour as well as mood and parental mediation throughout the day.

Post-survey. On the following Monday from which parent-adolescent dyads began their participation in the study, parents and adolescents who completed at least three experience sampling surveys received an email with the link to a post-survey. The questionnaire included self-report measures of general stress and internet media multitasking behaviors in the last week. Parents and teens were allowed to complete the post-survey between Monday and Wednesday of that week. In other words, answers were accepted up until ten days after the study for that parent/child dyad began (on the Wednesday, one week and half from the beginning of the study).

Reward & compensation. Before beginning the study, parents and teens were provided $1 dollar for completing the screener and $1 dollar on the date of the pre-survey. Thus, all parents and teens participating in the full study were rewarded with at least two dollars. If parents and teens both completed the pre-survey, post-survey and at least three of the experience sampling surveys with a good faith effort, they each received $20 or a total of $40 as a dyad. A good faith effort required that parents and teens pass at least four of the seven potential attention checks. Open-ended questions such as "what do you and your friends do for fun and why," were included on each survey (i.e., screener, pre-survey, experience sampling surveys from Tuesday to Friday, and the post-survey. Answers with less than three words, irrelevant information, or those which failed to provide reasoning for their answers were+ considered failing an attention check.

A good faith effort also required completing the questionnaires at an appropriate time. Completing experience sampling questionnaires or the pre-survey on the wrong date was problematic for data analysis. Therefore, those dyads in which parents or teens completed questionnaires on the incorrect data were also not provided the full study reward. Parents received their compensation via Product Report Card (a vendor used by Qualtrics) and adolescent received their compensation independently via a digital gift card sent to a personal email chosen by the adolescent.

Measures

Pre-survey.

Executive functioning. Executive functioning was defined as the general top-down processes or conscious effortful uses of cognitive capacities such as attention and working memory. The 12-item Attention-Related Cognitive Errors Scale (ARCES; α = .92) created by Carriere, Cheyne, and Smilek (2008) is a self-report measure of daily failures in attention and memory that asks participants about every day scenarios such as "I have accidentally mixed up targets of my action (e.g., pouring or putting something into the wrong container)" and " I have got to the fridge to get one thing (e.g., milk) and taken something else (e.g., juice)." This self-report instrument was used to assess executive functioning because of previous validation efforts that have demonstrated its relationships with executive functioning assessments tasks (Smilek, Carriere, & Cheyne, 2010). Participants rate how frequently they experience these

scenarios on a Likert-scale from (1) never to (5) very often. Therefore higher-scores on the ARCES reflects *less executive functioning*.

Self-regulation. Self-regulation is defined as the capacity for introspection, evaluation, and behavioral adjustment. For parents, self-regulation was measured via the adapted shortened self-regulation questionnaire (SSRQ; Carey, Neal & Collins, 2004) based on the Self-Regulation Questionnaire (SRQ; Brown, Miller, & Lawendowski, 1998). This SSRQ questionnaire includes 21-items asking about general or global experiences with self-regulation ($\alpha = .59$). This includes items such, "I'm able to accomplish goals I set for myself," "It's hard for me to notice when I've had enough (alcohol, food, sweets)" and "I give up easily." Participants indicate their responses on a 5-point Likert-scale from (1) strongly disagree to (5) strongly agree. For adolescents, self-regulation was measured via the Adolescent Self-Regulation Inventory (ASRI) by Moilanen (2007) ($\alpha = .94$), which involved 13-items rated from (1) not at all true for me to (5) really true for me. This included items such as, "when I'm bored, I fidget or can't sit still," and, "during a dull class, I have trouble forcing myself to start paying attention." Items on the SSRQ and ASRI were both scored such that increased scores reflected *lower self-regulation*.

Demographics.

Gender. Parents and children were asked whether they identify as male, female, transgendered males, transgendered females, or they prefer not to say.

Age. Participants were asked to report their age in years.

Household Income. One item asked parents to report on the household annual income; response options ranged from less than $20,000 to more than $250,000.

Race and Ethnicity. One categorical item was included to capture the race of both parent and children; options included White, Black, Hispanic, Latino, Native American, Asian, Arab, Pacific Islander and Other, where multiple choices are permitted.

Experience Sampling Method Survey. The media multitasking intensity questionnaire (MMTIQ; Guttman $\lambda6 = .84$) is a self-report experience sampling measure of media multitasking behaviors indicating the extent to which the respondent perceived the multitasking as involving difficult, relevant and intentional tasks. (See Chapter 3 for explication and validation of the MMTIQ.) The MMTIQ is intended for use within experience sampling methodologies in order to be sensitive to context and reduce recall errors. Specifically, participants were asked a set of questions that allowed them to narrate their experiences from the time 30 minutes ago. Thus, the questionnaire involved more questions than the 14-items that are ultimately used to specify the measurement model (Chapter 3 and Chapter 4). These other items fundamentally function to increase the respondent's ability to answer the vital 14-items. Additionally, the questionnaire was designed to be adaptive via Qualtrics' "Survey Flow" options, such that the questions a participant received depended upon their previous answers. For instance, if a participant stated that they did not engage in multiple media use, they

were not asked if they engaged in multiple media use that shared a goal. This design was chosen to reduce the cognitive load induced by the questionnaire. The need to reduce cognitive load and increase narrative were demonstrated in previous investigations of the MMTIQ (see Chapter 3).

For the 14 items that ultimately are statistically modeled within MMTIQ, participants were then asked questions about whether they engaged in multiple media or devices simultaneously, switched between applications, windows or tabs, or used media in the background. Each of these questions was first asked in a binary fashion (1) yes or (0) no. If participants reported engaging in these behaviors, they were then asked a follow-up question about how long they engaged in these behaviors based on the pilot in Chapter 3. Participants rated the frequency of this behavior in the last 30 minutes as either (1) a little of the time or about 5-10 minutes, (2) some of the time about 15 minutes, or (3) most of the time (20 or more minutes). Following these questions, participants were asked whether these media were used to complete challenging tasks, whether their primary goal was difficult, and whether they could complete their media tasks without even thinking. They were also asked about whether these media helped them complete a larger goal, made completing other tasks more enjoyable, and more specifically if the media they switched between or used simultaneously shared a goal. Finally, participants were asked whether they used media to take breaks and whether media served as unwanted distractions. Each of these questions was similarly first posed

as a binary (1) yes or (0) no, and then provided ordinal response options for rating their frequency within the 30-minute period of questions.

Post survey.

Perceived general stress. The revised 5-item version of the 10-item scale by Cohen et al. (1983) was used for perceived general stress ($\alpha = .89$). These items are rated on a 5-point Likert-scale from (0) never, (1) almost never, (2) sometimes, (3) fairly often, (4) very often. Items include statements such as, "In the last week, how often have you felt nervous and "stressed"?"

Internet media multitasking. The five-item scale developed by Reinecke et al. (2017) was used to measure media multitasking behaviors over the week ($\alpha = .87$). This questionnaire asks people to assess how frequently they engaged in the most common contexts for media multitasking (Jeong & Fishbein, 2017; Shih, 2013). These contexts are the use of Internet concurrently with a) other media, b) having conversations with people, c) eating a meal with another person, d) interacting with a romantic partner, and e) going out with friends.

CHAPTER SIX

DATA ANALYSIS PLAN

This chapter provides an overview and justification for the analysis of the data obtained within this dissertation. Thus, Chapter 6 explicates the relationship between psychometric and statistical choices and hypotheses identified in Chapter 4 and then the analyses used to test the hypotheses and research questions in Chapter 5. It begins by describing and justifying the specification of the measurement model to the MMTIQ, the specifications and evaluation of measurement model hypotheses proposed in Chapter 3 and the standards for the measurement model evaluation. It then transitions into describing and justifying the multi-level model and its use to evaluate the hypotheses between media multitasking intensity and executive functioning, self-regulation, stress, and development.

Measurement Model Specification

The previous iterations of the media multitasking questionnaire detailed in Chapter 3 led to a moderate security in a set of items as valid observations of media multitasking intensity and a few important revisions to the formatting and items. Therefore, in this dissertation, the first step was to formally specify a measurement model of the media multitasking intensity questionnaire that captures the relationship between the indicators (i.e., items) and the target attributes (i.e., construct) as defined in Chapter 2 and 3. In continuing to assess the best measurement model for the MMTIQ,

the current investigation involved evaluating several measurement models within the *item response theory* framework.

Item Response Theory

Item response theory (IRT) models are a set of psychometric models used to specify latent variables through a probabilistic relationship between items and responses – and thus, mathematically relates items and people to the target attribute. Latent models of measurement are those that conceptualize key concepts that are not directly observable (i.e., manifest), but can be measured (with some error) via the observations of other attributes or qualities (see Chapter 2 for further conceptualization). Models can include these observations as outcomes that are caused by latent constructs or as fundamentally causing the construct (Markus & Borsboom, 2013). It is contended that IRT models are the closest statistical complements to the theories of fundamental measurement proposed in the classical theory of measurement (described in detail in Chapter 2) because they examine conjoint measurement and invariance (see Bond & Fox, 2007, for more information). IRT models also allow scholars to secure some evidence of the objectivity and intersubjectivity of a measure (i.e., the quality of measure) as described by Maul et al. (2019).

Psychometrically, IRT is a framework in which expected response probabilities are used to develop and evaluate each of the observations (e.g., items or questions) used within a questionnaire or instrument as well as the questionnaire or instrument as a

whole. IRT reveals item and person characteristics by, "conjointly linking item parameters and latent trait values on the same measurement scale" (Sulis & Toland, 2017). Stated differently, items are designed to target information on the continuum of the latent variable of interest, and are estimated in relation to one another via individuals' response patterns and response probabilities (Bond & Fox, 2007). Item and person scores therefore are estimated on the same scale. Consequently, the variation in the observations of the property of an object and the variation in the values of the property derived from the measure become comparable. This improves the ability to assess hypotheses about the relationship between the variation in the attribute and the variation in properties of an object expressing this attribute (refer to Chapter 2). These models are informative for choosing items, and examining the reliability of the items and person-level estimates of the questionnaire across the continuum of the attribute (i.e., latent attribute scores).

IRT models were chosen for the current attribute of media multitasking intensity and its four dimensions because they allow evaluations of the items in the questionnaire as they provide information about *intensity* and about the current instance of *media multitasking (i.e., the specific person-context interaction).* Furthermore, it was hypothesized that a relationship existed between categorically observed variation of media multitasking *intensity* (i.e., the epistemic nature of the measure) and the continuous ontology of *intensity.* Thus, the nature of the attribute intensity and the

nature of the tool of observation (MMTIQ), which is an ecological momentary

assessment (i.e., a self-report assessment that requires in-the-moment evaluation of

one's experiences) with categorical response options, justified the use of IRT models.

Specifying Dichotomous and Polytomous Models

All 14 items in the MMTIQ were measured via both dichotomous (i.e., binary)

and polytomous (i.e., Likert-scales or multiple choices) item responses, as explained in

Chapter 5. For multiple attributes, participants must first identify whether a certain

behavior or attribute was present (dichotomous) and if so, they were then asked about

the frequency of length of time it was present (polytomous). IRT models' parameters

allow scholars to assess whether polytomous response options provide novel and

valuable information (i.e., over and above dichotomous measures, in this case) based on

the frequency of choosing response options in relation to people's scores on the latent-

attribute.

Rasch model. The Rasch model is the simplest and often considered the purest

form of the IRT model, which involves dichotomous observations or items with binary

response options. The Rasch model is considered to best complement and evaluate the

principles of fundamental measurement. The Rasch model involves modeling the

relationship between dichotomous items and a latent attribute via a logistic (S-shaped)

function. In the Rasch context, the logistic function reflects the probability of a positive

(e.g., yes or correct) response on an item given a person's level or score on a latent

attribute. This curve is called an item response function (IRF), which is also referred to as the item characteristic curve (ICC). The Rasch model involves necessary properties of specific objectivity and additivity.

Within the Rasch model and other IRT models, each item has two parameters estimated: item discrimination and item location (commonly known as item difficulty). Item discrimination (*a*) refers to the steepness of the curve of the IRF, signaling an item's power to differentiate across individuals at different levels of the latent trait. Item location (*b*) (on the latent trait) captures the minimal latent attribute level that corresponds to having a 50% chance of answering positively (i.e., yes). In Rasch models, item discrimination (*a*) is constrained to be equal across items. The item parameters (*a* and *b*) are modelled as fixed effects. The person parameter (θ) which identifies location on the latent attribute can be specified as either a fixed or random effect.

Partial credit model. There are multiple model options within the IRT framework for polytomous items; however, the partial credit model (PCM) is proposed to be the most suitable for multiple response options such as those on a Likert-scale. In PCM, each response option becomes another "step" in the value of the latent attribute, and therefore these options (e.g., "somewhat disagree and disagree") have parameters known as step difficulty parameters (Masters, 1982). When the probability of choosing each response option category is estimated separately and not cumulatively, it is known

181

as an adjacent categories IRT model with item-step parameters. These probabilities are still estimated as a logistic function of one person parameter and one or more item parameters. Each response option (i.e., a little of the time, some of the time, most of the time) or category is given its own threshold at the level of the latent attribute at which its probability of being selected exceeds that of the previous option. In the simplest form of the model, known as the one parameter (or 1-PL) partial credit model, thresholds of discrimination (*a*) are constrained across all items in the questionnaire, but location or difficulty is allowed to vary and is estimated. There are more complicated models such as 2-PL and 3-PL models, in which both discrimination and location are estimated, that are beyond the scope of this dissertation..

Model Dimensionality

IRT models generally assume unidimensionality, but there are ways to test whether measures provide meaningfully distinct information for more than one attribute (dimension). IRT models can be constrained and used to evaluate construct dimensionality through which multiple person parameters (θ) can be estimated from a single model. Therefore, each person parameter (θ) involves information from the entire test such that multiple dimensions involve unique information that are interpretable in context of each other.

