Age Differences in Theory of Mind: An Investigation of Neurocognitive, Health, and Demographic Predictors

by

Ashley L. Fischer

M.A., Simon Fraser University, 2010
B.Sc., University of Alberta, 2007

Dissertation Submitted in Partial Fulfillment of the
Requirements for the Degree of
Doctor of Philosophy

in the
Department of Psychology
Faculty of Arts and Social Sciences

SIMON FRASER UNIVERSITY

Summer 2015

Approval

Name: **Ashley L. Fischer**

Degree: **Doctor of Philosophy**

Title: ***Age Differences in Theory of Mind: An Investigation of Neurocognitive, Health, and Demographic Predictors***

Examining Committee: **Chair:** Dr. Robert McMahon
Professor

Dr. Wendy Loken Thornton
Senior Supervisor
Associate Professor

Dr. Allen Thornton
Supervisor
Associate Professor

Dr. Norm O'Rourke
Supervisor
Associate Professor
IRMACS Centre

Dr. Maureen Hoskyn
Internal Examiner
Associate Professor
Faculty of Education

Dr. Peter Graf
External Examiner
Professor
Department of Psychology
University of British Columbia

Date Defended/Approved: June 22, 2015

Ethics Statement

The author, whose name appears on the title page of this work, has obtained, for the research described in this work, either:

a. human research ethics approval from the Simon Fraser University Office of Research Ethics,

or

b. advance approval of the animal care protocol from the University Animal Care Committee of Simon Fraser University;

or has conducted the research

c. as a co-investigator, collaborator or research assistant in a research project approved in advance,

or

d. as a member of a course approved in advance for minimal risk human research, by the Office of Research Ethics.

A copy of the approval letter has been filed at the Theses Office of the University Library at the time of submission of this thesis or project.

The original application for approval and letter of approval are filed with the relevant offices. Inquiries may be directed to those authorities.

Simon Fraser University Library
Burnaby, British Columbia, Canada

update Spring 2010

Abstract

Theory of mind (ToM) is the ability to understand and reason about a variety of meta-cognitive and emotional mental states. Compared to young adults, older adults are more susceptible to reduced ToM, though the fundamental supporting processes are unclear. Earlier work demonstrates that neurocognitive performance, health status, and biological sex differences each contribute to ToM variability, yet no research has examined these predictors concurrently. In this dissertation we examined how these key predictors related to age differences in the cognitive and affective components of ToM. We tested 86 young (mean age = 19.8) and 85 older adults (mean age = 71.4) on standardized measures assessing neurocognitive performance and ToM. Predictor variables were derived from demographic information (sex), in-office blood pressure readings (pulse pressure or PP), and measures of three neurocognitive domains closely linked to ToM: executive functions, verbal comprehension, and episodic memory. We used path analysis to identify concurrent predictors of cognitive and affective ToM between age groups and partial invariance analyses to assess age differences in the strength of identified predictors. Our findings make several important contributions to this literature. We provide the first evidence that poor vascular health (high PP) directly predicts lower cognitive ToM across age groups, beyond other explanatory variables. Furthermore, in agreement with child development and cognitive neuroscience theory, we present the first neuropsychological evidence suggesting that cognitive ToM is a key predictor of affective ToM performance. Finally, while certain neurocognitive predictors of ToM are more salient in later life, we demonstrated that most predictors are shared between age groups and are equivalent in magnitude. Taken together, our study represents the most comprehensive investigation of predictors of ToM in aging to date, and suggests the value of continued investigation of ToM within a multidimensional framework.

Keywords: theory of mind; cognitive aging; blood pressure; sex differences; neurocognitive performance

For my grandmother, Clara.

Acknowledgements

This dissertation represents the culmination of a long and significant journey – one that would not have been possible without the contributions of many people along the way. These incredible supports have carried me through the ups and downs of completing my Ph.D. and have inspired me in so many ways. This work goes out to you.

It is with gratitude that I acknowledge the participants who eagerly volunteered their time and energy. I never ceased to be amazed by their enthusiasm and genuine interest in this research. Thanks also to the Social Sciences and Humanities Research Council of Canada for providing the financial support that enabled me to complete this project.

I am sincerely grateful to my senior supervisor, Dr. Wendy Loken Thornton for her dedication to my development as a graduate student, researcher, and clinician. I am fortunate to have benefited not only from your academic expertise but also your mentorship and advice that continue to influence my career in neuropsychology.

I would like to express my thanks to my supervisory committee Dr. Allen Thornton, Dr. Norm O'Rourke, and (for many years) Dr. Daniel Bernstein. Thank you for sharing your expertise and unique perspectives, and for your willingness to share constructive feedback. Your guidance and support shaped this project in many important ways.

I am incredibly lucky to have been part of the SFU Cognitive Aging Laboratory where I never lacked in support, inspiration, or friendship. I would like to express particular thanks to KB Zaidi, Scott Greenwood, and Janet Watts for dedicating their time and spirit to this project. To Sophie Yeung and Theone Paterson – thank you for paving the way and for your continued friendship, support, and so many memories. To my dear friends who were (and still are) my daily encouragement and inspiration – especially Kristina, Chantelle, Tania, Jenn, and Tyler – thank you for always being willing to lend an ear to listen, eye to read a draft, shoulder to cry on, evenings and weekends at coffee shops, or to go on long runs to de-stress. I cannot imagine this experience without you.

Finally, and most importantly, I am forever grateful to my dearest friends and family who have shared in every step of this journey, especially celebrating the successes. Special thanks to my brother Josh and sister Courtney who wholeheartedly supported my dreams and who were always willing to distract me with love and laughter. To my father Werner and mother Dianne – you have unconditionally loved, kept me grounded, and believed in me at times when I couldn't believe in myself. Thank you for your unwavering support. I am who I am today because of you. And last, to Chris – you have been my anchor since the day you selflessly encouraged me to hand in my graduate school application to SFU. Little did we both know what we had set in motion. Your love, faith, and confidence in me mean so much. Thank you for never letting go of my dreams.

Table of Contents

List of Tables

List of Figures

List of Acronyms & Initialisms

AD	Alzheimer's disease
bvFTD	Behavioural-variant frontotemporal dementia
CES-D	Centre for Epidemiological Studies Depression Scale
CI	Confidence interval
CFI	Comparative fit index
CVRF	Cardiovascular risk factor
CVLT-II	California Verbal Learning Test, 2nd Edition
DBP	Diastolic blood pressure
DKEFS	Delis-Kaplan Executive Functioning System
EC	Expected change
ES	Effect size
ICC	Intra-class correlation coefficient
KBIT-2	Kaufman Brief Intelligence Test, 2nd Edition
MAQ	Multidimensional Anxiety Questionnaire
MCI	Mild cognitive impairment
MI	Modification index
MMSE	Mini Mental Status Examination
PD	Parkinson's disease
PP	Pulse pressure
RSMEA	Root mean square error of approximation
SBP	Systolic blood pressure
SFU	Simon Fraser University
SRMR	Standardized root mean square residual
ToM	Theory of mind
WAIS-III	Wechsler Adult Intelligence Scale, 3rd Edition

Chapter 1. Introduction

Successfully navigating the social world relies on the ability to accurately predict what other people will think, feel, or do in various situations. For example, understanding that a loved one is upset over an earlier argument allows for a cautious approach to the topic and may lower the potential for further negative emotions. The capacity to consider another's beliefs, intentions, feelings, or motivations and use this knowledge to predict and explain behaviour is known as *theory of mind* (ToM) (Premack & Woodruff, 1978). Various terms exist to describe ToM, including mental state reasoning, mentalizing, cognitive empathy, and emotional awareness. Despite the wealth of data describing how humans interact with the social environment, the myriad of terms and disagreement regarding how to define and measure ToM creates ongoing challenges for researchers in this field. We use the terms 'ToM' and 'mental state reasoning' to refer to the ability to understand and reason about self- and other- mental state perspectives.

For most people, ToM is second nature and occurs without much insight into its process. Yet research suggests that ToM skills, as measured in the laboratory, are actually quite variable (Moran, 2013). In particular, older adults perform worse than young adults on measures of ToM, which can have clear and important consequences for daily social functioning (Bailey, Henry, & von Hippel, 2008; Henry, Phillips, Ruffman, & Bailey, 2013; Sandoz, Démonet, & Fossard, 2014). However, comparatively less is known about what factors underlie such differences. Concomitant reductions in age-sensitive neurocognitive abilities including executive functions, language, and episodic memory are one possibility (Apperly, Samson, & Humphreys, 2005), but these predictors are shown to account for less than 20% of overall variance in ToM performance (Ahmed & Miller, 2011; Bernstein, Thornton, & Sommerville, 2011; Fischer, Bernstein, & Thornton, 2014). More research is needed to clarify how and when neurocognitive processes predict ToM, and to determine the role of other potential explanatory variables. In a recent study (Fischer et al., 2014) we examined neurocognitive functions and vascular health as modifiers of ToM performance in a sample of community-residing

older adults. We observed strong associations between reduced ToM, older age, and worse performance on a composite measure of episodic memory and speed. We also demonstrated the first evidence that vascular health, an important predictor of general neurocognitive difficulties (Gifford et al., 2013), modified age reductions in ToM. These results provided evidence that concurrent examination of non-cognitive predictors affords a more nuanced view of how ToM changes with age. In this dissertation investigated ToM within a multidimensional framework wherein age differences were associated with fundamental neurocognitive, health, and demographic predictors.

1.1. Contemporary Perspectives on ToM

1.1.1. Age Differences and Clinical Relevance

Research portrays a consistent picture regarding the relevance of age differences in ToM: Older adults show significantly worse performance across measures of ToM compared to their younger peers (Cavallini, Lecce, Bottiroli, Palladino, & Pagnin, 2013; Pardini & Nichelli, 2009; see also Henry et al., 2013; Sandoz et al., 2014). Two points are noteworthy regarding these findings. First, the magnitude of age effects is considerable. In a recent meta-analysis Henry and colleagues (2013) demonstrated that effect sizes (ES) for age differences in ToM are at least moderate (mean weighted ES *r* = -.41; 95% CI [-.23 to -.48])[1]. As a practical benchmark, age effects in the ToM literature are similar in magnitude to those reported for prospective and retrospective memory (e.g., *r*'s = -.39 and -.52, respectively: Henry, MacLeod, Phillips, & Crawford, 2004). Second, age differences generalize across varied ToM measures and conceptual definitions (Henry et al., 2013). This point is critical in that it signifies that core difficulties in older adults' ToM performance are not driven by study-specific characteristics. Together this evidence is compelling and contradicts prior reports of maintained or

[1] Classified by the dissertation author according to O'Rourke, Hatcher, & Stepanski's (2005) recommendations for interpretation: | 0 < *r* < .20 | = "absent" degree of association, | .21 < *r* < .35 | = "low" degree of association, | .36 < *r* < .50 | = "moderate" degree of association, and | .51 < *r* < .80 | = "high" degree of association. Negative estimates indicate that older adults performed worse than young adults.

improved ToM in older age (e.g., Happé, Winner, & Brownell, 1998; Keightley, Winocur, Burianova, Hongwanishkul, & Grady, 2006; MacPherson, Phillips, & Della Sala, 2002).

Parallel to research in healthy aging, moderate to severe declines in ToM are reported among individuals with age-related neurodegenerative illness, including Alzheimer's disease (AD), Parkinson's disease (PD), and behavioural variant frontotemporal dementia (bvFTD) (Kemp, Després, Sellal, & Dufour, 2012; Sandoz et al., 2014). Studying ToM in these populations provides insight into how ToM changes as a function of age and disease process and may assist in refining theories about factors that are fundamental to ToM integrity in healthy aging. Individuals with AD, PD, and bvFTD often exhibit poor social skills and indifference toward others that is at least partly associated with poor ToM (Narme, Mouras, Roussel, Devendeville, & Godefroy, 2013a; Shany-Ur & Rankin, 2011). Specific patterns of ToM impairment are also useful diagnostic markers to distinguish among disease entities (e.g., AD versus bvFTD; Le Bouc et al., 2012). Performance distinctions may signal disease-specific neurodegenerative processes on the integrity of brain regions crucial for ToM (i.e., prefrontal structures and temporal-parietal cortex; Schlaffke et al., 2015), as well as correlated deficits in standard neurocognitive domains (Kemp et al., 2012).

Clinical interest is further motivated by evidence that ToM difficulties are observed in a large number of illness pathologies. Appendix A, Table A1 summarizes this literature. In populations where poor ToM is a prominent clinical feature (e.g., autism spectrum disorder, schizophrenia), difficulties are linked to decreased social engagement and reduced functional capacity (e.g., Fett et al., 2011). There is also preliminary evidence to suggest that poor ToM is associated with decreased participation in social activities among community-residing older adults (Bailey et al., 2008). Such outcome studies emphasize ToM as a potentially significant predictor of real-world behaviour. This idea is also reflected in the clinical practice of psychology and psychiatry, wherein the Diagnostic and Statistical Manual of Mental Disorders Fifth Edition (DSM-5; American Psychiatric Association, 2014) now lists social cognition (including ToM) as one of six neurocognitive domains to be considered when diagnosing *Minor and Major Neurocognitive Disorders* (i.e., mild cognitive impairment [MCI] and/or dementia). As a result, some authors have called for ToM to be assessed routinely in clinical neuropsychological evaluations (Adenzato & Poletti, 2013). Altogether the

collective interest in this topic highlights a need for more nuanced investigations regarding how ToM is maintained across adulthood.

1.1.2. The Multidimensional Structure of ToM

Our basic theoretical framework views ToM as a multidimensional construct comprised of two distinct, but partly overlapping, components[2]: 'Cognitive ToM' refers to the ability to reason about meta-cognitive beliefs, whereas 'affective ToM' refers to the ability to reason about emotional mental states (Shamay-Tsoory & Aharon-Peretz, 2007)[3]. What is central to both components is the *understanding of mental states*. Evidence supporting this conceptual distinction rests on differences in: (a) cognitive neuroscience theory; (b) developmental research in childhood and older age; (c) patterns of neuroanatomical localization and selective sparing and/or impairment observed in lesion case studies. Two primary conceptual perspectives are proposed to explain how humans come to understand others' mental states: 'Theory-theories' and 'simulation theories'. Theory-theories suggest that individuals use cognitive terms to describe unobservable mental states until an inference is formed, akin to building a scientific theory (Carruthers & Smith, 1996). In contrast, simulation perspectives suggest that individuals use their own minds as models to base inferences about others' mental states (Gallese & Goldman, 1998). The central premise of both perspectives is that specific neural architecture is reserved for, and activated by mental state inferences (i.e., mirror neurons and/or a core ToM network; Mahy, Moses, & Pfeifer, 2014a). Further, some authors speculate that these theories differentially explain cognitive and affective ToM, such that cognitive ToM involves theory-building processes and affective ToM involves simulation processes (Shamay-Tsoory, Tomer, Berger, Goldsher, & Aharon-

[2] We use the descriptor 'multidimensional' throughout this manuscript to refer to our conceptualization of ToM as componential or composite type of cognition whose development and use relies on input and interaction with a number of supporting resources.

[3] Affective ToM is often erroneously equated to empathy. Despite shared characteristics, the two constructs differ in their definition and function. Specifically, empathy occurs when the act of identifying another's emotions causes a change in one's own affective state that is an appropriate response. In contrast, affective ToM refers to an individual's active attempt to get "inside" another's emotional mind through a deliberative process and without necessarily experiencing that emotion (Decety & Jackson, 2004). Affective ToM is sometimes referred to in the literature as "cognitive empathy" (Dvash & Shamay-Tsoory, 2014).

Peretz, 2005). However, both theory-theories and simulation perspectives lack adequate explanation for the inter-relationship between cognitive and affective ToM and the potential contributions of supporting processes (Stone & Gerrans, 2006).

Like other types of componential cognition (e.g., everyday problem solving, Thornton, Deria, Gelb, Shapiro, & Hill, 2007), ToM is supported by perceptual precursors and neurocognitive processes in development and across the lifespan (Adolphs, 2003; Carlson & Moses, 2001; Mahy et al., 2014b). At the essence of any social interaction, humans first rely on basic perceptual skills such as face detection, gaze processing, emotion recognition, and joint attention, to gather and process social information (Adolphs, 2003; Stone & Gerrans, 2006). ToM requires information processing beyond what is extractable using elementary gaze detection or emotion recognition, though in-line interaction with these skills persists across the lifespan. As outlined by Wellman & Lagattuta (2000), children develop the ability to understand that others' beliefs may differ from their own by age three or four. This first-order cognitive ToM co-develops with language and executive functions (Astington & Jenkins, 1999; Carlson & Moses, 2001) and forms the basis for second-order cognitive ToM to develop by age six or seven (i.e., holding two or more mental states simultaneously). Advanced mental state reasoning, including the affective ToM attributions sarcasm, social faux pas, and moral judgment, emerges later in childhood and adolescence (Bosco, Gabbatore, & Tirassa, 2014; Vetter, Altgassen, Phillips, Mahy, & Kliegel, 2013). Shamay-Tsoory, Aharon-Peretz, and Perry (2009) describe the organizational structure of ToM as hierarchical and integrative: Individuals first require perceptual skills to detect social cues, followed by first- and second-order cognitive ToM to understand mental perspectives and infer behaviour, and empathy to experientially grasp interpersonal emotions. It is the integration of these processes that enables a functional affective ToM. This perspective is supported by evidence that children develop cognitive ToM at a younger age than affective ToM, as they must first understand others' beliefs in order to appreciate that beliefs guide emotions (Vetter et al., 2013). For example, to realize that a social faux pas has occurred, individuals need to represent two mental states: (1) that the other person does not know this is wrong to say, and (2) that the person hearing the comment might feel hurt. Thus, both cognitive and affective reasoning are utilized in affective ToM.

Dissociation between cognitive and affective ToM does not strictly imply that the components are always separate, rather that they *can be* separated (Dvash & Shamay-Tsoory, 2014). This notion is not captured by the early theories of ToM cited earlier. Shamay-Tsoory and colleagues (2009) outlined two plausible models to explain the cognitive-affective ToM relationship. The first model implies 'exclusivity'. In this model cognitive and affective ToM have separate neural bases—this is supported by evidence that selective brain lesions can impair cognitive ToM but spare affective ToM (and vice versa; Shamay-Tsoory & Aharon-Peretz, 2007). The second model implies 'dependence'. This model assumes that because cognitive ToM is more automatic and develops earlier than affective ToM, it might be a pre-condition to affective ToM. While empirical support is available for both models, the latter aligns best with neuroimaging and developmental research (Dvash & Shamay-Tsoory, 2014; Vetter et al., 2013). The broad ToM neural architecture comprises bilateral involvement of superior temporal sulci, temporal poles, and prefrontal cortex (Schlaffke et al., 2015). Each region plays a prominent but non-exclusive role in controlling ToM. Both cognitive and affective ToM depend on prefrontal cortex; however, where cognitive ToM relies on the larger ToM network, affective ToM relies specifically on contributions of ventromedial prefrontal cortex, where integration of cognitive and affective mental state information is thought to take place (Schlaffke et al., 2015). It is suggested that the ventromedial prefrontal cortex, through rich connections with the anterior insula, inferior parietal region and the limbic system, is uniquely positioned to regulate incoming emotional information and integrate it with cognitive representations (Shamay-Tsoory et al., 2009).

1.1.3. Measurement of ToM in Adulthood and Aging

Numerous measures exist to examine cognitive and affective ToM and vary by developmental complexity. Measures that tap simplistic ToM skills (first- and second-order ToM) are most appropriate for young children and adults with known cognitive difficulties, whereas measures that tap advanced ToM skills (e.g., social faux pas, moral judgment) are suitable for adolescents and healthy adults (Moran, 2013). In adult research, ToM measures generally fall into one of four categories: False belief reasoning (e.g., German & Hehman, 2006; Bernstein et al., 2011), reasoning about written stories or videos (e.g., Happé et al., 1998; Sullivan & Ruffman, 2004), social faux pas detection

(e.g., Stone, Baron-Cohen, & Knight, 1998), and emotional mental state reasoning (e.g., Baron-Cohen, Wheelwright, Hill, Raste, & Plumb, 2001). Each of these categories can be separated into first- and second-order question types; however, most studies have not analyzed data at this level of detail. Conclusions from review articles illustrate several additional limitations of existing ToM research (Kemp et al., 2012 and Moran, 2013). For one, the psychometric properties of ToM measures used in adult research are rarely reported. To our knowledge only six studies of ToM in aging report sample-specific reliability and most estimates concern inter-rater agreement (see Appendix B, Table B1). Little is known about the internal consistency or test-retest reliability of ToM scores, and what is reported varies widely (e.g., internal consistency: α = .49 to .85, Phillips, MacLean, & Allen, 2002; α = .91 to .93, Yeh, 2013). This point is concerning since scale reliability is central to statistical power and bears directly on interpretation (Henson, 2001). Further, despite similar operational definitions for ToM, most studies report low to moderate inter-test correlations (e.g., $|.01 < r < .58|$; Ahmed & Miller, 2011; Phillips et al., 2002; Rakoczy, Harder-Kasten, & Sturm, 2012; Saltzman, Strauss, Hunter & Archibald, 2000). While measurement error or discrepancies in test selection may play a role, it is also possible that these findings reflect the multidimensional nature of ToM. Erroneous conclusions may arise when studies conflate cognitive and affective ToM—constructs that hold differential relationships with age and neurocognitive performance. Even fewer studies have delineated between first- and second-order question types. One reason for this is that most adults show ceiling or near-ceiling performance on first- and second-order ToM questions. This has led researchers to rely on other populations such as typically developing children (e.g., Carlson & Moses, 2001), individuals with schizophrenia (e.g., Fett et al., 2011), and individuals with neurodegenerative illness (e.g., Le Bouc et al., 2012; Moreau et al., 2015) to gain insight into the organizational structure of ToM.