The 14-item MMTIQ was conceptualized to have four subdimensions, which may be collapsed into either two or three higher-order dimensions. Therefore, five

models were specified and compared: one unidimensional model, two two-dimensional models, a three-dimensional model, and a four-dimensional model. As dimensions are estimated, the items' discrimination become constrained to be equal within their specified dimension. The model-data fit or goodness of fit metrics were evaluated to identify whether the data supports or provides evidence for each of the models of media multitasking intensity, and then compare these models to identify if one provides the best explanation. See Figures 5 (two-dimensional), 13 (three-dimensional), as well as 10 (four-dimensional) to examine the hypothesized relationships between dimensions and items.
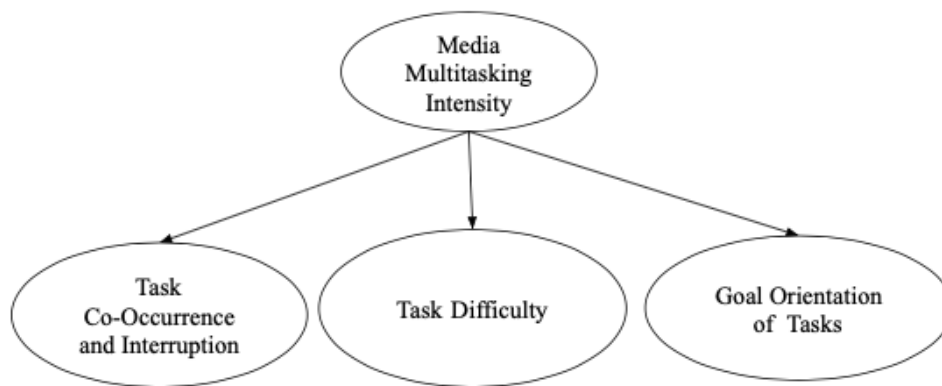
Figure 13. Media multitasking intensity 3-dimensional model in which items of relevancy and intentionality were combined.

These models were first estimated separately for cohorts of parents and adolescents for each time at which the MMTIQ was completed (i.e., at least three of

Tuesday, Wednesday, Thursday, and Friday). After estimating the models for parents

and adolescents separately, and between each time point, the models were then also

estimated for parents and adolescents across time points, and finally they were

estimated for the entire sample across time. These multiple tests evaluate whether the

four dimensional model fits well/best for parents and teens and for various periods of

time. The unidimensional, two dimensional, three dimensional, and four dimensional

models were specified for both dichotomous and polytomous item response options, for

a total of 64 models.

Model Evaluation

   Item and person fit statistics. Within the IRT framework, items can be

evaluated by examining patterns of responses to estimate the probability and degree of

misfit (Bond & Fox, 2007). Items that misfit differ from the expected pattern of item

location and *theta* or (θ) (person estimates on latent attributes). People who misfit

diverge from the expectations of the IRT model. Item and person fit statistics are

identified by squaring the standardized residuals from the mean response pattern. The

statistic involves the mean over respondents to summarize item fit, and the mean over

items to summarize person fit. There are two types of fit statistics: *infit* and *outfit*. Infit

statistics are information-weighted; they assess squared residuals from the means based

on modeled expectations. Infit is more sensitive to responses that are irregularly on-

target for expected response patterns. Infit statistics can be both standardized onto a *t* or

Z distribution or unstandardized. Outfit statistics are the unweighted squared residuals, which remain more sensitive to irregular outlying or off-target response patterns that are expressed both in a standardized (on a *t* or *Z* distribution) or unstandardized form. In other words, outfit refers to being "outlier sensitive fit," and infit refers to "information weighted" by the variance in the item responses. These statistics are used to identify whether items provide appropriate information about the target attribute and individuals in the sample.

Mean-square statistics (MNSQ) are expected to have a value of 1. Standardized values greater than 2 in absolute value indicate items with statistically significant misfit. The smaller the sample size, the harder it is to detect misfit. Wu and Adams (2013) recommend accounting for sample size and number of items (Wang & Chen, 2005) in the questionnaire using an asymptotic formula (i.e., the square root of two divided by sample size). For the current study with 14 items, the cut-off scores of unweighted fit mean-square begin at one plus or minus 0.14. Thus, the on-target interval of unweighted mean-square residual values will range between .86 and 1.14 (i.e., these "fit"). Mean-square values below .86 represent overfit to the model such that there is less noise in the data or that people's performances are better predicted than the model expected. Overfit does not contribute to a better measure, but does not denigrate a measure either. Underfitted items, or those with a mean-square residual value above 1.14, are more

erratic than expected by the model or there is more noise in the data than predicted. Underfit can degrade measurement.

Reliability. The IRT framework includes reliability estimates for both people and items. The person reliability index is the estimate of the replicability of a person's placement that can be expected if a particular sample of persons were to be given another set of suitable items measuring the same construct. The item reliability index is the estimate of the replicability of an item's placement on the latent attribute that can be expected if this particular set of items was provided to another sample of comparable ability (Bond & Fox, 2007). Similar to Cronbach's alpha, the estimates of either form of reliability are recognized as acceptable if they are above .70. There are two statistics of person-level reliability: the weighted likelihood estimation of ability (WLE) and expected a posteriori (EAP) measures.

Model fit statistics. Model-data fit is evaluated via reduced values for the Akaike information criterion (AIC), which estimates the relative quality of a statistical model, Bayesian information criterion (BIC), and the change in -2 log likelihood (-2LL) of the models from two hierarchically nested models (Bond & Fox, 2007; Toland, 2014).

Multigroup Comparison

Because parents and adolescents are two distinct populations whose responses and latent trait values will be compared on the MMTIQ, it is necessary to examine

measurement invariance across the two groups. Measurement invariance is the attribute of measures from which "the outcomes of a measurement procedure are causally determined solely by the measured attribute" (Maul, 2017). Statistically, measurement invariance is examined by evaluating if there is differential item functioning (DIF) for different groups. In other words, if a questionnaire or instrument were to have the property of measurement invariance, the parameters of the item response function would be equal across groups. Media multitasking intensity behaviors are theorized to differ between parents and adolescents, but their responses to the questionnaire may vary also due to their experience with media and reading comprehension capacities.

After unidimensional and multidimensional models were estimated and compared for model fit within parents and adolescents separately, as well as within each time point, the measurement model that demonstrated the best values on the goodness of fit indices was investigated for measurement invariance. In order to investigate measurement invariance as a property of the MMTIQ between parents and adolescents, a multigroup comparison was conducted. Parents and adolescents were specified as being derived from two different populations; two sets of item difficulty parameters were identified; and then these parameters were evaluated for DIF. Finally, the deviance statistics across the constrained and full models were evaluated. A full-information maximum-likelihood (FIML) multiple group analysis was conducted using the package of Multidimensional Item Response Theory (Mirt Version 1.30; Chalmers, 2019) within

the open-source software *R*. Invariance was established by estimating a series of models that sequentially increased the number of model constraints between the groups and then were compared across goodness of fit statistics as described above (Bond & Fox, 2007; Thissen, Steinberg, & Wainer, 1993).

Model Constraints

A series of four models was evaluated for appropriate model fit to establish measurement invariance. First, the model was run separately for the two groups, allowing the locations of items to differ across parents and teen. Second, "weak factorial invariance" or metric invariance was tested by restricting the slopes or the location of items on the latent attribute across parent and teen responses. This establishes that these latent constructs have the same relationship to the items across time within the groupings of parent and teen. Third, scalar invariance was first evaluated by restricting the intercepts of parents and teens on the latent attribute and then by restricting the mean values on the latent trait between parents and teens. Significant differences in the constrained models due to this restriction could potentially identify differential item functioning (DIF). Finally, the fully-constrained model involves constraining the variance of values on the latent trait between parents and adolescents.

Longitudinal Data Analysis: Specifying a Multilevel Measurement Model

The longitudinal data, which involves the repeated assessment (across 3 or 4 time periods) of media multitasking intensity within parents and teens, was collected in order make claims on the individual level (i.e., make claims about differences between individuals). By observing the latent media multitasking intensity that parents and teens were experiencing at the random time on the three or four weeknights, it becomes possible to make claims about individuals' tendency to engage in more intense media multitasking (see Figure 14). Multilevel item response theory models (MLIRT) were used to assess an individual's attribute of intense media multitasking by nesting these time period observations within a person. Stated differently, the random effects of person and time were both estimated, with the hypothesis that the individual had a greater random effect or explained more of the residual variance than the day on which the questionnaire was completed (i.e., more trait-like than state-like). The purpose of using a multi-level model was to be able to make inferences about individual's media multitasking attributes, including whether they were more stable or time and context dependent.

Figure 14. Visualization of the longitudinal and thus multi-level latent trait of media multitasking intensity.

Multilevel Modeling Requirements

There are two conditions that must be met in order to justify specifying random effects and a multilevel model. First, there must be nested units. This condition is met merely in that observations or responses to items are nested within individuals at each time (parent or adolescent) and across time (the ESM periods). Second, the clusters in which units are nested must create bias in standard error estimates such that there is significant shared variance across the groups (i.e., individuals). If so, when the higher-level unit is accounted for, the model then estimates a diminished or greater amount of variance in responses. The Intraclass Correlation Coefficient (ICC) estimates the amount of relevant variation explained by the higher-level unit. Particularly, the ICC is a statistic that estimates the proportion of total variance that is shared among units in the same cluster. As the ICC value increases, this indicates that variation in latent trait

190

scores is more affected by variance between clusters (e.g., in this case within individuals) than by differences within clusters (i.e., amongst observations or days on which the MMTIQ was administered).

Sulis and Toland (2017) recommend that if the ICC is significantly different from zero and greater than 5%, then it is valuable to compare nested single-level and multi-level (2-level) models via the Likelihood Ratio Test (LRT; Chi-Square) and proceed if the difference is significant. The ICCs for observing the MMTIQ responses as nested within individuals, dyads, days, and cohorts of teens or parents were evaluated through the Supplementary Item Response Theory Models (sirt Version 3.4-64; Robitzsch & Robitzsch, 2019) package in *R* to determine which if any of these clusters were justified to include when estimating the (individual-level) random effects.

Specifying Multilevel Models: Random and Fixed Effects

If in evaluating the ICC, there seemed to be enough evidence to suggest that it is necessary to specify a multilevel item response model, then the Multidimensional Item Response Theory (Mirt Version 1.30; Chalmers, 2019) package in *R* was used to specify and evaluate the random effects of individuals and time. A multilevel IRT model allows specification of random and fixed effects on item responses at each observation within each individual as well as random and fixed effects on latent attribute levels across observations and individuals. First, a baseline model with only a fixed intercept was estimated. Second, a multi-level model was specified with items as

191

fixed effects and individuals as random effects. This model investigates the explanatory value of items nested within individuals. Third, an additional random effect or random intercept of time was specified at the latent level. Fourth, latent fixed effects of the covariate teen (i.e., whether the individual was an adolescent or parent's executive functioning, self-regulation, the combination of executive functioning and self-regulation, the interaction between teen and executive functioning and self-regulation, stress within the month before beginning the study, stress during the week of the study, and self-report internet media multitasking were investigated separately. All models were compared via fit statistics to investigate the value of adding each fixed and random effect. Fixed effects were evaluated via the degree to which their addition to the model improved the fit statistics significantly from baseline and previous models that excluded it.

CHAPTER SEVEN

RESULTS

In this chapter of the dissertation, the results of examining the psychometric and structural statistical models described in Chapter 6 in relation to each hypothesis will be discussed. Before reviewing the results from examining the hypotheses from Chapter 4, the descriptive statistics of the covariates later included in the model are reviewed (see Table 17 and 18). These are the covariates for the parents and teens who completed the study across the whole week (pre-survey, at least 3 of the 4 ESM, and the post-survey). Thus, the MLIRT analyses which involved a covariate included a smaller subsample of 324 parent-adolescent dyads. The polytomous and dichomtous, unidimensional and multidimensional IRT analysis described in Chapter 6 included the full 648 dyads. Item descriptive statistics including item difficulty (i.e., the value of media multitasking intensity they help observe), discrimination, and standard deviation can be found under the results for Hypothesis 2 and within Table 28.

Table 17

Covariate Correlation Matrix

		Executive Functioning (EF)	Self-Regulation (SR)	EFSR	General Stress Before	General Stress After	Internet MMT
EF	Pearson's r	—					
	p-value	—					
	Upper 95% CI	—					
	Lower 95% CI	—					
SR	Pearson's r	0.37	—				
	p-value	< .001	—				
	Upper 95% CI	0.44	—				
	Lower 95% CI	0.30	—				
EFSR	Pearson's r	0.85	0.80	—			
	p-value	< .001	< .001	—			
	Upper 95% CI	0.87	0.83	—			
	Lower 95% CI	0.83	0.77	—			
Stress Before	Pearson's r	0.55	0.29	0.51	—		
	p-value	< .001	< .001	< .001	—		
	Upper 95% CI	0.60	0.36	0.57	—		
	Lower 95% CI	0.49	0.21	0.45	—		
Stress After	Pearson's r	0.44	0.31	0.46	0.60	—	
	p-value	< .001	< .001	< .001	< .001	—	
	Upper 95% CI	0.50	0.38	0.52	0.65	—	
	Lower 95% CI	0.37	0.23	0.39	0.54	—	
Internet MMT	Pearson's r	0.22	0.21	0.26	0.07	0.18	—
	p-value	< .001	< .001	< .001	0.080	< .001	—
	Upper 95% CI	0.29	0.29	0.33	0.15	0.25	—
	Lower 95% CI	0.14	0.13	0.18	-0.01	0.10	—

Table 18

Descriptive Statistics for Grand Mean Centered Covariates

	Executive Functioning (EF)	Self-Regulation (SR)	EFSR	General Stress Before Week	Genera l Stress After Week	Internet Media Multitasking
Valid	577	577	577	577	577	577
Missing	0	0	0	0	0	0
Mean	0.01692	0.09614	0.05403	-0.00245	0.009974	0.02070
Std. Deviation	0.7891	0.6971	0.6152	0.8557	0.8246	0.7769
Minimum	-1.660	-1.695	-1.680	-1.895	-1.685	-1.215
Maximum	2.340	1.767	1.676	2.105	2.315	1.785

Hypothesis 1: Dimensionality

The model fit statistics for the unidimensional, two-dimensional, three-dimensional, and four-dimensional models were estimated for teens and parents separately, and then within subgroups of teens and parents separated and estimated by the day of the week in which they were taken within the study. Within these sub-

samples, and in the overall sample, the four-dimensional model demonstrated the best

model fit as demonstrated by a lower loglikelihood, AIC, BIC, and Deviance (G2), as

compared to the three-dimensional, two-dimensional and unidimensional models. Thus,

the psychometric assessment or measurement model of media multitasking intensity

evidenced some theoretical security for the objectivity of the four dimensions identified.