Few neuropsychological studies involving healthy adults and older adults have clearly distinguished between cognitive and affective ToM (Duval, Piolino, Bejanin, Eustache, & Desgranges, 2011; Mahy et al., 2014b; Rakoczy et al., 2012; Wang & Su, 2013; Ze, Thoma, & Suchan, 2014). These studies primarily utilize measures requiring individuals to reason about mental states from short stories to assess cognitive ToM (e.g., Happé et al., 1998), and Baron-Cohen and colleagues' (2001) Reading the Mind in

the Eyes Test (RMET) or related instruments to assess affective ToM. Despite common measurement tools discrepant findings exist. Disagreement pertains to whether or not age effects are limited to cognitive ToM (Ze et al., 2014) or apply equally to both components (Duval et al., 2011; Rakoczy et al., 2012). Most evidence favours the latter notion. For instance, Duval and colleagues (2011) demonstrated that adults aged 61 to 83 years performed worse on measures of cognitive *and* affective ToM relative to young adults (21 to 34 years) and middle-aged adults (45 to 59 years). No differences were observed between the young and middle-aged groups, suggesting age differences are detectable in both components and become meaningful after age 60 (see also Bernstein et al., 2011; Charlton, Barrick, Markus, & Morris, 2009; Pardini & Nichelli, 2009). In a recent meta-analysis Henry and colleagues (2013) examined performance trends across measures of cognitive and affective ToM commonly used in aging research. They found moderate effect sizes across both ToM components favouring young adults: cognitive ToM, r = -.45, 95% CI [-.37 to -.53] and affective ToM, r = -.51, 95% CI [-.45 to -.58].

In sum, converging evidence from cognitive neuroscience, neuropsychology, child development, and neuroimaging research supports a multidimensional organization within ToM. Convergent lines of inquiry demonstrate that despite important conceptual dissociations, cognitive and affective ToM share a common, integrated structure. At present, researchers are merely beginning to examine ToM in a multidimensional way to assess how mental state reasoning is supported across the adult lifespan. Thus, the aim of this dissertation was to develop a multidimensional framework through which we comprehensively addressed how three key predictors (neurocognitive functions, blood pressure, and biological sex) related to age differences in cognitive and affective ToM.

1.2. Predictors of Age Differences in ToM

1.2.1. Neurocognitive Functions

Advancing age is accompanied by considerable variability in neurocognitive performance (Wilson et al., 2002). While some individuals maintain their level of neurocognitive functioning with age, others demonstrate declines that significantly impact their daily performance (Plehn, Marcopulos, & McLain, 2004). Declines typically

occur in fluid abilities including executive functions, episodic memory, and processing speed, while crystallized verbal skills and semantic knowledge remain intact (Salthouse, 2010). As noted earlier, debate surrounds whether ToM inferences are formed via automatic, implicit processes (i.e., theory-theory or simulation theory; Carruthers & Smith, 1996; Gallese & Goldman, 1998), or if inferences are supported by on-line interaction and input from perceptual skills and requisite neurocognitive resources (Apperly et al., 2005; Stone & Gerrans, 2006). If ToM is implicit, both cognitive and affective components should be relatively resistant to age-related change and improve steadily with accumulated social experience. In this case we would expect that (a) ToM should be weakly associated with neurocognitive functions across the lifespan, and/or (b) associations with neurocognitive functions might vary based on task demands. The latter idea suggests that for measures whose methodology requires ToM and other neurocognitive skills, "pure" age-invariant cognitive and affective ToM constructs would if parsed from construct-irrelevant variance. This contrasts with the neuropsychological conceptualization of ToM as a componential construct whose function is dependent on on-line input from perceptual skills and supporting neurocognitive resources, such as inhibition of the self-perspective, attention and working memory, correct recall of mental state information, and verbal expression. This latter framework would entail that individual differences in age-sensitive neurocognitive functions are important predictors of cognitive and affective ToM in both age groups.

In line with a neuropsychological approach to the conceptualization of ToM, we framed our review to summarize research supporting neurocognitive resources as essential to ToM across the lifespan. However, several problems with existing work limit the strength of this argument and are areas we sought to address in this dissertation. For one, knowledge about the relation of ToM to neurocognitive skills is mostly limited to research examining single neurocognitive domains (e.g., executive functions; Duval et al., 2011; Rakoczy et al., 2012). Less is known about how neurocognitive functions influence ToM when predictors are considered concurrently. Existing research is also predominately focused on associations between neurocognitive functions and *cognitive ToM* (Sandoz et al., 2014). Findings are mixed regarding whether neurocognitive functions similarly predict affective ToM. Finally, few studies have examined neurocognitive variables while also considering potential non-cognitive predictors of ToM

(see Charlton et al., 2009; Fischer et al., 2014). Based on these issues and our conceptualization of ToM as a multidimensional construct, we considered three domains with the strongest evidence documenting associations with ToM in adulthood: Executive functions, crystallized verbal skills, and episodic memory. Because thorough reviews are available elsewhere (Sandoz et al., 2014), we summarize supporting evidence below. Given the lack of standardized, recommended ToM measures in aging research we detail the findings of this literature in Appendix C, Table C1.

Executive Functions

Clear and consistent evidence links ToM to multiple executive functions, including response inhibition, working memory, attention, mental flexibility, updating, and abstract reasoning (Ahmed & Miller, 2011; Li et al., 2013; Mahy et al., 2014b; McKinnon & Moscovitch, 2007; Phillips et al., 2011). Response inhibition is necessary to filter out competing mental states (e.g., the self-perspective) and other irrelevant information in order to adopt another's mental perspective (Bailey & Henry, 2008). Because of this it has received the most attention in ToM research. Past work indicates that inhibition directly predicts cognitive ToM (Ahmed & Miller, 2011; Wang & Su, 2013) and mediates associations between age and cognitive ToM (Bailey & Henry, 2008). We know of only two studies that have examined links between inhibition and affective ToM. These suggest that increasing the inhibitory demands within a measure leads to poorer affective ToM (Mahy et al., 2014b). The fact that poor inhibition leads to reduced performance on measures of cognitive and affective ToM is consistent with Hasher and Zacks' (1998) inhibitory deficit hypothesis, which suggests that age-related declines in cognitive performance are attributable to difficulty filtering task-irrelevant information with age. Yet poor inhibition is related to worse ToM in childhood, adolescence, and young adulthood—not just in late life (Carlson & Moses, 2001; Vetter et al., 2013). Reported associations between ToM and other executive functions are limited to cognitive ToM, wherein modest associations exist between poor cognitive ToM and reduced mental flexibility (Li et al., 2013), working memory (McKinnon & Moscovitch, 2007), updating (Phillips et al., 2011), and abstract reasoning (Ahmed & Miller, 2013). Other studies report no association between cognitive ToM and executive functions (e.g., Bernstein et al., 2011; Keightley et al., 2006; MacPherson et al., 2002). It is generally unknown how executive resources aside from inhibition relate to affective ToM. Some authors have

further suggested that the focus on individual executive predictors may be too specific and that it is the totality, rather than individuality, of executive dysfunction that leads to poor ToM (Aboulafia-Brakha, Christe, Martory, & Annoni, 2011). This idea is supported by strong direct associations between reduced performance on composite 'executive function' variables and cognitive ToM (Charlton et al., 2009; Duval et al., 2011; Fischer et al., 2014) and moderate indirect associations with affective ToM (Rakoczy et al., 2012; Yeh, 2013).

Crystallized Verbal Abilities

There is clear support for an association between crystallized verbal abilities and both cognitive *and* affective ToM, such that higher verbal skills predict better mental state reasoning (Peterson & Miller, 2012). Significant associations with ToM are reported across measures ranging from basic vocabulary tests (Maylor, Moulson, Muncer, & Taylor, 2002; Slessor, Bull, & Phillips, 2007) to verbal composites of standardized intellectual test batteries (e.g., WAIS-III VIQ: Charlton et al., 2009). Charlton and colleagues (2009) demonstrated that verbal abilities explained 14% of the variance in older adults' performance on a story-based test of cognitive ToM, beyond age differences in story comprehension and neural integrity. Verbal abilities also predict ToM in young adults, even when efforts are made to mitigate language demands (Baker, Peterson, Pulos, & Kirkland, 2014; Peterson & Miller, 2012). In a recent meta-analysis, Peterson and Miller (2012) found that the association between verbal skills and affective ToM (r = 0.49; moderate) was considerably larger than the association between verbal abilities and cognitive ToM (r = 0.29; low). Regardless of differences in magnitude, it is apparent that crystallized verbal skills are closely linked to ToM across contexts; yet to date few studies have examined their influence alongside other potential predictors.

Episodic Memory

While changes in episodic memory are a hallmark feature of cognitive aging, support for an association between memory and ToM is equivocal. Several studies report links between episodic memory and cognitive ToM; however, it is unknown whether this relationship extends to affective ToM (Sandoz et al., 2014). Research is limited by a restricted focus on correlational approaches where ToM is measured using

tests that assess individuals' inferences about details recalled from short stories (Baglio et al. 2012; Castelli et al., 2011; Fischer et al, 2014; Maylor et al., 2002). This suggests that episodic memory involvement may be partly an artefact of the methodology used in story-based ToM paradigms. Studies of episodic memory and ToM also have failed to account for other potential explanatory neurocognitive variables. This is a significant drawback of the ToM literature and signifies that multi-predictor investigations of ToM are needed to elucidate if memory is a relevant predictor of age differences in cognitive and affective ToM.