Thus, the following data analyses will continue to model media multitasking *intensity* as

involving the dimensions of *sensory allocation to media, media task difficulty,*

intentional allocation to media, and *media task relevance.* See Tables (19-26 below).

Table 19

Model Fit Indices for Adolescents at Time 3

	Log Likelihood	AIC	BIC	Deviance	Parameters
Undimensional Rasch	-3359.78	6754.00	6822.00	6719.57	17
Two Dimensional Model	-3355.29	6752.59	6836.72	2378.15	21
Three Dimensional Model	-3278.11	6606.23	6706.39	2223.73	25
Four Dimensional Model	-3256.74	6573.49	6693.68	2180.98	30
Unidimensional Partial Credit Scale Model	-5019.54	10149.09	10369.43	5560.32	55
Two Dimensional Model	-5018.57	10153.13	10385.50	5558.33	58
Three Dimensional Model	-5013.28	10148.56	10392.95	5547.76	61
Four Dimensional Model	-4991.36	10104.72	10349.11	5503.91	61

Note. AIC = Akaike information criterion. BIC = Bayesian information criterion.

Table 20

Model Fit Indices for Teens at Time 4

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2981.01	5996.00	6064.00	5962.02	17
Two Dimensional Model	-2981.16	6004.32	6088.20	2094.15	19
Three Dimensional Model	-2751.36	5548.72	5640.58	1658.58	23
Four Dimensional Model	-2702.01	5460.02	5571.85	1559.85	28
Unidimensional Partial Credit Scale Model	-4266.25	8642.51	8862.18	4422.83	55
Two Dimensional Model	-4293.44	8702.88	8934.53	4477.17	58
Three Dimensional Model	-4256.25	8634.49	8878.12	4402.77	61
Four Dimensional Model	-4236.13	8594.26	8837.89	4362.55	61

Note. AIC = Akaike information criterion. BIC = Bayesian information criterion.

Table 21

Model Fit Indices for Teens at Time 5

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2830.57	5693	5757	5661.15	15
Two Dimensional Model	-2683.542	5405.085	5480.635	1589.02	19
Three Dimensional Model	-2654.165	5354.329	5445.785	1536.67	23
Four Dimensional Model	-2597.132	5250.264	5361.601	1420	28
Unidimensional Partial Credit Scale Model	-4206.313	8522.626	8741.325	4366.17,	55
Two Dimensional Model	-4205.874	8527.747	8758.376	4363.45,	58
Three Dimensional Model	-4188.247	8498.494	8741.052	4329.85	61
Four Dimensional Model	-4172.502	8467.005	8709.562	4294.83	61

Note. AIC = Akaike information criterion. BIC = Bayesian information criterion.

Table 22

Model Fit Indices for Teens at Time 6

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2691.35	5413.00	5472.00	5382.69	15
Two Dimensional Model	-2687.27	5412.54	5488.19	1640.56	19
Three Dimensional Model	-2652.88	5351.77	5443.34	1571.67	23
Four Dimensional Model	-2616.48	5288.95	5400.43	1498.87	28
Unidimensional Partial Credit Scale Model	-4163.98	8437.96	8656.94	4341.60	55
Two Dimensional Model	-4157.06	8430.12	8661.04	4327.71	58
Three Dimensional Model	-4149.22	8420.43	8663.30	4312.02	61
Four Dimensional Model	-4149.81	8421.62	8664.49	4313.21	61

Note. AIC = Akaike information criterion. BIC = Bayesian information criterion.

Table 23

Model Fit Indices for Parents at Time 3

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2326.77	4684.00	4740.00	4653.55	15
Two Dimensional Model	-2323.08	4684.16	4755.58	1482.05	19
Three Dimensional Model	-2300.76	4647.51	4733.96	1437.33	23
Four Dimensional Model	-2275.62	4607.25	4712.49	1387.07	28
Unidimensional Partial Credit Scale Model	-3669.16	7448.32	7655.06	3994.36	55
Two Dimensional Model	-2306.29	4650.59	4722.01	1448.47	58
Three Dimensional Model	-3659.59	7441.18	7670.48	3975.21	61
Four Dimensional Model	-3651.22	7424.44	7653.74	3958.47	61

Table 24

Model Fit Indices for Parents at Time 4

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2755.02	5540.00	5601.00	5510.04	15
Two Dimensional Model	-2751.52	5541.03	5617.79	1554.61	19
Three Dimensional Model	-2725.49	5496.98	5589.90	1502.50	23
Four Dimensional Model	-2672.90	5401.80	5514.93	1397.30	28
Unidimensional Partial Credit Scale Model	-4068.84	8247.68	8469.89	3833.79	55
Two Dimensional Model	-4069.99	8255.97	8490.31	3836.06	58
Three Dimensional Model	-4085.97	8293.93	8540.39	3868.01	61
Four Dimensional Model	-4054.38	8230.77	8477.22	3804.84	61

Table 25

Model Fit Indices for Parents at Time 5

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2719.64	5469.00	5530.00	5439.28	15
Two Dimensional Model	-2714.09	5466.19	5543.05	1616.27	19
Three Dimensional Model	-2682.94	5411.89	5504.92	1553.96	23
Four Dimensional Model	-2642.60	5341.21	5454.47	1473.25	28
Unidimensional Partial Credit Scale Model	-4060.84	8231.69	8454.16	4049.19	55
Two Dimensional Model	-4065.34	8246.68	8481.29	4058.15	58
Three Dimensional Model	-4040.18	8202.35	8449.10	4007.82	61
Four Dimensional Model	-4024.07	8170.14	8416.88	3975.59	61

Table 26

Model Fit Indices for Parents at Time 6

	Log Likelihood	AIC	BIC	Deviance	Parameters
Unidimensional Rasch	-2610.75	5252	5312	5221.51	15
Two Dimensional Model	-2595.112	5228.225	5304.762	1458.86	19
Three Dimensional Model	-2582.129	5210.257	5303.018	1424	23
Four Dimensional Model	-2539.794	5135.588	5248.379	1348.22	28
Unidimensional Partial Credit Scale Model	-3942.697	7995.394	8217.214	3929.86	55
Two Dimensional Model	-3953.677	8023.354	8257.273	3951.96	58
Three Dimensional Model	-3935.706	7993.411	8239.429	3893.99	61
Four Dimensional Model	-3923.218	7968.436	8214.455	3867.68	61

Research Question 1: Dichotomous or Polytomous

In identifying whether the dichotomous or polytomous models were more appropriate for the data, the category response curves from the Partial Credit Model were examined. The purpose was to evaluate whether each category was identified as the most probable choice for some value of the latent attribute media multitasking *intensity*. In examining both the test information curves and the category response curves, it appeared that the response options "a little of the time" and "some of the time" did not have a point on the latent attribute of intensity at which they were most probable. As such, it appeared that the partial credit model, which included these response options, did not provide more information because the options "never" or "most of the time" were providing the most information. Therefore, this sample of parents' and teens' responses over one week supported findings from Study 4 mentioned in Chapter 3. That is, the dichotomous measure was a better fit.

Measurement invariance was examined between parents and adolescents for the four dimensional dichotomous model across their answers throughout the week (see Table 27). There was no significant difference (χ_2 (14) = 18, p = .21) between the configural model, which estimates the model within both subgroups, and the metric model, which constrains variance in slopes. While the model demonstrated metric invariance such that the factor loadings of items remained the same between parents and teens, there was a significant difference in model fit once the intercepts were fixed or

constrained in both subsamples (χ_2 (20) = 444.29, $p < .001$). Only weak or metric

measurement invariance was demonstrated between parents and teens.

Table 27

Multiple Group Comparisons, Examining Measurement Invariance Amonst Parents and Teens

	Log Likelihood	AIC	BIC	Chi-Square Change	Df	p value
Configural Model	-20958.40	42084.80	42594.05			
Metric Model (additional contraint of slopes)	-20967.40	42074.80	42499.17	18.00	14	0.207
Scalar Model 1 (additional contraint of intercepts)	-21191.04	42482.09	42785.21	444.29	20	<.001
Scalar Model 2 (additional contraint of means)	-21191.04	42482.09	42785.21	0.00	0	1
Full Constrained Model (additional contraint of variance)	-20993.25	42100.54	42437.99	395.60	6	<.001

Hypothesis 2: Individual Differences

The second hypothesis predicted that individual differences (i.e., there would be

a significant amount of shared variance across observations within an individual) would

explain media multitasking intensity across the week of the study. This also was used to

provide evidence for the need to create a multilevel item response theory (MLIRT)

model. In order to evaluate this hypothesis, it was necessary to examine if the ICCs, or

intra class coefficients between items over the week as nested within individuals, was

greater than .05. In the current sample, the ICC for individuals was .48, which suggests

48% of the variance was explained between individuals rather than within individuals

(See Table 34). Thus, hypothesis two was supported. In additional analyses, the

individual ICC (48%) was larger than that of the dyad (.42), time (.08), or the grouping

dummy variable of parent and adolescent (.14). This means that that individuals shared

more variance within themselves in that an individual engaged in more similar media

multitasking over time than between individuals variance; individuals media

multitasking intensity is more stable.

Table 28

Multilevel 2PL model

	parameter	Mean	SD	MAP	Rhat	effSize	Q5	Q95
1	b[1]	0.41	0.037	0.41	1.09	27.4	0.35	0.47
2	b[2]	0.99	0.06	0.96	3.97	11.2	0.90	1.09
3	b[3]	1.92	0.09	1.89	3.64	5.3	1.78	2.07
4	b[4]	1.34	0.05	1.33	1.57	13.4	1.24	1.42
5	b[5]	0.85	0.04	0.86	1.26	13.7	0.80	0.91
6	b[6]	0.33	0.03	0.34	1.01	130.3	0.29	0.38
7	b[7]	-0.23	0.02	-0.23	1.08	155.9	-0.27	-0.19
8	b[8]	1.48	0.05	1.49	1.38	16.6	1.40	1.57
9	b[9]	1.60	0.11	1.53	1.28	4.2	1.45	1.87
10	b[10]	2.09	0.10	2.12	3.3	7.6	1.92	2.23
11	b[11]	1.18	0.05	1.20	1.7	18.1	1.09	1.25
12	b[12]	0.29	0.03	0.30	1.05	132.8	0.25	0.34
13	b[13]	0.56	0.03	0.56	1.05	99.3	0.52	0.62
14	b[14]	3.34	0.19	3.48	5.9	2.5	3.02	3.59
15	a[1]	1.02	0.05	1.01	1.23	11.8	0.93	1.10
16	a[2]	1.60	0.10	1.53	6.2	4.1	1.46	1.78
17	a[3]	1.89	0.10	1.85	4.4	4.8	1.72	2.06
18	a[4]	0.96	0.05	0.97	1.58	12.7	0.87	1.04
19	a[5]	0.46	0.04	0.45	1.65	33	0.39	0.52
20	a[6]	0.36	0.03	0.35	1.13	60	0.31	0.41
21	a[7]	0.28	0.03	0.28	1.15	48.2	0.24	0.34
22	a[8]	0.52	0.04	0.53	1.03	42	0.46	0.58
23	a[9]	1.02	0.07	1.01	1.3	6.6	0.93	1.16
24	a[10]	1.51	0.08	1.56	2.79	7.3	1.36	1.63
25	a[11]	0.72	0.04	0.69	1.34	25.3	0.65	0.79
26	a[12]	0.82	0.03	0.81	1.1	31.1	0.76	0.87
27	a[13]	0.79	0.04	0.79	1.14	31.4	0.73	0.85
28	a[14]	2.06	0.14	2.18	7.37	3.5	1.82	2.22
29	sigma1	0.91	0.02	0.91	1.21	34.1	0.87	0.94
30	sigma2 (level	0.87	0.03	0.87	1.15	54.4	0.82	0.92
31	ICC	0.48	0.02	0.49	1.28	58.4	0.45	0.52
32	mu.b	1.16	0.26	1.16	1	500	0.71	1.56
33	omega.b	0.97	0.20	0.88	1.01	378.6	0.69	1.33
34	sigma.b[1]	0.49	0.06	0.48	1.45	6.7	0.39	0.59
35	sigma.b[2]	0.55	0.07	0.49	2.98	7.7	0.46	0.67
36	sigma.b[3]	0.52	0.07	0.54	1.78	9.5	0.40	0.62
37	sigma.b[4]	0.53	0.08	0.48	1.71	6.3	0.43	0.67
38	sigma.b[5]	0.53	0.05	0.52	1.15	20.7	0.46	0.61
39	sigma.b[6]	0.52	0.05	0.52	1.1	13.4	0.45	0.60
40	sigma.b[7]	0.67	0.05	0.66	1.11	23	0.60	0.76
41	sigma.b[8]	0.58	0.08	0.63	2.32	5.8	0.43	0.69
42	sigma.b[9]	0.86	0.12	0.80	1.29	5.2	0.70	1.12
43	sigma.b[10]	0.63	0.07	0.62	1.63	9.4	0.50	0.75
44	sigma.b[11]	0.53	0.11	0.52	3.17	2.7	0.31	0.66
45	sigma.b[12]	0.53	0.04	0.52	1.22	27.6	0.48	0.59
46	sigma.b[13]	0.49	0.06	0.44	1.31	10.5	0.41	0.58
47	sigma.b[14]	0.36	0.06	0.36	1.04	6	0.28	0.47
48	sigma.a	0.33	0.02	0.32	3.86	5.7	0.31	0.36
49	omega.a	0.561	0.109	0.51	1.02	346.5	0.41	0.76

Hypothesis 3: Individual Differences in Executive Functioning

It was predicted that media multitasking intensity would have a positive relationship with individual differences in executive functioning. Particularly, deficiencies of executive functioning would predict more sensory resource allocation to media, and more difficulty of media tasks, but less intentional media and task relevance. After including the grand-mean centered scores for executive functioning as a predictor, which would have a fixed latent effect on the four dimensions of media multitasking intensity, this hypothesis was *not* supported. Importantly, including executive functioning in the model after controlling for dyad did not improve the overall model fit from the baseline significantly according the -2LL loglikelihood ratio test (LRT), which is calculated via chi-square (χ_2 (4) = -53.32, p = 1). However, it significantly improved the model once dyad ID was included (χ_2 (4) = 51.70, $p < .001$).