As outlined in this review and in Appendix C, there is strong evidence supporting neurocognitive performance as a key predictor of cognitive and affective ToM. At present, firm conclusions regarding the relative importance of individual neurocognitive domains are obscured by the array of ToM tests used and researchers' tendency to isolate single predictors. In this dissertation we investigated the unique and shared contributions of executive functions, crystallized verbal abilities, and episodic memory to age differences in cognitive and affective ToM. Because recent investigations report that neurocognitive functions account for less than 20-30% of the variance in both cognitive and affective ToM performance (e.g., Bernstein et al., 2011; Fischer et al., 2014) we also investigated two categories of non-cognitive variables with the potential to explain additional variance: Blood pressure and biological sex.

1.2.2. Blood Pressure

Hypertension, or high blood pressure (defined as systolic blood pressure ≥ 140 mmHg and/or diastolic blood pressure ≥ 90 mmHg; Nichols, O'Rourke, & Vlachopoulos, 2011), is a major public health concern due to its high prevalence and relation to increased risk of disease and disability (Kearney et al., 2005). Blood pressure is measured using two primary indicators: Systolic blood pressure (SBP) and diastolic blood pressure (DBP). Differentiating SBP and DBP is important due to their varied prognostic significance in older age: After age 60, elevated SBP is accompanied by falling or stabilized DBP (Khattar, Swales, Dore, Senior, & Lahiri, 2001). High blood pressure is the most important modifiable risk factor for cardiovascular disease and stroke and is a leading contributor to mortality worldwide (World Health Organization, 2009). As of 2007, nearly one out of every five Canadians aged 20 to 79 was diagnosed

with hypertension and many others thought to be unaware of their condition (Wilkins et al., 2010). Although hypertension is a lifespan disease, prevalence rates rise substantially in later life. The Public Health Agency of Canada (2010) estimated that crude rates for diagnosed hypertension range from 43.4% in adults aged 60 to 64 to 71.3% in adults aged 85+, with slightly higher rates in women. Hypertension is also highly comorbid with other chronic illness. Individuals with poorly controlled blood pressure are more likely to experience type 2 diabetes, dyslipidemia, obesity, and renal disease (Chobanian et al., 2003), all which are independently linked to neurocognitive decline (Elias, Elias, Sullivan, Wolf, & D'Agostino, 2003; Thornton et al., 2007; Yaffe, Barrett-Connor, Lin, & Grady, 2002; Yeung, Fischer, & Dixon, 2009). Epidemiological research suggests that individuals with uncontrolled blood pressure have an elevated risk of AD and vascular cognitive impairment compared to their normotensive peers (Gorelick et al., 2011; Nagai, Hoshide, & Kario, 2010).

Uncontrolled blood pressure is also associated with neurocognitive difficulties prior to the onset of stroke or dementia (Birns & Kalra, 2009). Understanding how blood pressure impacts neurocognitive functions is critical to mitigating clinical consequences such as reduced quality of life and loss of independence. Recent evidence suggests that elevated SBP/DBP *and* low SBP/DBP may be similarly detrimental to neurocognitive health (e.g., Birns & Kalra, 2009; Qiu, Winblad, & Fratiglioni, 2005; Yeung & Thornton, 2011). Blood pressure-related neurocognitive difficulties are often subtle but can occur across multiple domains. Gifford and colleagues' (2013) meta-analysis of eight cross-sectional studies indicated that high blood pressure predicted worse performance on cognitive screening tests (r = -.11, 95% CI [-.18 to -.04]) and worse episodic memory (r = -.20, 95% CI [-.28 to -.12]), whereas low blood pressure predicted worse attention (r = .14, 95% CI [.03 to .25]). Links between high SBP, high DBP, and executive functions are also commonly reported (Birns & Kalra, 2009). What is striking about the findings of Gifford and colleagues is that associations remained significant after adjusting for age, education, and comorbid cardiovascular risk factors (CVRFs). The small to moderate effects reported by these authors are consistent other work (Qiu et al., 2005; Waldstein, 2003). Another key point is that the neurocognitive domains impacted by blood pressure overlap with those linked to poor ToM, which suggests that variable blood pressure may similarly be associated with altered ToM performance (Fischer et al., 2014).

Because of the risk jointly conferred by high SBP and low DBP, many researchers advocate that a third indicator, pulse pressure (PP), is a superior marker for current and incipient cognitive decline (Yasar, Ko, Nothelle, Mielke, & Carlson, 2011). Calculated as SBP minus DBP, PP rises sharply with advancing age, thereby capturing the divergent influences of SBP and DBP (Franklin, 2004). PP is considered a proxy measure for arterial stiffening, an age-related loss of elasticity caused by structural and cellular changes in large central arteries (Nichols et al., 2011). Independent of SBP, DBP, and other CVRFs, high PP (> 60 mmHg) predicts cardiovascular morbidity and all-cause mortality (Chobanian et al., 2003). High PP also predicts worse neurocognitive performance in healthy older adults on cognitive screening and across individual domains, including ToM (Fischer et al., 2014; Singer, Trollor, Baune, Sachdev, & Smith, 2014), and is associated with increased risk of future neurocognitive decline (McFall et al., 2014). In the Women's Health and Aging Study II, women aged 70 to 80 years with PP ≥ 84 mmHg at baseline had five times greater incidence of impaired executive functions and episodic memory over the subsequent 9-years than women with lower baseline PP (Yasar et al., 2011). Across studies, PP shows the strongest associations with executive functions and processing speed (McFall et al., 2014).

Earlier we reported initial evidence that high PP is indirectly associated with lower cognitive ToM (Fischer et al., 2014). This suggests that blood pressure effects may be more consequential than previously acknowledged and extend to social-emotional functioning. In our previous research (Fischer et al., 2014) we examined PP as a modifier of cognitive ToM in 66 community-residing older adults. Reduced ToM was associated with older age, poor episodic memory, slowed processing speed, and poor working memory; however, associations between ToM and working memory were attenuated when PP was added as a predictor. A novel finding of this work was that PP modified ToM performance. Specifically, older adults with elevated PP showed the strongest associations between reduced ToM and lower episodic memory/speed. We also observed a direct association between high PP and worse ToM that did not meet conventional guidelines for statistical significance within the relatively small sample (r = −.23, p = .06). These findings draw attention to blood pressure as a potentially important contributor to age differences in ToM performance. Thus, in this dissertation we sought

to determine whether PP was a key predictor of cognitive and affective ToM alongside neurocognitive performance and biological sex.

1.2.3. Sex Differences

Sex differences in ToM performance are observed early in childhood, such that girls tend to acquire and master ToM concepts at a younger age than boys (Charman, Ruffman, & Clements, 2002). In addition, epidemiological research shows that clinical disorders characterized by core deficiencies in ToM are more frequently diagnosed among boys (DSM-IV male to female diagnosis ratio: 4 to 1 in autism and 11 to 1 in Asperger Syndrome; Gillberg, Cederlund, Lamberg, & Zeijlon, 2006). It is important to note that evidence linking biological sex to ToM variability pertains to the acquisition and use of affective ToM—the relationship between sex and cognitive ToM is largely unstudied. Research supporting a female advantage on measures of affective ToM is consistent across childhood, adolescence, and young adulthood, with some authors suggesting this relationship may be mediated by empathy (Ibanez et al., 2013). Kirkland, Peterson, Baker, Miller, & Pulos (2013) reported a significant female advantage in affective ToM performance across 40 studies of typically developing adults. The mean weighted effect size was small and not moderated by language of test administration, country of study, or research group (i.e., Hedge's g = .18, 95% CI [.12 to .24]), suggesting that sex differences for affective ToM were robust. These behavioural findings are further corroborated by neuroimaging research showing that women utilize additional brain regions underlying emotion and self-referential thinking when making ToM inferences as compared to same-aged men (Christov-Moore et al., 2014).

In aging research, the two studies that have examined sex differences on false-belief measures of cognitive ToM report null associations (Franco & Smith, 2013; Sullivan & Ruffman, 2004). To our knowledge, no research has specifically examined associations between sex and affective ToM in later life. Decisions to exclude sex as a predictor may be influenced by the small sample sizes that are typical in aging research; however, many aging studies report disproportionate enrolment of women. Key questions remain regarding the role of sex differences in explaining age differences ToM, and whether or not it is critical process when neurocognitive performance and health status are considered concurrently.

1.3. The Current Study

Convergent evidence supports ToM as a complex, multidimensional construct with overlapping, but non-identical cognitive and affective components. In the preceding review we emphasized that possible factors supporting ToM are diverse and involve inter-related neurocognitive, health, and demographic variables. To date research is only beginning to examine ToM within a multidimensional framework and assess how external variables support and maintain ToM across the lifespan. Thus, we sought to expand directly on the findings of Fischer and colleagues (2014) to address three primary gaps in research: (1) Using four common measures, we examined the breadth and magnitude of age differences across cognitive and affective ToM components; (2) We explored the multidimensional structure of ToM by identifying the neurocognitive, vascular health, and biological sex variables that are fundamental to supporting ToM in young and older adults; (3) By comparing the strength of association between predictors shared between age groups, and by identifying non-shared predictors, we determined whether age differences existed in predictors of cognitive and affective ToM.

Chapter 2. Objectives and Hypotheses

2.1. Age Differences in Theory of Mind

Our first objective was to clarify the presence and magnitude of age differences in ToM. We extended past research by considering multiple measures of cognitive and affective ToM. We predicted that older adults would show poorer cognitive ToM and poorer affective ToM than young adults. Based on previous studies (e.g., Henry et al., 2013), we expected medium to large effect size differences by age between groups.

2.2. Identifying Predictors of Theory of Mind

Our second objective was to identify predictors of cognitive and affective ToM. We expected that some predictors would be consistent across age groups, while others would apply only to older adults. Specific hypotheses are outlined in the sections below and in Figure 2.1.