Hypothesis 3a. Poorer executive functioning predicted more sensory resource allocation to media across the week (β = .20). Similarly, the lower one's executive functioning, the more difficult one reported their tasks to be (β = .35). Thus, hypothesis 3a was supported.

Hypothesis 3b. Poorer executive functioning demonstrated a positive relationship with intentional allocation of sensory resources (β = .28), as well as a positive relationship with relevancy of tasks in which participants reported to engage (β = .16). Thus, hypothesis 3b was not supported.

In addition to the lack of significant information added to the model, the expected relationships between executive functioning and dimensions were not supported. Though a positive relationship was expected and found between lower executive functioning and sensory resource allocation and task difficulty, executive functioning did not demonstrate the expected negative relationship with intentional allocation and relevancy; indeed, it showed positively relationships. This supports the meta-analysis conducted by Wiradhany and Nieuwenstein (2017), which found that the effect sizes of the relationship between executive functioning and media multitasking frequency, when controlling for sample size, neared zero. Yet, this could support hypotheses that media multitasking is a consequence of poorer executive functioning due to a breadth-bias, in which more tasks are perceived as relevant and potentially even intentional (Ophir et al., 2009).

Hypothesis 4: Individual Differences in Self-Regulation

In hypothesis four, self-regulation, a specific sub-type or use of executive functioning, was predicted to have a significant relationship with the dimensions of media multitasking intensity. Specifically, deficiencies in self-regulation would predict more sensory resource allocation and task difficulty within the week, but less intentional sensory resource allocation, and task relevance. Including self-regulation as a latent fixed effect significantly improved model fit (χ_2 (4) = 18203.42, $p < .001$). However, hypothesis four was only partially supported.

Hypothesis 4a. Deficiencies in self-regulation predicted a small amount of

additional sensory allocation to media (β = .14), and a moderate amount of more

difficult tasks (β = .22). Thus, hypothesis 4a was supported.

Hypothesis 4b. However, deficiencies in self-regulation did not predict less

intentional sensory resource allocation (β = .21) to media or task relevance (β = .18).

Thus, hypothesis 4b was not supported.

The current findings suggest that self-regulation impacts sensory resource

allocation and the evaluation of the difficulty of tasks such that more difficult tasks are

involved and more sensory resource allocation occurs when self-regulation is lower.

However, this does not relate to intentional sensory allocation or the relevance of the

media involved.

Hypothesis 5: The Explanatory Power of Executive Function vs. Self-
Regulation

Hypothesis five involved the prediction that self-regulation, which involves the

use of attentional resources and working memory to monitor, evaluate, and adjust

behavior, would be a better explanatory variable than executive functioning, which is

more general. However, because the constructs are related in nature, it was necessary to

analyze this hypothesis by first creating a construct that combines executive functioning

and self-regulation and comparing this construct to executive functioning and self-

regulation separately.

In examining the explanatory power of the combined predictor, the grand-mean centered average of the executive functioning and self-regulation were entered into the model and first compared to the baseline model, similar to the analyses described for hypotheses four and five. The addition of this combined construct significantly improved the model, as reflected in the change test (χ_2 (4) =75.851, $p < .001$).

Following this model, the explanatory power of the grand-mean centered combined predictor of the two were entered into the model and compared to the model testing the influence of executive functioning model and the self-regulation separately. The combined model would have the executive functioning model nested within it and therefore provided an appropriate comparison. The model with the combined executive functioning and self-regulation latent fixed effects was significantly better than the model that included executive functioning alone as the predictor (χ_2 (0) = 24.15, $p <$ .001). Thus far, this suggests that adding self-regulation into the model was an improvement.

Finally, the examination of hypothesis five resulted in comparing the model with self-regulation (grand-mean centered) to this model with the combined executive functioning and self-regulation predictor. In comparison to the latent fixed effects of the combined forms of attentional functions, self-regulation alone performed worse. Again, the combination of executive functioning and self-regulation as a single, combined

predictor of the dimensions of media multitasking intensity significantly improved the fit of the model than self-regulation alone (χ_2 (0) = 41.24, $p < .001$).

Therefore, hypothesis five was *partially supported.* While it is true that self-regulation is not a better predictor than executive functioning, self-regulation as a component of executive functioning increases explanatory power. This is evaluated as partial support of hypothesis in addition to executive functioning's lack of significant improvement to the baseline. Stated differently, self-regulation is not a better predictor alone, but self-regulation is a valuable predictor in that it is a significant predictor, but is most valuable to the model when it is contextualized as a component of executive functioning.

Hypothesis 6: Adolescents and Media Multitasking Intensity

Hypothesis six predicted that being an adolescent would moderate the relationship between self-regulation and the four dimensions of media multitasking intensity. Adolescents, who have yet to develop their self-regulatory capacities, were expected to report *more* sensory resource allocation to media, *more* difficult media tasks, *less* intentional sensory resource allocation, and *less* task relevancy.

This hypothesis was evaluated by examining the predicted value of adding the variable of teen, the grand centered mean of self-regulation, and finally an interaction between the variable of adolescent and self-regulation. In order to examine this, teens were first introduced into the model in addition to self-regulation to evaluate if

including the variable of teen improved overall significance. The inclusion of teen as a predictor significantly improved model fit (χ_2 (4) = 17.583, $p < .001$). Thus, the effect of the interaction between teen and self-regulation was then examined.

The model including the grand centered mean for self-regulation, the predictor of teen, and the interaction between teen and self-regulation demonstrated a significantly better fit than the baseline model (χ_2 (12) = 69.32, $p < .001$), better than the model with the mean for self-regulation alone (χ_2 (8) = 34.71, $p < .01$), and better than the model with the two separate measures of self-regulation and teen (χ_2 (8) = 17.13, $p < .001$). Thus, adolescence significantly moderates the impact of self-regulation's impact on media multitasking intensity.

In the model, teens engage in less sensory resource allocation to media, less difficult tasks, and less intentional sensory resource allocation than parents. Lesser self-regulation in general still predicted more media multitasking, more difficult tasks, more intention sensory resource allocation, and more task relevancy. However, amongst teens, the lesser the self-regulation the lesser the sensory resource allocation to media tasks, more intentional sensory resource allocation, and yet less relevance amongst tasks in which they engaged, in comparison to parents with less self-regulation capacity. However, there were no significant differences in difficulty of tasks predicted by the interaction between self-regulation and adolescence. Thus, these results find that less

self-regulation for teens impacts them such that they allocate less sensory resources to media use, but do so with more intention and less relevance.

Hypothesis 7: Media Multitasking Intensity and Stress

In hypothesis seven, the relationship between the dimensions of media multitasking *intensity* and stress were investigated in order to examine directionality and the strength of the relationship. Particularly, hypothesis 7a predicted that increased engagement in sensory resource allocation to media and increased engagement in difficult tasks were expected to be positively predictive of stress at the end of the week (i.e., in the post-survey). On the other hand, hypothesis 7b predicted that increased engagement in intentional sensory allocation and task relevancy would be predictive of less stress at the end of the week (post-survey). This hypothesis was first examined by regressing stress at the end of the week on media multitasking intensity throughout the week (i.e., over time using MLIRT). Then, it was tested evaluating the predictive power of stress at the beginning of the week for media multitasking intensity throughout the week. Finally, stress at the beginning of the week was included in the model in addition to self-reported stress at the end of the week, such that the model predicting stress at the end of the study would control for stress at the beginning of the study.

Thus, there are three overall model fits to the data that first assessed the change in stress explained by media multitasking intensity. By adding the grand centered mean stress at the end of the week in which the study occurred, the overall model fit statistics

showed significant improvement (χ_2 (4) = 70.348, $p < .001$). Similarly, after including

the grand-centered mean of stress at the beginning of the study into the model, the

overall model fit again improved significantly (χ_2 (4) = 47.466, $p < .001$). The grand-

centered mean of stress at the beginning of the week/study predicted more sensory

resource allocation, more difficulty, more intentional resource allocation, and more

relevant tasks. Finally, a model was specified with the grand-centered mean of stress at

the beginning of the week and the grand-centered mean of stress at the end of the week.

This model significantly again improved upon the baseline model including dyad (χ_2

(8) = 73.275, $p < .001$), as well as the model involving stress at the beginning of the

study (χ_2 (4) = 25.81, $p < .001$), however, it was not significantly better than the model

involving only stress at the end of the week (χ_2 (4) = 2.93, $p = .57$).

Hypothesis 7a. In the final model, hypothesis 7a received support: a positive

relationship was expected between sensory allocation and task difficulty and stress.

Greater shared sensory resource allocation to media ($\beta = .18$), and more difficult tasks

($\beta = .29$) both predicted increased stress at the end of the week.

Hypothesis 7b. Intentional sensory resource allocation ($\beta = .21$) and greater

task relevance ($\beta = .11$) was associated with more stress at the end of the week. Thus,

hypothesis 7b was not supported.

CHAPTER EIGHT:

DISCUSSION AND FUTURE DIRECTIONS

Discussion

The environment and thus daily experiences of individuals around the globe have become dependent upon technology. This dependency has raised new concerns for the cognitive and emotional well-being of users, especially for adolescents who are the most avid users. While the evidence accumulates that media has become an essential tool and fundamental aspect of daily life, its critical role in a variety of human needs and goals has created more interference and distraction from the most valuable aspects of life: relationships, work, and personal time. Yet, as the concern about and necessity for investigating the role of media dependency and a multiple media environment in daily life has increased, the challenges of investigating media and evaluating its impact have also grown or at least become more apparent. As the forms of media, variety of media content, and access to media increase and continue to proliferate, the nature of media and its effects appear to evolve. Theories explaining the mechanisms of media's effects are challenged by these transformations. Simultaneously, the security of the claims or validity of measures of media have been confronted by the mounting difficulty of accurately representing the features of media enough to explain its effects. *In response to the existing challenges with theory and measurement development, the current dissertation developed and evaluated a theory of media multitasking intensity*

211

through an iterative process of validating a measure and a longitudinal multi-survey study of parent-adolescent dyads. The findings evidence the profound value of intertwining these pillars of empirical results: theory and measurement.

In its most recent iteration, the media multitasking intensity questionnaire (MMTIQ) demonstrated evidence that the revisions to theory, structure, and content secured a more valid measure. The media multitasking intensity questionnaire, the measure validated and iterated upon, psychometrically reflected the theorized four dimensions (See Figure 10). S*ensory resource allocation to media, task difficulty, intentional sensory resource allocation,* and *task relevance* appear to best explain the variation in a random sample of media experiences as self-reported by parents and adolescents on multiple weeknights.

First, the psychometric fit of models of the data revealed that the frequency in which people engaged in behaviors indicative of media multitasking intensity did not provide more valuable information. Rather, it appeared that simply the presence or absence of these features was most informative. When examining the presence of *media multitasking intensity* indicators, the measure demonstrated the same properties (i.e., similar ranges of variation and means) regardless of weeknight. However, the averages of these four dimensions and variation of values within them differed between parents and teens. Despite, the potential caution this necessitates, these findings provide evidence of some degree of intersubjectivity within the measure.

Second, participants' media multitasking intensity was partially explained by individual differences. The current study found that media multitasking intensity demonstrated some stability across the weeknights but also suggested that the media multitasking intensity is indeed partially explained by individual tendency. The variation of responses was greater between individuals than within individuals. However, it is important to note that there was still some significant variation within one person across the time periods. These findings could indicate that media multitasking intensity, at least when measured systematically similar periods of time (i.e., on weeknights, within the most recent half hour), is consistent within a person because of the consistency of the types of contexts they experience. The intricacies of the interaction between individual and context require further investigation before it is possible to make larger claims about the stability, tendency, or habits of individuals to engage in more or less intense media multitasking. It is unclear if the causes of differences in media multitasking intensity are variation in individuals due to their contexts or if it is an outcome of a more internal, trait-like quality.

Third, cognitive abilities such as executive functioning and self-regulation predict engagement in more intense media multitasking, and are most powerful as a single combined predictor. The current findings demonstrate lesser conscious directed attention and working-memory capacity as well as the lesser capacity for self-monitoring, evaluation, and adaptation impact media multitasking intensity. Yet, their

213

impact on intentionality and task relevance was not as predicted. Lower executive functioning and lesser self-regulation functioning both predicted more sensory resource allocation, more difficult tasks, more intentional sensory resource allocation, and more relevancy. However, it was predicted that the lower executive functioning and less self-regulation would predict *less* intentionality and *less* relevancy. Thus, this raises concerns about whether this is explainable by the perception of intentional allocation and relevancy that is problematic or whether the engagement of media multitasking with difficult tasks is unlikely to occur if the tasks are not engaged intentionally or due to perceived relevancy. Yet, the findings gain interpretability and thus secure more validity in relation to previous iterations evaluating the measure and attribute of intensity.

Based on previous investigations (Study 1 and 2 described in Chapter 3), it appears that in difficult conditions, it is more common to allocate sensory resources to media more intentionally and to do so with tasks that are more relevant. This provides support for the "law of less work", limited capacity models, and threaded cognition (Wang et al., 2015; Yeykelis et al., 2014), which suggests that it may not be cognitively plausible or likely for people to engage in the most cognitively taxing or most intense forms of media multitasking theorized in Figure 11. Though it is possible to theorize that some people *may* engage in the most extreme form of media multitasking intensity, this is relatively rare and may require a different study design to examine it. Thus, the

most intense form of media multitasking intensity, engaging in more sensory allocation more intentionally and with more difficult and relevant tasks, should be further investigated. Previous research explicated in Chapter 3 in addition to that of Wang et al. (2015) may provide more clarity, but the security of these claims still warrants further investigation. Nonetheless, deficient executive functioning and less self-regulation separately and in combination both predict more intense media multitasking behaviors.

Fourth, adolescents' media multitasking intensity significantly differs from their parents' and is partially explainable by deficiencies in self-regulation. In comparison to the previous research, which suggests that teens engage in media multitasking more frequently than other age ranges, the current study found that adolescents engaged in less sensory resource allocation amongst media and other tasks. However, they allocated their resources to less difficult tasks, and engaged in their tasks less intentionally. There was no significant difference in the relevance of their tasks. Thus, they engaged in *less intense* media multitasking than their parents. However, teens with poorer self-regulation functioning engaged in less sensory resource allocation to media tasks and engaged in them more intentionally, but with significantly less relevant tasks. The difference between parents' and teens' relevancy of tasks depends on lower self-regulation capacities. It appears lower self-regulation within adolescents is not predictive of sensory allocation to multiple media tasks or media tasks that are interruptive or overlapping with other tasks, but rather of the intentional engagement in

irrelevant media tasks. The current findings support and echo Reinecke et al. (2018)'s findings that increased access to the Internet and thus various mediated content can invite procrastination, rather than active engagement in multiple media or multiple tasks. They similarly found that self-regulation predicts procrastination behaviors and thus increased depression and anxiety. This may suggest that less self-regulation does not explain when *media multitasking intensity* is greatest amongst teens, but explains when it creates opportunities to procrastinate.

Finally, higher media multitasking intensity predicted more stress. The current study found that even when controlling for stress at the beginning of the study, stress was positively caused by to engaging in more sensory resource allocation to media in addition to other tasks, more difficult tasks, more intentionally allocated sensory resources, and more relevant tasks. This relationship with intentionality and relevancy was not expected. Extremely demanding or extremely high media multitasking intensity may be very uncommon. In other words, they are currently suppressed in the existing study. A ceiling effect may be occurring such that the most harmful combination of these dimensions involves the intentional, relevant, difficulty, sensory resource allocation amongst media and/or other tasks. These findings imply that the most stressful forms of media multitasking intensity are those in which sensory resource allocation amongst media and other tasks are done intentionally because they are relevant to difficult goals. This suggests that technostress or the digital stress

216

experienced by parents and teens alike may be an outcome of increased involvement of media, if not multiple media, in completing their goals. In comparison to previous frameworks or explanations of media multitasking intensity, the current findings suggest that it is not most harmful because it is an unintended and irrelevant behavior (i.e., uncontrollable distraction). Rather, these findings suggest that the technologically and media-saturated environment has created new challenges for accomplishing daily goals that currently have the capacity to overwhelm.

Future Directions

The current findings demonstrate that the theory of media multitasking intensity must incorporate the various competing factors that impact the integration of media tasks into daily life as opposed to focusing on only the interference of media tasks with daily life. The dissertation secures evidence of the validity of a four-dimensional model of media multitasking intensity (i.e., co-occurrence and interruption, task difficulty, task intentionality, and task relevancy) that predicts experiences of stress and is explained by executive functioning and self-regulation capacity. Still, the role of intentionality and relevance in contributing to stress and in relation to these explanatory factors warrants further research to examine whether low intentional and low relevance is uncommon or less harmful. Future research should continue to examine the causal mechanisms that predict media multitasking intensity and explain its effect on well-being. In addition, the current investigation of media multitasking limited the observations of sensory resource

allocation to media on only weeknights. This may have impacted the relationships that were investigated: weeknights involve relaxation with family and friends as well as extracurricular activities and homework for teens. Finally, in the vein of continuing to wed theory development with measurement validation, future research should investigate triangulations of methods of observing media multitasking intensity and well-being, including passive-observation, reports from close friends and family, and uses of biological measures. Though weak measurement invariance and thus evidence of intersubjectivity was moderately secured, these future investigations should continue to validate the measure in various theoretically-relevant subpopulations. In combination, these efforts could continue to build security in the evidence of validity of the media multitasking intensity questionnaire (MMTIQ) and the claims that scholars can assert based on their empirical findings.

Conclusion

The current dissertation examined the existing media effects literature and identified the weak role of measurement development and validation in theory development. It thus contended that this explains existing inconsistent findings within the literature (Chapter 2). Thus, it proposed an iterative process of measure validation and theory development, and evidenced its value. By creating measures and theories as reflections of one another or "mirrors" that become sources of feedback for one another, the advancement of knowledge develops in relation to the security of claims and with a

greater degree of cohesiveness. The development of the media multitasking intensity questionnaire and the iterative process of validation became the foundation for a theory of media multitasking intensity (Chapters 3 and 4), which more securely captures complex variation in a sociocognitively sensitive and novel construct. These developments are a consequence of leveraging the measure's evidence of validity or the lack thereof to guide clear hypothesis testing and thus understanding of the construct.

Within the specific contexts of media multitasking research, the current dissertation demonstrated that media multitasking intensity negatively affects impacts well-being, predicting increases in stress. The most detrimental experiences emerge as those in which more sensory resources are intentionally allocated to media tasks, along with tasks that are difficult and relevant. Unexpectedly, teenagers engaged in lower media multitasking intensity than their parents. Media multitasking intensity, accordingly, appears to reflect the necessary integration of media multitasking into existing goals particularly for parents rather than distraction. Despite, being stressful, this does not imply that all media multitasking is harmful. The future is not necessarily doomed to being overwhelmed and stressed due to technology. Individuals are adapting not only to the stimuli in a media-saturated environment, but are also learning how to leverage media for their goals. They may need guidance and practice with prioritizing or reassessing the relevance of their media.

References

Afifi, T. D., Davis, S., Merrill, A. F., Coveleski, S., Denes, A., & Afifi, W. (2015). In the wake of the great recession: Economic uncertainty, communication, and biological stress responses in families. *Human Communication Research*, *41*(2), 268–302. https://doi.org/10.1111/hcre.12048

Afifi, T. D., Granger, D. A., Joseph, A., Denes, A., & Aldeis, D. (2015). The influence of divorce and parents' communication skills on adolescents' and young adults' stress reactivity and recovery. *Communication Research*, *42*(7), 1009-1042.

Afifi, T. D., Merrill, A. F., & Davis, S. (2016). The theory of resilience and relational load. *Personal Relationships*, *23*(4), 663-683.

Afifi, T. D., Zamanzadeh, N. N., Harrison, K., & Acevedo Callejas, M. (2017). WIRED: The impact of media and technology use on stress (cortisol) and inflammation (interleukin IL-6) in fast paced families. *Computers in Human Behavior, in press*.

Almond, R., Steinberg, L., & Mislevy, R. (2002). Enhancing the design and delivery of assessment systems: A four-process architecture. *The Journal of Technology, Learning and Assessment*, *1*(5). [if online only, provide URL]

Alzahabi, R., & Becker, M. W. (2013). The association between media multitasking, task-switching, and dual-task performance. *Journal of Experimental Psychology. Human Perception and Performance*, *39*(5), 1485–95. https://doi.org/10.1037/a0031208

Andersen, K., de Vreese, C. H., & Albæk, E. (2016). Measuring media diet in a high-

choice environment - testing the list-frequency technique. *Communication

Methods and Measures*, *10*(2–3), 81–98.

https://doi.org/10.1080/19312458.2016.1150973

AERA (1999). Standards for educational and psychological testing. *New York:

American Educational Research Association*.

Barkley, R. A. (2012). *Barkley deficits in executive functioning scale--children and

adolescents (BDEFS-CA)*. New York, NY: The Guilford Press.

Barratt, E. S. (1983). The biological basis of impulsiveness: The significance of timing

and rhythm disorders. *Personality and individual differences*, *4*(4), 387-391.

Barratt, E. S., Stanford, M. S., Kent, T. A., & Alan, F. (1997). Neuropsychological and

cognitive psychophysiological substrates of impulsive aggression. *Biological

Psychiatry, 41,* 1045–1061. http://dx.doi.org/ 10.1016/S0006-3223(96)00175-8

Baumeister, R. F., & Vohs, K. D. (2016). Strength model of self-regulation as limited

resource: Assessment, controversies, update. *Advances in Experimental Social

Psychology* (1st ed., Vol. 54). Elsevier Inc.

https://doi.org/10.1016/bs.aesp.2016.04.001

Baumgartner, S. E., & Sumter, S. R. (2017). Dealing with media distractions: An

observational study of computer-based multitasking among children and adults in

the Netherlands. *Journal of Children and Media*, *27*(May), 1–19.

Baumgartner, S. E., Weeda, W. D., van der Heijden, L. L., & Huizinga, M. (2014). The

relationship between media multitasking and executive function in early

adolescents. *The Journal of Early Adolescence*, *34*(8), 1120–1144. https://doi.org/10.1177/0272431614523133

Beatty, P. C., & Willis, G. B. (2007). Research synthesis: The practice of cognitive interviewing. *Public Opinion Quarterly*, *71*(2), 287-311.

Beauchamp, M. R., & McEwan, D. (2017). Response processes and measurement validity in health psychology. In B. D. Zumbo & A. M. Hubley (Eds.), *Understanding and investigating response processes in validation research* (pp. 13-30). Vancouver, BC, Canada: Springer International Publishing

Becker, M. W., Alzahabi, R., & Hopwood, C. J. (2013). Media multitasking is associated with symptoms of depression and social anxiety. *Cyberpsychology, Behavior, and Social Networking*, *16*(2), 132-135.

Bonanno, G. A. (2004). Loss, trauma, and human resilience: Have we underestimated the human capacity to thrive after extremely aversive events? *American Psychologist, 59*, 20-28.

Bond, T. G., & Fox, C. M. (2007). *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, NJ, US.

Borsboom, D. (2005). *Measuring the mind: Conceptual issues in contemporary psychometrics*. New York, NY: Cambridge University Press.

Borsboom, D. (2008). Latent variable theory. *Measurement: Interdisciplinary Research & Perspective*, *6*(1–2), 25–53. https://doi.org/10.1080/15366360802035497

Borsboom, D., Mellenbergh, G. J., & Heerden, J. Van. (2003). The theoretical status of latent variables, *110*(2), 203–219. https://doi.org/10.1037/0033-295X.110.2.203

Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2004). The concept of validity. *Psychological Review*, *111*(4), 1061–1071. https://doi.org/10.1037/0033-295X.111.4.1061

Borsboom, D., Cramer, A. O. J., Kievit, R. A., Scholten, A. Z., & Franic, S. (2009). The end of construct validity. *The Concept of Validity,* 135–170. [is this a book chapter? Need to format that correctly]

Boukes, M., & Vliegenthart, R. (2017). A general pattern in the construction of economic newsworthiness? Analyzing news factors in popular, quality, regional, and financial newspapers. *Journalism*. https://doi.org/10.1177/1464884917725989 [Is this a journal? Need vol(issue), pp-pp.]

boyd, D. (2014). *It's complicated: The social lives of networked teens*. New Haven, CT: Yale University Press.

Brasel, S. A., & Gips, J. (2011). Media multitasking behavior: Concurrent television and computer usage. *Cyberpsychology, Behavior and Social Networking*, *14*(9), 527–534. https://doi.org/10.1089/cyber.2010.0350f

Bridgman, P. W. (1927). The compressibility and pressure coefficient of resistance of ten elements. In *Proceedings of the American Academy of Arts and Sciences* 62(8), 207-226.

Brown, T. A. (2014). *Confirmatory factor analysis for applied research*. [Loc:] Guilford Publications.

Brown J.M., Miller, W.R., & Lawendowski, L.A. (1999). The self-regulation questionnaire. In L. VandeCreek & T. L. Jackson (Eds.), *Innovations in clinical*

practice: A sourcebook (pp. 281–292) Sarasota, FL: Professional Resource Press/Professional Resource Exchange.

Browne, M. W., & Cudeck, R. (1989). Single sample cross-validation indices for covariance structures. *Multivariate Behavioral Research*, *24*(4), 445-455.

Bucy, E. P., & Tao, C.C. (2007). The mediated moderation model of interactivity. *Media Psychology*, *9*(3), 647–672. https://doi.org/10.1080/15213260701283269

Byrne, B. M. (2008). Testing for multigroup equivalence of a measuring instrument: A walk through the process. *Psicothema*, *20*(4), 872-882.

Campbell, N.R. (1920). *Physics: The elements*. New York, NY: Cambridge University Press.

Carey, K. B., Neal, D. J., & Collins, S. E. (2004). A psychometric analysis of the self-regulation questionnaire. *Addictive Behaviors*, *29*(2), 253-260. [etc…I won't reformat any more]

Cardoso-Leite, P., Kludt, R., Vignola, G., Ma, W. J., Green, C. S., & Bavelier, D. (2015). Technology consumption and cognitive control: Contrasting action video game experience with media multitasking. *Attention, Perception, & Psychophysics*, *78*(1), 218–241. https://doi.org/10.3758/s13414-015-0988-0

Carrier, L. M., Black, V., Vasquez, L., Miller, A. D., & Rosen, L. D. (2015). Executive function in risky online behaviors by adolescents and young adults. In L. D. Rosen, N. A. Cheever, & L. M. Carrier (Eds.), *The Wiley Handbook of Psychology, Technology, and Society* (pp. 117-141). Hoboken, NJ: Wiley-Blackwell.