Shared Predictors *(solid arrows in Figure 2.1)*

First, given strong evidence linking executive functions to ToM in young and older adults (e.g., Ahmed & Miller, 2011; Carlson & Moses, 2001; Mahy et al., 2014b), we expected that, for both age groups, better executive functions would predict better cognitive and affective ToM. Second, given ample research linking crystallized verbal abilities to ToM (e.g., Baker et al., 2014; Peterson & Miller, 2012), we expected that, for both age groups, better verbal comprehension would predict better cognitive and affective ToM. Third, because published associations with memory are limited to cognitive ToM (Baglio et al. 2012; Fischer et al, 2014), we expected that, for both age groups, better episodic memory would predict better cognitive ToM. Fourth, considering clear evidence for between sex differences on tests of affective ToM (but not cognitive

ToM: Kirkland et al., 2013), we expected that, for both age groups, women would exhibit better affective ToM. Fifth, given research suggesting that cognitive ToM is a developmental precursor for affective ToM (Shamay-Tsoory et al., 2009), we expected a moderate positive association between cognitive ToM and affective ToM across groups.

Predictors in Older Age *(dashed arrows in Figure 2.1)*

Given extensive literature linking uncontrolled blood pressure to neurocognitive difficulties (Birns & Kalra, 2009; Gifford et al., 2013), we predicted that high PP would predict worse executive functions and worse episodic memory. Expanding on our previous research (Fischer et al., 2014) we expected that high PP would also predict poorer cognitive and affective ToM. We anticipated these relationships for the older sample only given the greater prognostic significance of PP in later life and the restricted range of PP that is typically observed in early adulthood (Yasar et al., 2011).

2.3. Age Differences in Predictors of Theory of Mind

Our third objective was to identify age differences in predictors of cognitive and affective ToM. Considering strong evidence for associations between neurocognitive performance and ToM in advancing age (Sandoz et al., 2014), we expected that the strength of associations between (a) executive functions and cognitive ToM, (b) executive functions and affective ToM, and (c) episodic memory and cognitive ToM would be significantly greater for older adults than for young adults.

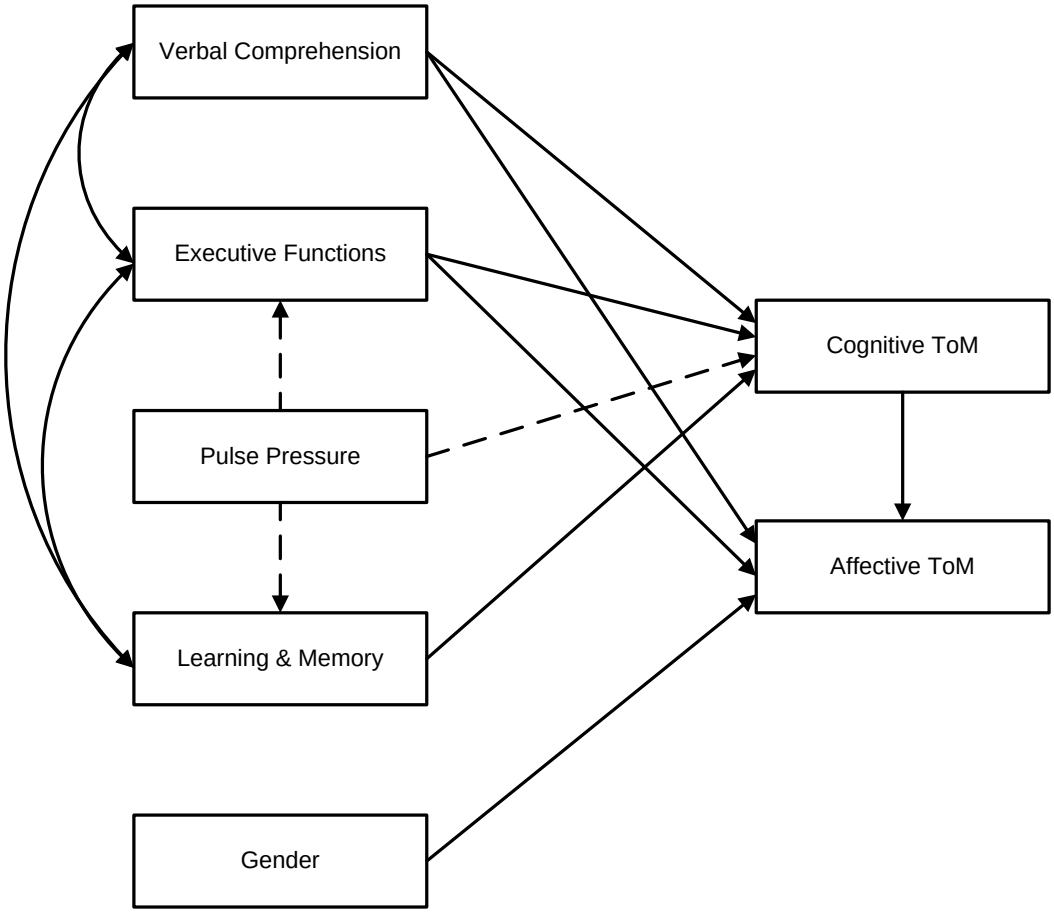


Figure 2.1 Hypothesized path model for predictors of cognitive ToM and affective ToM.

Note. Rectangles depict observed variables. Solid single-headed arrows represent the hypothesized direction of relationships we predicted for both age groups. Solid double-headed arrows indicate shared error covariances. The single-headed dashed arrows represent the hypothesized direction of non-shared associations we predicted for older adults. We present hypotheses for free parameters only (i.e., those predictor relationships that are hypothesized and estimated in the path model; indicated by arrows in Figure 2.1). We did not make predictions regarding the fixed parameters given a lack of supporting theoretical rationale for associations between these variables (i.e., possible relationships between variables that are not represented by arrows in Figure 2.1).

Chapter 3. Methods

3.1. Participants

We recruited two independent samples of adults living in metro Vancouver. The older adult sample comprised 90 adults 60+ years of age through flyers posted at community recreation facilities, free online volunteer postings, and advertisements in local community newspapers. We recruited this broad range given increased rates of cardiovascular illness in this cohort (Public Health Agency of Canada, 2010) and to ensure representative sampling of the community-residing aging population. The young adult sample comprised 93 undergraduate students aged 17 to 30 enrolled in introductory psychology courses at Simon Fraser University (SFU). All participants met the following inclusion criteria: (a) independently provided informed consent, (b) no impairments in vision, hearing, or other sensory/motor functions that interfered with testing, and (c) a minimum Grade 6 education to ensure reading level adequate for the questionnaire and neurocognitive protocol. We screened English fluency using an acculturation measure that examined language preferences for speaking, thinking, reading, and writing. To be eligible, participants needed to indicate English as their preference for at least three of these categories (see also Fischer et al., 2014; Paterson, O'Rourke, Elmer, Shapiro, & Thornton, 2011; Thornton et al. 2007).

In addition, exclusion criteria included: (a) diagnosis of dementia or MCI by a physician and/or a score of less than 26 on the Mini-Mental State Examination (MMSE; Folstein, Folstein, & McHugh, 1975; Roper, Bieliauskas, & Peterson, 1996); (b) current significant psychotic illness; (c) current illness or organ failure known to affect the central nervous system (e.g., kidney failure; Parkinson's disease, Huntington's disease, epilepsy); (d) history of significant stroke; (e) history of significant head injury (defined by a loss of consciousness > 15 minutes); or (f) alcohol consumption of greater than 3 drinks daily. We established these criteria based on past research documenting these variables to negatively influence neurocognitive performance and/or attenuate effects

attributable to blood pressure. We screened for visual acuity with a set lower limit of 20/70 in both eyes (corrected or non-corrected).

Prior to data analysis, we excluded five participants (4 young, 1 older) following examiners' notes documenting low or inconsistent effort during neurocognitive testing. Based on initial data screening (Cook's D, standardized DFBETAS, scatterplots of externally studentized residuals against centred leverage values), we excluded seven participants who were extreme multivariate outliers across the set of dependent variables (3 young, 4 older; see Appendix D). Thus, our final sample comprised a total of 171 participants (86 young adults, 85 older adults). Appendix D, Figure D1 presents detailed information about participant recruitment and screening strategies.

3.2. Materials and Assessment Procedure

After obtaining informed consent, we tested all participants individually on a two and one-half hour battery that assessed standard neurocognitive functions and ToM. Trained graduate students and an undergraduate honour's student conducted all of the testing under the supervision of Dr. Wendy Thornton. Testing was conducted in quiet rooms at the SFU Burnaby or SFU Surrey campuses, dependent on participants' travel preferences. Prior to the testing session, participants completed questionnaires assessing background demographics, medical history, and self-ratings of current depressive and anxiety symptoms. We standardized the administration order of our test battery: At the start of each session, prior to any neurocognitive testing, we measured participants' resting blood pressure. We then administered the various neurocognitive and ToM measures. Young adults (students) received course credit for their participation, and older adults received a $20 honorarium. All protocol included in this study was pre-approved by the SFU Research Ethics Board.

3.2.1. Blood Pressure Protocol

At the start of the testing session, we obtained four blood pressure readings for each participant on the non-dominant arm unless medically contraindicated. We measured blood pressure using an automatic oscillometric upper arm monitor validated

by the British Hypertension Society (Microlife BP 3AC1-1PC). Participants sat quietly with their feet flat on the floor. We took an initial reading to ensure comfort and familiarity with the procedure. After a five-minute rest break, we took three additional readings separated by one-minute rest intervals. This procedure adheres to published standards for in-office research blood pressure assessment (Campbell, Joffres, & McKay, 2005). As an outcome measure, we calculated the average PP (SBP – DBP) over the final three readings for each participant.

3.2.2. Questionnaire Protocol

Demographics and Health

We administered a self-report questionnaire addressing participant demographics and history of medical illness and treatment. We were specifically interested in obtaining information about current diagnoses of hypertension and other CVRFs (i.e., type 2 diabetes, dyslipidemia, cardiovascular disease), and where possible, verification of diagnosis by prescription treatment and/or lifestyle interventions. We requested that participants bring in the pill bottles or pharmacy receipts of any current medications for verification purposes.

Depressive and Anxiety Symptoms

We used the Center for Epidemiologic Studies Depression Scale (CES-D; Radloff, 1977) to assess self-reported depressive symptoms. The CES-D consists of 20 items that assess depressive affect, somatic symptoms, self-ratings of wellbeing, and interpersonal adjustment over the preceding one-week. Participants rated each item on a 4-point scale according to how often a symptom was experienced: 0 = *rarely or none of the time* to 3 = *most or all of the time*. Raw scores greater than or equal to 16/60 are suggested to demarcate clinically significant depressive symptoms (Radloff, 1977). Items comprising the CES-D have high internal consistency reliability in community-residing Canadian adults aged 19 to 91 (α = .87; Paterson et al., 2011).