Carriere, J. S., Cheyne, J. A., & Smilek, D. (2008). Everyday attention lapses and memory failures: The affective consequences of mindlessness. *Consciousness and cognition*, *17*(3), 835-847.

Carver, C. S. (1998). Resilience and thriving: issues, models, and linkages. *Journal of Social Issues*, *54*(2), 245–266. https://doi.org/10.1111/j.1540-4560.1998.tb01217.x

Chalmers, R. P. (2012). mirt: A multidimensional item response theory package for the R environment. *Journal of Statistical Software*, *48*(6), 1-29.

Chang, H. (2010). Operationalism. *Stanford Encyclopedia of Philosophy*. Retrieved from http://www.science.uva.nl/~seop/archives/spr2010/entries/operationalism/%5Cnpapers3://publication/uuid/D1CD2295-1A71-4982-87B1-B6271989405D

Chang, Y. (2016). Why do young people multitask with multiple media? Explicating the relationships among sensation seeking, needs, and media multitasking behavior. *Media Psychology*, *0*(0), 1–19. https://doi.org/10.1080/15213269.2016.1247717

Chein, J. M., Wilmer, H. H., & Sherman, L. E. (2017). Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Frontiers in Psychology*, *8*(April), 1–16. https://doi.org/10.3389/fpsyg.2017.00605

Cheung, G. W. & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance, *Structural Equation Modeling: A Multidisciplinary Journal*, *9*(2), 233-255.

Cingel, D. P. & Krcmar, M. (2013). Predicting media use in very young children: The role of demographics and parent attitudes, *Communication Studies, 64*(4), 374-394.

Clairmont, A. (2020). *Towards an Anthropology of Measurement* (Unpublished doctoral dissertation). UC Santa Barbara, Santa Barbara, CA.

Clairmont, A., Wolf, M., & Maul, A. (2019) "The prevention and detection of deception in self-report survey data. In *Basic Elements of Survey Research in Education: Addressing the Problems Your Advisor Never Told You About.* IAP.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A global measure of perceived stress. *Journal of health and social behavior*, 385-396

Conrad, F. G., & Blair, J. (2009). Sources of error in cognitive interviews. *Public Opinion Quarterly*, *73*(1), 32-55.

Converse, P. E. (2006). The nature of belief systems in mass publics (1964). *Critical review*, *18*(1-3), 1-74.

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98–104.

Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16,* 297–334.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological bulletin,* 52(4), 281.

Csikszentmihalyi, M. (2000). Happiness, flow, and economic equality. *American Psychologist, 55*(10), 1163-1164. http://dx.doi.org/10.1037/0003-066X.55.10.1163

de Vreese, C. H., & Neijens, P. (2016). Measuring Media Exposure in a Changing

Communications Environment. *Communication Methods and Measures*, *10*(2–3),

69–80. https://doi.org/10.1080/19312458.2016.115044

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135-168.

Drew, D., & Weaver, D. (1990). Media attention, media exposure, and media

effects. *Journalism Quarterly*, *67*(4), 740-748.

Duff, B. R. L., Yoon, G., Wang, Z., & Anghelcev, G. (2014). Doing it all: An

exploratory study of predictors of media multitasking. *Journal of Interactive

Advertising*, *14*(1), 11-23.

Eisenberg, N. (2017). Commentary: What's in a word (or words) – on the relations

among self-regulation, self-control, executive functioning, effortful control,

cognitive control, impulsivity, risk-taking, and inhibition for developmental

psychopathology – reflections on Nigg (. *Journal of Child Psychology and

Psychiatry and Allied Disciplines*, *58*(4), 384–386.

https://doi.org/10.1111/jcpp.12707

Ellis, B. J., & Del Giudice, M. (2014). Beyond allostatic load: rethinking the role of

stress in regulating human development. *Development and Psychopathology*,

26(1), 1–20. https://doi.org/10.1017/S0954579413000849

Ercikan, K., & Pellegrino, J. W. (Eds.). (2017). *Validation of score meaning for the next

generation of assessments: The use of response processes*. New York, NY:

Routledge.

Evans, S. K., Pearce, K. E., Vitak, J., & Treem, J. W. (2016). Explicating affordances:

A conceptual framework for understanding affordances in communication

research, *Journal of Computer-Mediated Communication, 22*(1), 35–52. https://doi.org/10.1111/jcc4.12180

Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological methods*, *4*(3), 272.

Fan, Y., Gong, S., Wang, Y., & Wang, Z. (2017). The effects of learning task and learners' metacognition on college students' media multitasking behaviors, *Advances in Psychology, 7*(7), 895-902.

Fields, S., Edens, J. F., Smith, S. T., Rulseh, A., Donnellan, M. B., Ruiz, M. A., ... & Douglas, K. S. (2015). Examining the psychometric properties of the Barratt Impulsiveness Scale–Brief Form in justice-involved samples. *Psychological Assessment*, *27*(4), 1211.

Fikkers, K. M., Piotrowski, J. T., & Valkenburg, P. M. (2017). A matter of style? Exploring the effects of parental mediation styles on early adolescents' media violence exposure and aggression. *Computers in Human Behavior, 70,* 407–415.

Fishbein, M., & Hornik, R. (2008). Measuring media exposure: An introduction to the special issue. *Communication Methods and Measures, 2*(1–2), 1–5.

Fisher, J. T., & Keene, J. R. (2019). Attention, working memory, and media multitasking Proceedings from ICA '19: *The International Communication Association*. Washington D.C.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive–developmental inquiry. *American psychologist*, *34*(10), 906.

Folkman, S. (2013). Stress: appraisal and coping. In *Encyclopedia of behavioral medicine* (pp. 1913-1915). Springer New York.

Folkman, S., Lazarus, R. S., Gruen, R. J., & DeLongis, A. (1986). Appraisal, coping, health status, and psychological symptoms. *Journal of personality and social psychology*, *50*(3), 571.

Fox, J., & Moreland, J. J. (2015). The dark side of social networking sites: An exploration of the relational and psychological stressors associated with Facebook. *Computers in Human Behavior*, *45*(April), 168–176. https://doi.org/10.1016/j.chb.2014.11.083

Floyd, K., Pauley, P. M., & Hesse, C. (2010). State and Trait Affectionate Communication Buffer Adults' Stress Reactions. *Communication Monographs*, *77*(February 2015), 618–636. https://doi.org/10.1080/03637751.2010.49879

Gentile, D. A., & Bushman, B. J. (2012). Reassessing media violence effects using a risk and resilience approach to understanding aggression. *Psychology of Popular Media Culture*, *1*(3), 138.

Gorin, J. S. (2007). Reconsidering issues in validity theory. *Educational Researcher*, *36*(8), 456-462.

Green, S. B., Lissitz, R.W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, *37*, 827–838.

Gross, J. J. (Ed.) (2014). The handbook of emotional regulation (2nd edition). New York, NY: The Guilford Press.

Hagger, M. S., Wood, C., Stiff, C., & Chatzisarantis, N. L. (2009). The strength model of self-regulation failure and health-related behaviour. *Health Psychology Review*, *3*(2), 208-238.

Hargittai, E., Neuman, W. R., & Curry, O. (2012). Taming the information tide: Perceptions of information overload in the American home. *The Information Society*, *28*(3), 161-173.

Hedstrom, A. and Irwin, M. (2017). Mobile Experience Sampling Method (MESM). In The International Encyclopedia of Communication Research Methods (eds J. Matthes, C. S. Davis and R. F. Potter). doi:10.1002/9781118901731.iecrm0157

Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). Experience sampling method: Measuring the quality of everyday life. Thousand Oaks, CA, US: Sage Publications, Inc.

Hu, L. T., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural equation modeling: a multidisciplinary journal*, *6*(1), 1-55.

Humphry, S. (2013). Understanding measurement in light of its origins. *Frontiers in psychology*, *4*, 113.

Jake-Schoffman, D. E., Turner-McGrievy, G., & Walsemann, K. M. (2017). Wired: parent–child relationship quality and recreational media use in a diverse sample of US children and adolescents. *Journal of Children and Media*, *11*(3), 347-357.

Jennings, N. (2017). Media and families: Looking ahead. *Journal of Family Communication*, *17*(3), 203–207. https://doi.org/10.1080/15267431.2017.1322972

Jeong, S. H., & Fishbein, M. (2007). Predictors of multitasking with media: Media

    factors and audience factors. *Media Psychology*, *10*(3), 364-384.

Jeong, S. H., & Hwang, Y. (2016). Media multitasking effects on cognitive vs.

    attitudinal outcomes: A meta-analysis. *Human Communication Research*, *42*(4),

    599–618. https://doi.org/10.1111/hcre.12089

Jerit, J., Barabas, J., Pollock, W., Banducci, S., Stevens, D., & Schoonvelde, M. (2016).

    Manipulated vs. Measured: Using an Experimental Benchmark to Investigate the

    Performance of Self-Reported Media Exposure. *Communication Methods and*

    Measures, *10*(2–3), 99–114. https://doi.org/10.1080/19312458.2016.1150444

John, O. P., & Gross, J. J. (2004). Healthy and unhealthy emotion regulation:

    Personality processes, individual differences, and life span development. *Journal*

    of Personality, *72*(6), 1301–1333. https://doi.org/10.1111/j.1467-

    6494.2004.00298.x

Judd, T., & Kennedy, G. (2011). Measurement and evidence of computer-based task

    switching and multitasking by "Net Generation" students. *Computers &*

    Education, *56*(3), 625–631. https://doi.org/10.1016/j.compedu.2010.10.004

Kane, M. T. (2006). Validation. In R.L. Brennan (Ed.), *Educational measurement,* 4th

    edition (pp. 17-64). Santa Barbara: Greenwood Publishing Group.

Karr-Wisniewski, P., & Lu, Y. (2010). When more is too much: Operationalizing

    technology overload and exploring its impact on knowledge worker

    productivity. *Computers in Human Behavior*, *26*(5), 1061-1072.

Katz, E., & Foulkes, D. (1962). On the use of mass media for escape: Clarification of a

    concept. *Public Opinion Quarterly, 26*, 377-388.

Kline, R. B. (2015). *Principles and practice of structural equation modeling*. New York, NY: Guilford publications.

Kobayashi, T., & Boase, J. (2012). No such effect? The implications of measurement error in self-report measures of mobile communication use, *Communication Methods and Measures, 6*(2), 126-143, doi:10.1080/19312458.2012.679243

Konijn, E. A., Veldhuis, J., Plaisier, X. S., Spekman, M., Hamer, A. D. (2015). Adolescent development and psychological mechanisms in interactive media use. *The Handbook of the Psychology of Communication Technology*, 332–364.

Kononova, A. (2013). Multitasking across borders: A cross-national study of media multitasking behaviors, its antecedents, and outcomes. *International Journal of Communication*, *7*, 23.

Kononova, A., & Chiang, Y. H. (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*, 31-41.

Kononova, A., Joo, E., & Yuan, S. (2016). If I choose when to switch: Heavy multitaskers remember online content better than light multitaskers when they have the freedom to multitask. *Computers in Human Behavior, 65*, 567-575.

Kononova, A., McAlister, A., & Oh, H. J. (2018). Screen overload: Pleasant multitasking with screen devices leads to the choice of healthful over less healthful snacks when compared with unpleasant multitasking. *Computers in Human Behavior*, *80*, 1-11.

Lang, A. (*2009*). The limited capacity model of motivated media message processing. In R. L.

Nabi & M. B. Oliver (Eds.), *The SAGE handbook of media processes and effects* (pp. 193-204).

Lang, A. (2013). Discipline in Crisis? The Shifting Paradigm of Mass Communication Research.

Communication Theory, *23*(1), 10–24. https://doi.org/10.1111/comt.12000

Lang, A., & Chrzan, J. (2015). Media Multitasking: Good, Bad, or Ugly?. *Annals of the International Communication Association, 39*(1), 99-128.

LaRose, R. (2010). The problem of media habits. *Communication Theory, 20*, 194–222.

LaRose, R. (2015). The psychology of interactive media habits. *The Handbook of the Psychology of Communication Technology.*, 365–383. https://doi.org/10.1002/9781118426456.ch16

Lau, W. W. (2017). Effects of social media usage and social media multitasking on the academic performance of university students. *Computers in human behavior*, *68*, 286-291.

Lavee, Y., McCubbin, H. I., & Patterson, J. M. (1985). The double ABCX model of family stress and adaptation: An empirical test by analysis of structural equations with latent variables. *Journal of Marriage and the Family*, 811-825.

Lavie, N. (1995). Perceptual load as a necessary condition for selective attention. *Journal of Experimental Psychology: Human Perception and Performance, 21*(3), 451-468. http://dx.doi.org/10.1037/0096-1523.21.3.451

Lazarus, R. S. (1993). Coping theory and research: past, present, and future. *Psychosomatic medicine*, *55*(3), 234-247.

Lazarus, R. S., & Folkman, S. (1984). Coping and adaptation. *The handbook of behavioral medicine*, 282-325.

Lee, A. R., Son, S. M., & Kim, K. K. (2016). Information and communication technology overload and social networking service fatigue: A stress perspective. *Computers in Human Behavior*, *55*, 51–61. https://doi.org/10.1016/j.chb.2015.08.011

Lemola, S., Perkinson-Gloor, N., Brand, S., Dewald-Kaufmann, J. F., & Grob, A. (2014). Adolescents, electronic media use at night, sleep disturbance, and depressive symptoms in the smartphone age. *Journal of Youth and Adolescence*, *44*(2), 405–418. https://doi.org/10.1007/s10964-014-0176-x

Len-Ríos, M. E., Hughes, H. E., McKee, L. G., & Young, H. N. (2016). Early adolescents as publics: A national survey of teens with social media accounts, their media use preferences, parental mediation, and perceived Internet literacy. *Public Relations Review*, *42*(1), 101-108.

Levine, L. E., Waite, B. M., & Bowman, L. L. (2012). Mobile media use, multitasking and distractibility. *International Journal of Cyber Behavior, Psychology and Learning (IJCBPL)*, *2*(3), 15-29.

Levitin, D. J. (2014). *The organized mind: Thinking straight in the age of information overload*. New York, NY: Penguin Random House.

Levy, D. M., Wobbrock, J. O., Kaszniak, A. W., & Ostergren, M. (2012). The effects of mindfulness meditation training on multitasking in a high-stress information environment. In *Proceedings of Graphics Interface 2012* (pp. 45-52). Canadian Information Processing Society.

234

Lim, S., & Shim, H. (2016). Who multitasks on smartphones? Smartphone multitaskers'
 motivations and personality traits. *Cyberpsychology, Behavior, and Social
 Networking*, *19*(3), 223-227.

Livingstone, S., Ólafsson, K., Helsper, E. J., Lupiáñez-Villanueva, F., Veltri, G. A., &
 Folkvord, F. (2017). Maximizing opportunities and minimizing risks for children
 online: The role of digital skills in emerging strategies of parental
 mediation. *Journal of Communication*, *67*(1), 82-105.

Loh, K. K., & Kanai, R. (2016). How has the Internet reshaped human cognition*? The
 Neuroscientist*, 22(5). https://doi.org/10.1177/1073858415595005

MacKenzie, S. B., Podsakoff, P. M., & Podsakoff, N. P. (2011). Construct measurement
 and validation procedures in MIS and behavioral research: Integrating new and
 existing techniques. *MIS quarterly*, *35*(2), 293-334.

Magen, H. (2017). The relations between executive functions, media multitasking and
 polychronicity. *Computers in Human Behavior*, *67*, 1-9.

Mari, L. (2013). A quest for the definition of measurement. *Measurement*, *46*(8), 2889-
 2895.

Mari, L., Carbone, P., Giordani, A., & Petri, D. (2017). A structural interpretation of
 measurement and some related epistemological issues. *Studies in History and
 Philosophy of Science Part A*, *65*, 46-56.

Mari, L., Maul, A., Irribarra, D. T., & Wilson, M. (2016). A meta-structural
 understanding of measurement. *Journal of Physics: Conference Series*, *772*(1),
 12009. https://doi.org/10.1088/1742-6596/772/1/012009

Mari, L., Maul, A., Irribarra, D. T., & Wilson, M. (2017). Quantities, quantification, and the necessary and sufficient conditions for measurement. *Measurement*, *100*, 115-121.

Markus, K., & Borsboom, D. (2013). *Frontiers of test validity theory.* New York, NY: Routledge.

Masters, G. N. (1982). A Rasch model for partial credit scoring. *Psychometrika*, *47*(2), 149-174.

Matthews, G., & Campbell, S. E. (2010). Dynamic relationships between stress states and working memory. *Cognition and Emotion*, *24*(2), 357-373.

Maul, A. (2017). Rethinking Traditional Methods of Survey Validation. *Measurement*, *15*(2), 1–19. https://doi.org/10.1080/15366367.2017.1348108

Maul, A. (2013). On the ontology of psychological attributes. *Theory & Psychology*, *23*(6), 752–769. https://doi.org/10.1177/0959354313506273

Maul, A., Mari, L., & Wilson, M. (2019). Intersubjectivity of measurement across the sciences. *Measurement*, *131*, 764-770.

Maul, A., & McGrane, J. (2017). As Pragmatic as Theft Over Honest Toil: Disentangling Pragmatism From Operationalism. *Measurement*, *15*(1), 1–3. https://doi.org/10.1080/15366367.2017.1342484

Maul, A., Torres Irribarra, D., & Wilson, M. (2016). On the philosophical foundations of psychological measurement. *Measurement*, *79*, 311–320. https://doi.org/10.1016/j.measurement.2015.11.001

Mayer, J. D., & Gaschke, Y. N. (1988). The experience and meta-experience of mood. *Journal of personality and social psychology*, *55*(1), 102.

236

McDaniel, B. T. (2015). "Technoference": Everyday intrusions and interruptions of

    technology in couple and family relationships. In C. J. Bruess (Ed.), *Family*

    communication in the age of digital and social media. New York: Peter Lang

    Publishing.

McDaniel, B. T., & Radesky, J. S. (2017). Technoference: Parent distraction with

    technology and associations with child behavior problems. *Child Development*,

    89(1), 100-109.

McEwen, B. S. (2016). Central role of the brain in stress and adaptation: allostasis,

    biological embedding, and cumulative change. In *Stress: Concepts, Cognition,*

    Emotion, and Behavior (pp. 39-55).

McGrane, J. A. (2015). Stevens' forgotten crossroads: the divergent measurement

    traditions in the physical and psychological sciences from the mid-twentieth

    century. *Frontiers in Psychology*, *6*, 431.

Messick, S. (1989). Validity. In R. L. Linn (Ed.), *Educational measurement* (pp. 13–

    103). Washington, DC: American Council on Education and National Council on

    Measurement in Education.

Messick, S. (1995). Validity of psychological assessment: validation of inferences from

    persons' responses and performances as scientific inquiry into score meaning.

    American Psychologist, *50*(9), 741.

Michell, J. (1997). Quantitative science and the definition of measurement in

    psychology. *British journal of Psychology*, *88*(3), 355-383.

Michell, J. (2004). The place of qualitative research in psychology. *Qualitative research*

    in Psychology, *1*(4), 307-319.

Michell, J. (2008). Is Psychometrics Pathological Science? *Measurement: Interdisciplinary Research & Perspective*, *6*(1–2), 7–24. https://doi.org/10.1080/15366360802035489

Michell, J. (2013). Constructs, inferences, and mental measurement. *New Ideas in Psychology*, *31*(1), 13–21. https://doi.org/10.1016/j.newideapsych.2011.02.004

Michell, J. (2014). *An introduction to the logic of psychological measurement*. New York, NY: Psychology Press.

Millsap, R. E. (1997). Invariance in measurement and prediction: Their relationship in the single-factor case. *Psychological Methods*, *2*(3), 248–260. https://doi.org/10.1037/1082-989X.2.3.248

Minear, M., Brasher, F., McCurdy, M., Lewis, J., & Younggren, A. (2013). Working memory, fluid intelligence, and impulsiveness in heavy media multitaskers. *Psychonomic Bulletin & Review*, *20*, 1274–81. https://doi.org/10.3758/s13423-013-0456-6

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. *ETS Research Report Series*, *2003*(1), i-29. https://doi.org/10.1002/j.2333-8504.2003.tb01908.x

Mislevy, R. J. (2009). Validity from the perspective of model-based reasoning. *The concept of validity: Revisions, new directions and applications*, In R. W. Lissitz (Ed.), The concept of validity: Revisions, new directions, and applications (pp. 83–108). Charlotte, NC: Information Age.

Misran, S., & Stokols, D. (2012). Psychological and health outcomes of perceived

information overload. *Environment and Behavior*, *44*(6), 737–759.

https://doi.org/10.1177/0013916511404408

Moilanen, K. L. (2007). The adolescent self-regulatory inventory: The development and

validation of a questionnaire of short-term and long-term self-regulation. *Journal

of Youth and Adolescence*, *36*(6), 835-848.

Moy, P., & Murphy, J. (2016). Problems and prospects in survey research. *Journalism

& Mass Communication Quarterly*, *93*(1), 16–37.

https://doi.org/10.1177/1077699016631108

Muthén, L. K., & Muthén, B. O. (2017). 1998–2017. *Mplus User's Guide.* Los Angeles,

CA: Muthén & Muthén.

Nabi, R.L., & Oliver, M.B. (Eds.) (2009). *The SAGE handbook of media processes and

effects.* Thousand Oaks, CA: SAGE Publications.

Nathanson, A. I. (2015). Media and the family: Reflections and future directions.

Journal of Children and Media, *9*(1), 133–139.

https://doi.org/10.1080/17482798.2015.997145

Niederdeppe, J. (2014). Conceptual, empirical, and practical issues in developing valid

measures of public communication campaign exposure. *Communication

Methods and Measures, 8*(2), 138-161. DOI: 10.1080/19312458.2014.903391

Niederdeppe, J. (2016) Meeting the Challenge of Measuring Communication Exposure

in the Digital Age. *Communication Methods and Measures*, *10*(2-3), 170-172,

DOI: 10.1080/19312458.2016.1150970

Nigg, J. T. (2017). Annual Research Review: On the relations among self-regulation, self-control, executive functioning, effortful control, cognitive control, impulsivity, risk-taking, and inhibition for developmental psychopathology. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(4), 361–383. https://doi.org/10.1111/jcpp.12675

Nunnally, J. C. (1978). *Psychometric theory.* New York: McGraw-Hill.

Ophir, E., Nass, C., & Wagner, A. D. (2009). Cognitive control in media multitaskers. *Proceedings of the National Academy of Sciences of the United States of America, 106*(37), 15583–7. https://doi.org/10.1073/pnas.0903620106

Oremus, M., Cosby, J. L., & Wolfson, C. (2005). A hybrid qualitative method for pretesting questionnaires: The example of a questionnaire to caregivers of Alzheimer disease patients. *Research in nursing & health*, *28*(5), 419-430.

Padilla-Walker, L. M., Coyne, S. M., & Fraser, A. M. (2012). Getting a high-speed family connection: Associations between family media use and family connection. *Family Relations*, *61*(3), 426–440. https://doi.org/10.1111/j.1741-3729.2012.00710.x

Pea, R., Nass, C., Mehul, L., Ranke, M., Kumar, A., Bamford, H., … Zhou, M. (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. https://doi.org/10.1037/a0027030

Perse, E. M., & Lambe, J. (2016). *Media effects and society*. New York: Routledge.

Pool, M. M., Koolstra, C. M., & Voort, T. H. (2003). The impact of background radio and television on high school students' homework performance. *Journal of Communication*, *53*(1), 74-87.

PEW Research Center, Internet & Technology (2018) *Internet/Broadband Fact Sheet.* Retrieved from https://www.pewinternet.org/fact-sheet/internet-broadband/

PEW Research Center, Internet & Technology (2018) *Mobile Fact Sheet*. Retrieved from https://www.pewinternet.org/fact-sheet/mobile/

Potter, W. J., (2011). Conceptualizing mass media effect, *Journal of Communication*, *61*(5), 896–915, https://doi.org/10.1111/j.1460-2466.2011.01586.x

Potter, W. J. (2018). A review and analysis of patterns of design decisions in recent media effects research. *Review of Communication Research*, *6*, 1-29.

Potter, W. J., Riddle, K. (2007). A content analysis of the media effects literature. *Journalism & Mass Communication Quarterly, 84*(1), 90-104.

Preide, C. & Farrall, S. (2011). Comparing results from different styles of cognitive interviewing: 'Verbal probing' vs. 'Thinking aloud'. *International Journal of Social Research Methodology*, *14*(4), 271-287. DOI: 10.1080/13645579.2010.523187

Przybylski, A. K., & Weinstein, N. (2017). A large-scale test of the goldilocks hypothesis: Quantifying the relations between digital-screen use and the mental well-being of adolescents, *Psychological Science, 28*(2), 204-215. https://doi.org/10.1177/0956797616678438

Radesky, J., Eisenberg, S., Kistin, Caroline, J., Gross, J., Block, G., Zuckerman, B., & Silverstein, M. (2016). Overstimulated consumers or next-generation learners?

Parent tensions About child mobile technology use. *Annals of Family Medicine*, *14*, 503–508. https://doi.org/10.1370/afm.1976.

Radesky, J. S., Peacock-Chambers, E., Zuckerman, B., & Silverstein, M. (2016). Use of mobile technology to calm upset children: associations with social-emotional development. *JAMA pediatrics*, *170*(4), 397-399.

Rains, S. A., Levine, T. R., & Weber, R. (2018). Sixty years of quantitative communication research summarized: lessons from 149 meta-analyses, *Annals of the International Communication Association*, *42*(2), 105-124. https://doi.org/10.1080/23808985.2018.1446350

Ralph, B. C. W., & Smilek, D. (2017). Individual differences in media multitasking and performance on the n-back. *Attention, Perception, & Psychophysics,* 582–592. https://doi.org/10.3758/s13414-016-1260-y

Raykov, T. (2001). Bias of coefficient alpha for fixed congeneric measures with correlated errors. *Applied Psychological Measurement*, *25*, 69–76.

Reeves, B., & Nass, C. I. (1996). *The media equation: How people treat computers, television, and new media like real people and places*. Cambridge, UK: Cambridge University Press.

Reeves, B., Yeykelis, L., & Cummings, J. J. (2016). The use of media in media psychology. *Media Psychology*, *19*(1), 49–71. https://doi.org/10.1080/15213269.2015.1030083

Reinecke, L., Aufenanger, S., Beutel, M. E., Dreier, M., Quiring, O., Stark, B., … Dreier, M. (2017). Digital stress over the life span: The effects of communication load and internet multitasking on perceived stress and psychological health

impairments in a German probability sample. *Media Psychology*, *20*(1), 90-115. https://doi.org/10.1080/15213269.2015.1121832

Reinecke, L., & Eden, A. (2017). Media use and recreation: Media-induced recovery as a link between media exposure and well-being. In L. Reinecke & M.B. Oliver (Eds.), The Routledge handbook of media use and well-being: International perspectives on theory and research on positive media effects (pp. 106-117). New York: Routledge.

Reinecke, L., & Eden, A. (2017). Editorial media use and well-being. *Media Psychology, 29(3), 111-114.*

Reinecke, L., Meier, A., Aufenanger, S., Beutel, M. E., Dreier, M., Quiring, O., ... & Müller, K. W. (2018). Permanently online and permanently procrastinating? The mediating role of Internet use for the effects of trait procrastination on psychological health and well-being. *new media & society*, 20(3), 862-880.