We used the Multidimensional Anxiety Questionnaire (MAQ; Reynolds, 1999) to assess self-reported anxiety symptoms. The MAQ consists of 40 items that assess

physiological symptoms of anxiety, social withdrawal, feelings of worry and/or fear, and negative affect over the preceding one-month. Participants rated each question on a 4-point scale from 1 = *almost never* to 4 = *almost all of the time*. Raw scores range from 40 to 160 and can be converted to T-scores and compared to normative community and college-based samples to indicate severity of anxiety. T-scores of 64 to 70 indicate moderate clinical anxiety and T-scores greater than 71 indicate severe clinical anxiety. Items comprising the MAQ have very high internal consistency (α = .96) in both community and clinical samples (Reynolds, 1999).

3.2.3. Neurocognitive Protocol

Executive Functions and Attention

We used the Color-Word Interference subtest from the Delis-Kaplan Executive Function System to assess response inhibition (DKEFS CW; Delis, Kaplan, & Kramer, 2001). The baseline conditions of this measure required participants to name the colours of printed ink squares (C1) and read printed words (C2) as quickly as possible. In the inhibition condition (C3), participants viewed a page of colour words printed in discordantly coloured ink. They were instructed to name the colour of each printed word while supressing their automatic response to read the word itself. We recorded participants' latency to perform each condition in seconds. As recommended by Delis and colleagues (2001), we subtracted the latencies of C3 - C1 to obtain a non-speeded measure of response inhibition. We reversed-coded scores so that higher scores indicated better performance. DKEFS CW items have adequate internal consistency reliability (α = .75; Delis et al., 2001), though it is suggested that contrast score reliability may be weaker among older adults (e.g., α = .70 for adults aged 20 to 49 years, and α = .21 in adults aged 50 to 89 years: Crawford, Sutherland, & Garthwaite, 2008).

We also used the Letter-Number Sequencing subtest from the Wechsler Adult Intelligence Scale – 3rd Edition to measure working memory (WAIS-III LN; Wechsler, 1997). This measure required participants to listen to sequences of numbers and letters and recall each sequence in alphanumeric order. We used the number of sequences correctly recalled as our outcome measure of working memory. Items comprising the WAIS-III LN subtest demonstrate adequate test-retest reliability in adults aged 52 to 80

years (two-week interval: r = .75 and four-week interval r = .74; Lemay, Bédard, Rouleau, & Tremblay, 2004).

To measure basic auditory attention, we used the WAIS-III Backward Digit Span subtest (WAIS-III DS; Wechsler, 1997). This measure required participants to listen to sequences of numbers and recall each sequence in the reverse order that it was heard. We used the number of sequences correctly recalled as our outcome measure of auditory attention. Items from the WAIS-III DS subtest demonstrate high internal consistency across clinical populations (α = .92; Zhu, Tulsky, Price, & Chen, 2001)[4].

Verbal Comprehension

We used the Kaufman Brief Intelligence Test – 2nd Edition Verbal Knowledge subtest (KBIT-2 VK; Kaufman & Kaufman, 2004) to measure receptive verbal skills. With this measure, participants viewed pages containing six pictures each. Examiners read one word per page, and participants indicated which picture best represented the meaning of each target word. We used the number of correctly identified target words as our outcome measure of verbal comprehension. Items comprising the KBIT-2 Verbal Scale demonstrate high internal consistency reliability (α = .91). Similarly, test-retest reliability (four-week interval) is excellent for adults aged 22 to 59 years (α = .89) and those aged 60 to 89 years (α = .92) (Kaufman & Kaufman, 2004)[5].

Episodic Memory

We used the California Verbal Learning Test–II to measure verbal episodic memory (CVLT-II; Delis, Kramer, Kaplan, & Ober, 2000). Participants learned 16 words over five trials. Recall was assessed after each learning trial, after a very brief delay, and again after a 20-minute delay. We used the following scores as our outcome measures of episodic memory: The total items learned over trials 1-5 (learning), the total items recalled at a short delay free recall trial (short-term verbal retention), and the total items recalled at a long delay free recall trial (long-term verbal retention). Items comprising the

[4] Includes both Forward (not used in this dissertation) and Backward Digit Span subtests.

[5] Includes both Verbal Knowledge and Riddles (not used in this dissertation) subtests.

CVLT-II demonstrate good internal consistency in healthy adults aged 16 to 88 years (.80 < α < .84; Woods, Delis, Scott, Kramer, & Holdnack, 2006).

3.2.4. Theory of Mind Protocol

Strange Stories Test

We used the Strange Stories test to assess participants' understanding of advanced cognitive ToM concepts (Happé et al., 1998; hereafter referred to as *STORIES*). This test was originally developed for use in research among individuals with autism spectrum disorder (Happé, 1994), but is now one of the most commonly used measures to assess ToM in aging (Henry et al., 2013). Consistent with procedures outlined in Fischer et al., 2014, we selected eight ToM stories from those published by Happé and colleagues (1998) based on aging-relevant content. The stories required participants to reason about a variety of meta-cognitive mental states (e.g., decide whether a character was lying, infer the intentions of a character playing on another character's sympathy). Participants viewed the stories alongside black and white drawings depicting significant story details. We encouraged participants to take the reading time necessary to ensure complete understanding. For each story, once participants indicated that they had read and understood its content, we asked one critical question to assess mental state reasoning: *Why did [the character] say/do that?* We recorded responses verbatim and scored them according to criteria published by Happé and colleagues (1998): 2 = complete and accurate response, 1 = partial or implied response and 0 = incorrect or irrelevant response. When more than one answer was provided we credited the most accurate response. Similarly, if a response contained both mental state and non–mental state inferences, it was scored for the mental state. We used the total summed score as an outcome measure of cognitive ToM. We provide examples of the stories used and associated response criteria in Appendix E, Table E1.

Reading the Mind in the Eyes Test

We used the revised version of the RMET to assess participants' understanding of advanced affective ToM concepts (Baron-Cohen et al., 2001). The RMET was developed for use in autism research but is now routinely used to assess variability in

affective ToM in both child developmental and adult research (Kirkland et al., 2013). Participants viewed 36 black-and-white photographs of the eye region of human faces. Each pair of eyes was standardized for size (15 cm x 6 cm) and portion of the face that was shown (top of eyebrows to midway down the ridge of the nose). Participants selected from four possible descriptors the word they felt best represented the emotional mental state depicted in each set of eyes. We provided participants with a glossary of all descriptors and encouraged them to reference this at any point throughout the test. Participants made their responses by circling their choice of descriptor on a response sheet. We allowed participants to take as much time as needed to complete the test. We scored responses as correct (1-point) and incorrect (0-points) and used the total summed score as an outcome measure of affective ToM. We provide examples of stimuli used in the RMET in Appendix E, Figure E1.

Yoni Test

We used the Yoni Test (Shamay-Tsoory & Aharon-Peretz, 2007) as a third measure of simple and advanced ToM concepts. This measure contains nested cognitive and affective ToM test conditions—that is, the two conditions are completed within a single administration and are standardized to differ only by cognitive and affective content. The Yoni Test is used widely in neuroimaging research with young and older adults (e.g., Schlaffke et al., 2015; Shamay-Tsoory & Aharon-Peretz, 2007) and more recently in several neuropsychological studies assessing ToM in clinical populations (e.g., AD and bvFTD: Narme et al., 2013a; PD: Narme et al., 2013b). Protocol and instructions for the Yoni Test are publically available at http://sans.haifa.ac.il/downloads.html.

We adapted the original protocol for computer administration using E-Prime 2.0 software (Schneider, Eschman, & Zuccolotto, 2002) and to reflect content specific to our primary objectives. Specifically, we modified four of the original trials (two cognitive, two affective) as teaching trials to facilitate participants' familiarity with the task format. We also removed an additional 12 trials whose content reflected competitive emotions such as "gloat" and "envy," thereby evaluating participants' knowledge of the consequences of another person's fortune, rather than pure reasoning about cognitive and affective mental states (see Shamay-Tsoory, Tibi-Elhanany, & Aharon-Peretz, 2007).

After these modifications, the Yoni Test version used in this study comprised 34 trials that assessed first- and second-order cognitive ToM (*YONI C-TOM*) and 33 trials assessing first- and second-order affective ToM (*YONI A-TOM*). In each trial, participants viewed a character named Yoni who was surrounded by four coloured pictures belonging to a single category (e.g., fruits, chairs, transportation, faces). Participants selected which picture Yoni was referring to in his mental state, based on information conveyed in simple sentences and facial cues such as eye gaze or mouth expression. For a subset of trials, Yoni's eye gaze was directed straight ahead (~35% trials) or toward an incorrect target (~5% trials), to ensure that participants understood the task and did not respond solely to gaze direction without considering other cues. For the cognitive ToM trials, all cues were emotionally neutral, whereas for the affective ToM trials, facial and verbal cues comprised positive and negative valences. Participants were allowed as much time as needed to complete the measure. We scored responses as correct (1-point) versus incorrect (0-points) and used the total summed scores for YONI C-TOM and YONI A-TOM as our outcome measures representing cognitive and affective ToM. We provide examples of Yoni Test protocol in Appendix E, Figure E2.

Table 3.1 Summary of Neurocognitive Measures by Conceptual Domain

Domain/subdomain	Measure(s)	Acronym
Executive Functions & Attention		
Response inhibition	DKEFS Color-Word Contrast Score	*DKEFS CW*
Working memory	WAIS-III Letter Number Sequencing	*WAIS LN*
Basic auditory attention	WAIS-III Backwards Digit Span	*WAIS DS*
Crystallized Verbal Abilities		
Verbal comprehension	KBIT-2 Verbal Knowledge	*KBIT VK*
Episodic Memory		
Learning	CVLT-II Trials 1-5	*CVLT 15*
Short-term verbal retention	CVLT-II Short Delay Free Recall	*CVLT SDFR*
Long-term verbal retention	CVLT-II Long Delay Free Recall	*CVLT LDFR*
Theory of Mind		
Cognitive ToM	Strange Stories	*STORIES*
	Yoni Test – Cognitive ToM trials	*YONI C-TOM*
Affective ToM	Reading the Mind in the Eyes Test	*RMET*
	Yoni Test – Affective ToM trials	*YONI A-TOM*

Chapter 4. Analytic Strategy

4.1. Data Preparation

We inspected the data for fit between the distributions of variables of interest and the assumptions of path analysis[6] (see Appendix D). Rates of missing data were negligible: Total missing data equalled seven out of approximately 3600 cells (or ~ .001%). We imputed these cells by linear interpolation using ordinary least squares regression to predict the missing values from the last complete observation prior to the missing data and the first complete observation after the missing data. We conducted analyses on the imputed data with a final N = 171. For all analyses we used SPSS 22.0 (IBM Corp, 2013) and AMOS 21.0 software (IBM Corp, 2012).

4.2. Sample Characterization & Initial Analyses

Prior to addressing our primary objectives, we examined group differences in participant health and demographic characteristics. To characterize the sample we examined the following variables known to influence blood pressure and/or neurocognitive performance: CVRFs commonly comorbid with hypertension (type 2

[6] Path analysis holds a number of requirements concerning the nature of the data and the theoretical model itself. As summarized in O'Rourke & Hatcher (2013), assumptions concerning the nature of the data require that: (a) predicted variables should be assessed on interval or ratio levels of measurement; (b) predicted variables should be continuous and assume a minimum of five values; (c) data should assume a multivariate normal distribution and any deviations from normality should be addressed using transformation or the deletion of outliers; (d) variables should be free of multi-collinearity; and (e) variables should be measured without error. Note that condition (e) cannot be assumed for many of our neurocognitive tests. To control for this, we modeled error terms with each predicted variable as recommended by O'Rourke & Hatcher (2013) and Byrne (2013). Assumptions concerning the theoretical model include: (a) relationships among variables should be linear and additive; (b) all known and non-trivial predictors should be included in the model as independent variables; and (c) the model must be over-identified (i.e., contains more correlations than predictor variables).

diabetes, dyslipidemia, cardiovascular disease: Anderson, Odell, Wilson, & Kannel, 1991; Elias et al., 2003; Yaffe et al., 2002; Yeung et al., 2009); body mass index (Gunstad et al., 2007); alcohol consumption (Sesso, Cook, Buring, Manson, & Gaziano, 2008); use of anti-hypertensive medication (Gorelick et al., 2011); educational attainment and English as an additional language status (Heaton, Ryan, & Grant, 2009); depressive and anxiety symptoms (Castaneda, Tuulio-Henriksson, Marttunen, Suvisaari, & Lönnqvist, 2008). We analyzed age group differences and effect sizes for these variables and our neurocognitive variables using chi-squared tests/coefficient phi for categorical variables (small ES $\Phi \leq .10$; medium ES $\Phi \geq .30$; large ES $\Phi \geq .50$) and independent samples *t*-tests/Cohen's *d* for continuous variables (small ES $d \leq .20$; medium ES $d \geq .50$; large ES $d \geq .80$; very large ES $d \geq 1.30$; Cohen, 1988).

Given very little published data on psychometric properties of ToM tests we examined score reliability for measures used in this dissertation. To ensure that individual items reflected the same construct as their respective scale scores, we deleted those items with low response variability or negative/very low item-total correlations (i.e., $r < .10$; Meyers, Gamst, & Guarino, 2013). Once these items were removed we assessed the reliability of the remaining items for each measure. For the RMET, YONI C-TOM, and YONI A-TOM we computed intraclass correlation coefficients (ICC) to determine internal consistency. For the STORIES we computed ICC estimates to determine the level of agreement between the three raters who independently scored participants' responses. For all estimates, we computed ICC coefficients using a two-way mixed model for absolute agreement (Shrout & Fleiss, 1979). Finally, to reduce the number of independent variables in the path models in a meaningful way, we created composite variables by summing the *z*-scores of the ToM and neurocognitive variables, respectively, based on theoretical dissociations among constructs in neuropsychological and cognitive aging literature.

4.3. Primary Research Objectives

First, to evaluate the presence and magnitude of age differences in cognitive and affective ToM, we used independent samples *t*-tests to identify age differences in

performance across the individual and composite ToM variables. Effect sizes for these contrasts are reported as Cohen's *d* (Cohen, 1988).

Second, to identify predictors of cognitive and affective ToM in each age group, we used path analysis to examine the fit of the hypothesized model in Figure 2.1 to the young and older data, using separate models for each age group (Byrne, 2013). We also considered whether alternative models resulted in a better fit (i.e., the addition or deletion of a predictor or path from Figure 2.1). We made decisions about alternative models by evaluating the size and direction of the expected change (EC) statistics and modification indices (MI) (Saris, Satorra, & Sörbom, 1987). As per the recommendations of Byrne (2013), our final models were those that maximized the following criteria: (1) overall fit to the data; (2) meaningfulness within the ToM literature; (3) theoretical relevance; (4) parsimony. Once a good fitting model was identified we interpreted the beta weights for the individual predictors of ToM.

We evaluated model fit using three goodness-of-fit statistics based on the recommendations of Byrne (2013) and O'Rourke and Hatcher (2013): (1) the Comparative Fit Index (CFI) to evaluate incremental fit (i.e., the extent to which the hypothesized model fits the *sample* data); (2) the Root Mean Square Error of Approximation (RMSEA) and its associated confidence intervals (CI) to evaluate parsimony fit (i.e., the extent to which the hypothesized model fits the *population* data); (3) the Standardized Root Mean Square Residual (SRMR) to assess absolute fit (i.e., comparison between the fit of the hypothesized model versus no model). For CFI, values greater than .95 indicate good fit between the model and sample data (Byrne, 2013). For RMSEA and SRMR, values below .055 denote "good" model fit and less than .08 are considered adequate (Browne & Cudeck 1993; O'Rourke & Hatcher, 2013).

Finally, we used partial invariance analyses to test the equivalence of any predictors of ToM that were nested between age groups (i.e., group differences in predictors; Byrne, 2004, 2013). We first tested the fit of the final models for each age group in a single multi-group analysis. Results from this step reflected how the hypothesized predictors fit the data with no cross-group constraints imposed and served as a baseline fit value against which we compared all subsequent models. We then assessed whether any shared predictors were invariant (i.e., equivalent across groups)

by testing a hierarchy of nested models and specifying that the beta weights of predictors be held equal across models for each age group. We tested the models in an increasingly restrictive fashion wherein we cumulatively held in place the constraints for invariant predictors, thereby providing a rigorous test of cross-group equality (Byrne, 2004). We evaluated incremental fit by comparing the change in chi-squared values ($\Delta\chi^2$) between each model and the baseline fit to our a priori criteria of $\alpha \leq .05$ for each change in the degrees of freedom. For any model, a significant $\Delta\chi^2$ suggested a statistically meaningful age group difference in the regression path under question (Byrne, 2004, 2013).

Table 4.1 Interpretative Guidelines for Goodness-of-Fit Statistics

Fit Statistic	Interpretation
CFI	Values > .95 indicate good model fit
SRMR	Values < .055 indicate good model fit
RSMEA	Values < .055 indicate good model fit

Chapter 5. Results

5.1. Initial Analyses

5.1.1. Sample Characteristics

In Table 5.1 we present demographic and health data by age group. The samples were equivalent in sex distribution, but differed in other demographic variables. Similar proportions of young and older participants were born in North America; however, young adults had more ethnically diverse backgrounds (Φ = .55; large ES) and were more likely to report speaking English as an additional language (EAL) (Φ = -.20; medium ES). Young adults endorsed greater symptoms of depression (mean = 11.87, SD = 7.96, *d* = .70; medium ES) and anxiety (mean = 66.45, SD = 14.39, *d* = .97; large ES). Totals of 26.7% of young and 10.5% of older adults met the cut-off for clinical depression (16/60; Radloff, 1977: χ^2 = 7.34 [*df* = 1], *p* < .01, Φ = -.20; medium ES). Similarly, 12.7% of young and 8.2% of older adults displayed at least moderate clinical anxiety (T-score ≤ 64; Reynolds, 1999: χ^2 = .94 [*df* = 1], *p* = .33; Φ = -.07; small ES). Higher emotional distress among university samples is often reported and may reflect different situational stressors or better emotional regulation in older age (Jorm, 2000).

The older sample was cognitively healthy, as evidenced by clinically non-significant mean MMSE scores (i.e., mean = 29/30; see Roper et al., 1996). As expected with advancing age, older adults had greater diagnosed CVRFs (|.22 < Φ < .44|; medium ES). Also as expected, older adults also had higher rates of diagnosed hypertension (Φ = .48) and higher average SBP, DBP, and PP (|.83 < *d* < 1.33|; large to very large ES). These patterns are consistent with published linear trends for worsened cardiovascular health with advancing age (e.g., Chobanian et al., 2003; Dahle, Jacobs, & Raz, 2009).

Table 5.1 Demographic and Health Characteristics by Age Group

Variable	Participants (*N* = 171)		χ^2/t	Effect Size	
	Young Adults (*n* = 86)	Older Adults (*n* = 85)		*d*	ϕ
Age	19.80 (2.24)	71.40 (5.46)	-	-	-
Range	*17 - 27*	*64 - 87*	-	-	-
Education[a]	13.26 (1.21)	14.33 (2.27)	-3.87***	.51	
Range	*12 - 16*	*9 - 20*	-	-	-
Female[b] **(%)**	73.3	69.4	0.31		-.04
Ethnicity[b] **(% Caucasian)**	45.3	95.3	50.95***		.55
Birthplace[b] **(% foreign born)**	23.3	32.9	1.99		-.11
EAL[b] **(%)**	27.9	11.8	6.99**		-.20
MMSE	-	29.16 (1.12)	-	-	-
CES-D[a]	11.87 (7.96)	6.82 (6.27)	4.61***	.70	
MAQ[a]	66.45 (14.39)	53.79 (11.16)	6.43***	.97	
SBP (mmHg)[a]	111.83 (10.91)	129.82 (15.36)	-8.84***	1.35	
Range	*92 – 140*	*95 – 171*	-	-	-
DBP (mmHg)[a]	67.74 (6.75)	74.51 (8.98)	-5.57***	.85	
Range	*54 – 87*	*58 – 99*	-	-	-
PP (mmHg)[a]	44.08 (8.11)	55.16 (12.60)	-6.84***	1.05	
Range	*25 – 67*	*34 – 85*	-	-	-
Cardiovascular Risks (% diagnosed)					
Hypertension[b]	0	37.6	39.83***		.48
Type2 diabetes[b]	0	8.2	7.39**		.21
High cholesterol[b]	2.3	34.1	29.11***		.41
CVD[b]	0	15.3	14.24***		.29
Anti-HTN use[b]	0	30.6	31.02***		.43

Note: We present means and standard deviations as M (SD). Birthplace (% foreign born) = reported birthplace outside of North America; EAL = reported English as an additional language; MMSE = Mini-Mental Status Examination (range = 0 to 30); CES-D = Centre for Epidemiological Studies Depression Scale (range: 0 to 60); MAQ = Multidimensional Anxiety Questionnaire (range = 40 to 160); SBP = mean systolic blood pressure; DBP = mean diastolic blood pressure; PP = mean pulse pressure; CVD = cardiovascular disease. Anti-HTN use = current confirmed use of anti-hypertensive medication.