Reinecke, L., & Oliver, M. B. (2016). Media Use and Well-Being. *The Routledge Handbook of Media Use and Well-Being*, *29*, 111–114. https://doi.org/10.4324/9781315714752.ch1

Rice, R. E. (1990). Computer-mediated communication system network data: theoretical concerns and empirical examples. *International Journal of Man-Machine Studies*, *32*(6), 627-647.

Rice, R.E., Evans, S. K., Pearce, K. E., Sivunen, A., Vitak, J., & Treem, J. W. (2017). Organizational media affordances: Operationalization and associations with media use, *Journal of Communication*, *67*(1,) 106–130. https://doi.org/10.1111/jcom.12273

Rice, R.E., Hagen, I., & Zamanzadeh, N. N. (2018) College students' media mastery: Paradoxes in using computers and mobile phones. *American Behavioral Scientist, 62(9), 1229-1250.*

Rice, R. E., Zamanzadeh, N.N., & Hagen, I. (2019) Media mastery by College Students: A Typology and Review. In S. J. Yates & Ron E. Rice (Eds.) *The Oxford Handbook of Digital Technology and Society*. New York, NY: Oxford University Press.

Risko, E. F., Buchanan, D., Medimorec, S., & Kingstone, A. (2013). Everyday attention: mind wandering and computer use during lectures. *Computers & Education, 68*, 275-283. Doi: 10.1016/j.compedu.2013.05.001

Robitzsch, A., & Robitzsch, M. A. (2019). Package 'sirt'

Rodriguez, M. C., & Maeda, Y. (2006). Meta-Analysis of coefficient alpha. *Psychological Methods*, *11*(3), 306–322.

Rosen, L. D., Lim, A. F., Carrier, L. M., & Cheever, N. A. (2011). An empirical examination of the educational impact of text message-induced task switching in the classroom: Educational implications and strategies to enhance learning. *Psicología educativa*, *17*(2), 163-177.

Rosen, L. D., Whaling, K., Carrier, L. M., Cheever, N. a, & Rokkum, J. (2013). The Media and Technology Usage and Attitudes Scale: An empirical investigation. *Computers in Human Behavior*, *29*(6), 2501–2511. https://doi.org/10.1016/j.chb.2013.06.006

Sanbonmatsu, D. M., Strayer, D. L., Medeiros-Ward, N., & Watson, J. M. (2013). Who multi-tasks and why? Multi-tasking ability, perceived multi-tasking ability,

impulsivity, and sensation seeking. *PloS One, 8*(1), e54402.

https://doi.org/10.1371/journal.pone.0054402

Salvucci, D. D., & Bogunovich, P. (2010). Multitasking and monotasking: the effects of

mental workload on deferred task interruptions. *Proceedings of the SIGCHI

Conference on Human Factors in Computing Systems*, 85–88.

https://doi.org/10.1145/1753326.1753340

Salvucci, D. D., Taatgen, N. a, & Borst, J. (2009). Toward a unified theory of the

multitasking continuum: From concurrent performance to task switching,

interruption, and resumption. Chi, 1819–1828.

https://doi.org/10.1145/1518701.1518981

Schilhab, T. (2017). Adaptive smart technology use: The need for meta-self-regulation.

Frontiers in Psychology, *8*(MAR), 8–11. https://doi.org/10.3389/fpsyg.2017.00298

Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*,

350–353.

Schneider, F. M., Reich, S., & Reinecke, L. (2017). Methodological challenges of

POPC for communication research. In P. Vorder, D. Hefner, L. Reinecke, and C.

Klimmt (Eds.) *Permanently online, permanently connected. Living and

communicating in a POPC world* (pp. 29-39). New York, New York: Routledge.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys:

Experiments on question form, wording, and context*. Thousand Oaks, CA: Sage.

Schutten, D., Stokes, K. A., & Arnell, K. M. (2017). I want to media multitask and I

want to do it now: Individual differences in media multitasking predict delay of

gratification and system-1 thinking. *Cognitive Research: Principles and Implications*, *2*(1), 8. https://doi.org/10.1186/s41235-016-0048-x

Seger, J., & Potts, R. (2017). Construct Validity of Adults' Retrospective Memory for Childhood TV Viewing: Self-Reported Early Exposure to Violent TV Programs Predicts Current Aggression. *Communication Methods and Measures*, *11*(1), 31–48. https://doi.org/10.1080/19312458.2016.1271114

Segijn, C. M., Voorveld, H. A. M., Vandeberg, L., Pennekamp, S. F., & Smit, E. G. (2017). Insight into everyday media use with multiple screens. *International Journal of Advertising, 36*(5), 779–797. https://doi.org/10.1080/02650487.2017.1348042

Segijn, C. M., Xiong, S., Duff, B. R. (2019). Manipulating and measuring media multitasking: Implications of previous research and guidelines for future research. *Communication Methods and Measures*.

Shih, S. I. (2013). A Null Relationship between media multitasking and well-being. PLoS ONE, 8(5).

Smallwood, J., & Schooler, J. W. (2006). The restless mind. *Psychological bulletin*, *132*(6), 946.

Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of cronbach's alpha. *Psychometrika*, *74*(1), 107–120. https://doi.org/10.1007/s11336-008-9101-0

Sireci, S. G. (2009). Packing and unpacking sources of validity evidence: History repeats itself again. *The Concept of Validity: Revisions, New Directions, and Applications.*, 19–37.

Slater, M. D. (2004). Operationalizing and analyzing exposure: The foundation of media effects research. *Journalism & Mass Communication Quarterly*, *81*(1), 168-183.

Slater, D. M., Peter, J. & Valkenburg, P. M. (2015). Message variability and heterogeneity: A core challenge for Communication research, *Annals of the International Communication Association*, *39*(1), 3-31.

Smilek, D., Carriere, J. S., & Cheyne, J. A. (2010). Failures of sustained attention in life, lab, and brain: ecological validity of the SART. *Neuropsychologia*, *48*(9), 2564-2570.

Sparks, G. G. (2015). *Media effects research: A basic overview, 5th edition*. Boston, MA: Cengage Learning.

Statista (2018). *Internet use worldwide. – Statistics & Facts.* Retrieved from https://www.statista.com/topics/1145/internet-usage-worldwide/

Statista (2018). *Internet of Things – Statistics & Facts.* Retrieved from https://www.statista.com/topics/2637/internet-of-things/

Stevens, S. S. (1946). On the theory of scales of measurement. *Science, 103(*2684)*, 677-680.*

Steinberg, L., Icenogle, G., Shulman, E. P., Breiner, K., Chein, J., Bacchini, D., … Takash, H. M. S. (2017). Around the world, adolescence is a time of heightened sensation seeking and immature self-regulation. *Developmental Science*. https://doi.org/10.1111/desc.12532

Steinfield, C., Ellison, N. B., & Lampe, C. (2008). Social capital, self-esteem, and use

    of online social network sites: A longitudinal analysis. *Journal of Applied*

    Developmental Psychology, *29*(6), 434-445.

Sulis, I., & Toland, M. D. (2017). Introduction to multilevel item response theory

    analysis: Descriptive and explanatory models. *The Journal of Early*

    Adolescence, *37*(1), 85-128.

Szumowska, E., Popławska-Boruc, A., Kuś, J., Osowiecka, M., & Kramarczyk, J.

    (2018). When frequent media multitaskers perform worse and when they do not:

    The role of self-regulation ability and strategy manipulation. *Computers in Human*

    Behavior, *83*, 184-193.

Tal, E. (2013). Old and New Problems in philosophy of measurement. *Philosophy*

    Compass, *12*, 1159–1173

Tal, E. (2016). How Does Measuring Generate Evidence? The Problem of

    Observational Grounding. *Journal of Physics: Conference Series*, *772*, 12001.

    https://doi.org/10.1088/1742-6596/772/1/012001

Thissen, D., Steinberg, L., & Wainer, H. (1993). Detection of differential item

    functioning using the parameters of item response models. In P. W. Holland & H.

    Wainer (Eds.), *Differential item functioning* (pp. 67-113). Hillsdale, NJ, US:

    Lawrence Erlbaum Associates, Inc.

Toland, M. D. (2014). Practical guide to conducting an item response theory

    analysis. *The Journal of Early Adolescence*, *34*(1), 120-151.

Treem, J. W. & Leonardi, P. M. (2013). Social Media Use in Organizations: Exploring the Affordances of Visibility, Editability, Persistence, and Association, Annals of the International Communication Association, *36*(1), 143-189,

Turkle, S. (2011). *Alone together: Why we expect more from technology and less from each other*. New York, NY: Basic Books.

Uncapher, M. R., Thieu, M. K., & Wagner, A. D. (2016). Media multitasking and memory: Differences in working memory and long-term memory. *Psychonomic Bulletin & Review*, *23*(2), 483–90. http://doi.org/10.3758/s13423-015-0907-3

Valkenburg, P. M., & Peter, J. (2013). The differential susceptibility to media effects model. *Journal of Communication*, *63*(2), 221–243. https://doi.org/10.1111/jcom.12024

Valkenburg, P. M., Peter, J., & Walther, J. B. (2016). Media Effects: Theory and Research. *Annual Review of Psychology*, *67*(1), 315–338. https://doi.org/10.1146/annurev-psych-122414-033608

Van Cauwenberge, A., Schaap, G., & Van Roy, R. (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,* 100-109.

Van der Schuur, W. A., Baumgartner, S. E., Sumter, S. R., & Valkenburg, P. M. (2015). The consequences of media multitasking for youth: A review. *Computers in Human Behavior,* 53, 204–215. https://doi.org/10.1016/j.chb.2015.06.035

Voorveld, H. A., & van der Goot, M. (2013). Age differences in media multitasking: A diary study. *Journal of Broadcasting & Electronic Media, 57*(3), 392-408.

Voorveld, H. A., & Viswanathan, V. (2015). An observational study on how situational

    factors influence media multitasking with TV: The role of genres, dayparts, and

    social viewing. *Media Psychology*, *18*(4), 499-526.

Walsh, F. (2016). Family resilience: a developmental systems framework. *European*

    Journal of Developmental Psychology, *13*(3), 313-324.

Wang, W. C., & Chen, C. T. (2005). Item parameter recovery, standard error estimates,

    and fit statistics of the WINSTEPS program for the family of Rasch

    models. *Educational and Psychological Measurement*, *65*(3), 376-404.

Wang, Z., Irwin, M., Cooper, C., & Srivastava, J. (2015). Multidimensions of media

    multitasking and adaptive media selection. *Human Communication Research*, *41*,

    102–127. https://doi.org/10.1111/hcre.12042

Wang, Y., Niiya, M., Mark, G., Reich, S., & Warschauer, M. (2015). Coming of age

    (digitally): An ecological view of social media use among college students.

    Proceedings of CSCW, 571–582. doi:10.1145/2675133.2675271

Warren, R. (2016). "Someday you'll have children just like you": what tomorrow's

    parents can teach us about parental mediation research. *Journal of Children and*

    Media, *10*(2), 173-180.

Widaman, K. F., Ferrer, E., & Conger, R. D. (2010). Factorial invariance within

    longitudinal structural equation models: Measuring the same construct across

    time. *Child development perspectives*, *4*(1), 10-18.

Willis, G. B. (2005). Cognitive interviewing in practice: think-aloud, verbal probing and

    other techniques. *Cognitive interviewing: a tool for improving questionnaire*

    design. London: Sage Publications.

Wilmer, H. H., & Chein, J. M. (2016). Mobile technology habits: patterns of association among device usage, intertemporal preference, impulse control, and reward sensitivity. *Psychonomic Bulletin & Review*, 1–8. https://doi.org/10.3758/s13423-016-1011-z

Wilmer, H. H., Sherman, L. E., Chein, J. M. (2017). Smartphones and cognition: A review of research exploring the links between mobile technology habits and cognitive functioning. *Frontiers in Psychology*, *8*(April), 1–16. https://doi.org/10.3389/fpsyg.2017.00605

Wiradhany, W., & Nieuwenstein, M. R. (2017). Cognitive control in media multitaskers: Two replication studies and a meta-analysis. *Attention, Perception, & Psychophysics, 79*(8), 2620-2641. https://doi.org/10.3758/s13414-017-1408-4

Wolf, M. G., Ihm, E., Maul, A., & Taves, A. (2019). Survey Item Validation. In M. Stausber & S. Engler (Eds.) *The Handbook of Research in the Study of Religion* (2nd ed). New York, NY: Routledge.

Wood, W., Tam, L., & Guerrero Witt, M. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, *88(*6),918-933.

Wu, M., & Adams, R. J. (2013). Properties of Rasch residual fit statistics. *Journal of Applied Measurement*, *14*(4), 339-355.

Xu, S., Wang, Z. (Joyce), & David, P. (2016). Media multitasking and well-being of university students. *Computers in Human Behavior*, *55*, 242–250. https://doi.org/10.1016/j.chb.2015.08.040

Yeykelis, L., Cummings, J. J., & Reeves, B. (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), 167–192.

http://doi.org/10.1111/jcom

Yeykelis, L., Cummings, J. J., & Reeves, B. (2018). The fragmentation of work, entertainment, E-Mail, and news on a personal computer: Motivational predictors of switching between media content. *Media Psychology*, *21*(3), 377-402.

Zhang, W. (2014). Learning variables, in-class laptop multitasking and academic performance: A path analysis. *Computers & Education*, *81*, 82–88. https://doi.org/10.1016/j.compedu.2014.09.012

Zimmerman, D. W., Williams, R. H., Zumbo, B. D., & Ross, D. (2005). Louis Guttman's contributions to classical test theory. *International journal of testing*, *5*(1), 81-95.

Zuckerman, M. (1984). Sensation seeking: A comparative approach to a human trait. *Behavioral and Brain Sciences, 7*(3), 413-471. http://dx.doi.org/10.1017/S0140525X00018938

Zumbo, B. D., & Chan, E. K. (Eds.). (2014). *Validity and validation in social, behavioral, and health sciences.* New York: Springer International Publishing.