[a] *p* value and Cohen's *d* derived from *t*-test (continuous data; small ES ≤ .20; medium ES ≥ .50; large ES = .80; very large ES ≥ 1.30). [b] *p* value and phi coefficient (ϕ) derived from χ^2 test (binary categorical data; small ES ≤ .10; medium ES ≥ .30; large ES ≥ .50; Cohen, 1988).

*$p < .05$, **$p < .01$, ***$p < .001$.

5.1.2. Reliability of Theory of Mind Measures

We examined item-level properties of our ToM measures. We deleted 22 out of 111 original items (or 19.8%) that demonstrated poor response variability or very low item-total correlations (i.e., $r < .10$; Meyers et al., 2013), as it was questionable whether these items reflected the same construct assessed by the total scores (see Appendix F, Table F1). The percentage of items we deleted per ToM measure is similar to that reported by other authors (e.g., Söderstrand & Almkvist, 2012). We then calculated scale reliability. Across the full sample, reliability for cognitive ToM measures ranged from good to excellent: STORIES test items demonstrated excellent inter-rater agreement across three independent raters ($ICC_{(3,1)} = .95$; after deleting one item with poor response variability), and YONI C-TOM test items demonstrated good internal consistency ($ICC_{(3,1)} = .86$; after deleting three items with poor response variability and three items with low item-total agreement). Reliability for affective ToM measures was lower: YONI A-TOM test items demonstrated adequate internal consistency ($ICC_{(3,1)} = .74$; after deleting two items with poor response variability and three items with low item-total agreement) and RMET test items demonstrated poor internal consistency ($ICC_{(3,1)} = .48$; after deleting two items that poorly discriminated among foil and target response options and eight items with low item-total agreement). Table 5.2 summarizes the final psychometric properties after item deletion. While our estimates are lower than recommended psychometric standards (Henson, 2001), they are comparable to other recently published data on affective ToM measures (e.g., internal consistency reliability: $\alpha = .64$, Söderstrand & Almkvist, 2012, $\alpha = .61$, Vellante et al., 2013; test-retest reliability [one-year interval] ICC = .63; Fernández-Abascal, Cabello, Fernández-Berrocal, & Baron-Cohen, 2013). Further discussion of ToM reliability is presented in Appendix G.

Table 5.2 Psychometric Properties for the Theory of Mind Measures

Test	Possible score range	Actual score range	Internal consistency ICC [95% CI][a]	Inter-rater agreement ICC [95% CI][b]
STORIES	0 – 14	5 – 14	-	.95 [.93, .95]
RMET	0 – 26	9 – 25	0.48 [.36, .59]	-
YONI C-TOM	0 – 28	10 – 28	0.86 [.83, .89]	-
YONI A-TOM	0 – 28	16 – 28	0.74 [.68, .80]	-

[a] For the RMET, YONI C-TOM, & YONI A-TOM measures, we present internal consistency as the intraclass correlation coefficients $ICC_{(3,1)}$ for mixed effects models (average measures).

[b] For STORIES, we present inter-rater agreement as the intraclass correlation coefficient, $ICC_{(3,1)}$ for a mixed effects model (average measures) assessing the consistency between three independent raters.

5.1.3. Descriptive Analyses and Data Reduction

Theory of mind data

Univariate statistics were within acceptable ranges for each ToM measure; however data for Yoni Test data had negatively skewed distributions (see Appendix F, Table F2; Tabachnick & Fidell, 2013). Based on supporting theory, we created composite measures of cognitive and affective ToM by converting data on the individual measures to *z*-scores and summing them (Edgington, 1995). The *Cognitive ToM* composite included items from the two measures that assessed inferences to meta-cognitive mental states (STORIES & YONI C-TOM), and the *Affective ToM* composite included items from the two measures that assessed inferences to emotional mental states (RMET & YONI A-TOM). The *z*-score composites had good univariate properties, thus stabilizing potential influences of skewness in the raw data (Tabachnick & Fidell, 2013). To better understand the implications of data reduction and to assess potential interpretive advantages, we compared the results obtained using *z*-score composites with logarithmic and square root transformed data. No meaningful difference was observed. Thus, we retained the *z*-score composites for all analyses as this provided the most parsimonious interpretation of our results in the context of existing ToM literature.

Table 5.3 presents inter-test correlations between ToM measures for the full sample. The degree of association between measures of cognitive ToM was very low or

absent (r = .17, p = .02), whereas the association between measures of affective ToM was low (r = .24, p < .01; classified via O'Rourke et al., 2005). The generally low inter-test associations we observed are consistent with estimates published in other recent research (e.g., .11 < r < .14, Ahmed & Miller, 2011; .34 < r < .42, Bailey & Henry, 2008; -.01 < r < .42, Phillips et al., 2002; r = .27, Rakoczy et al., 2012), and may reflect measurement error or poor test design (see discussion for details).

Table 5.3 Correlation Matrix for Theory of Mind Variables

	STORIES	RMET	YONI C-TOM	YONI A-TOM	COGNITIVE TOM
			Full Sample		
STORIES	-				
RMET	.17*	-			
YONI C-TOM	.19*	.14	-		
YONI A-TOM	.14	.24**	.70***	-	
Cognitive ToM	.77***	.20*	.77***	.55***	-
Affective ToM	.20*	.79***	.53***	.79***	.47***

Note: *Cognitive ToM* and *Affective ToM* reflect the composite z-score variables. For all variables higher scores indicate better performance.

*p < .05, **p < .01, ***p < .001.

Neurocognitive Variables

Table 5.4 presents mean scores by age group across the individual neurocognitive measures. As expected, young adults outperformed older adults on all measures, except verbal comprehension, where older adults performed better (d = 1.89; very large ES). The magnitude of age differences was large (d ≥ .80; learning and memory, working memory) to very large (d ≥ 1.30; inhibition). Age differences in simple auditory attention approached a medium effect size (d = .48). This pattern of results is consistent with past research examining age differences in cognitive performance (Drag & Bieliauskas, 2010; Salthouse, 2010).

Table 5.4 Mean Neurocognitive Performance by Age Group

Test	Young (*n* = 86)	Older (*n* = 85)	*t*-test		Effect Size	Correlation with age
	Mean (SD)	Mean (SD)	*t*	*p*	*d* [95% CI]	*r*
KBIT VK[a]	49.20 (3.41)	55.01 (2.71)	-12.29	< .01	1.90 [1.44, 2.36]	.68***
CVLT 15[a]	55.85 (6.66)	45.25 (11.77)	7.24	< .01	1.12 [.31, 2.54]	-.51***
CVLT SDFR[a]	12.57 (2.42)	9.11 (3.38)	7.70	< .01	1.19 [.75, 1.62]	-.52***
CVLT LDFR[a]	12.72 (2.42)	9.72 (3.47)	6.55	< .01	1.01 [.57, 1.46]	-.46***
WAIS LN	11.91 (2.69)	9.59 (2.22)	6.14	< .01	.95 [.58, 1.31]	-.44***
WAIS DS	7.95 (2.40)	6.86 (2.19)	3.10	< .01	.48 [.14, .82]	-.23**
DKEFS CW[a,b]	17.94 (6.27)	32.80 (12.31)	9.93	< .01	1.53 [.80, 2.99]	-.62***

Note: KBIT VK = KBIT-2 Verbal Knowledge (range: 0 to 60); CVLT 15 = CVLT-II Trials 1-5 (range: 0 to 80); CVLT SDFR = CVLT-II Short Delay Free Recall (range: 0 to 16); CVLT LDFR = CVLT-II Long Delay Free Recall (range: 0 to 16); WAIS LN= WAIS-III Letter Number Sequencing (range: 0 to 21); WAIS DS = WAIS-III Backwards Digit Span (range: 0 to 15); DKEFS CW = DKEFS Color-Word Interference Contrast Score. All measures except that noted in (b) below are coded so that higher scores indicate better performance. Correlations with age are represented as Pearson's *r* correlation coefficients for the association between each variable and age group (0 = young adults; 1 = older adults).

[a] Unequal variances assumed. *p* value and Cohen's *d* derived from *t*-test (continuous data; small ES ≤ .20; medium ES ≥ .50; large ES ≥ .80; very large ES ≥ 1.30; Cohen, 1988).

[b] DKEFS CW scores are a timed contrast measure. On this measure higher scores in the Mean (SD) columns indicate slower or worse performance. These scores were reverse coded to calculate the correlation with age (last column) and for all primary analyses to maintain a consistent metric with the other neurocognitive scores (i.e., where higher scores represent better performance).

* $p < .05$, ** $p < .01$, *** $p < .001$.

Table 5.5 presents inter-test correlations among the neurocognitive measures for the full sample. Because detailing precise relationships between ToM and specific neurocognitive abilities was not our central aim and to meaningfully reduce the number of independent variables for path analysis, we created composite neurocognitive variables based on theoretical distinctions (e.g., Salthouse, 2009). Individual patterns of correlations supported this procedure: The three executive functioning measures displayed low to moderate associations (|.26 < *r* < .41|), and the three memory measures displayed moderate to high associations (|.40 < *r* < .87|; classified via O'Rourke et al., 2005). We distinguished between composites of Learning & Memory and Executive Functions for the following reasons. First, there is strong theoretical rationale that episodic memory and executive functions are separate constructs that also display differential relationships with age and with brain morphology (e.g., Dahle et al., 2009; Salthouse, 2009). Second, these domains are differentially related to variability in ToM performance in both young and older adults (Fischer et al., 2014; Mahy et al., 2014b; Sandoz et al., 2014). To this end, we converted created summed *z*-score composites representing Learning & Memory (CVLT 15, CVLT SDFR, & CVLT LDFR) and Executive Functions (WAIS LN, WAIS DS & DKEFS CW; Edgington, 1995). These composites were used to index neurocognitive performance in all subsequent analyses.

Correlations

Table 5.6 presents Pearson correlation coefficients between cognitive ToM, affective ToM, age, Learning & Memory, Executive Functions, and PP. Associations with sex (M/F) reflect Spearman's rank-order correlations. Given extensive research documenting age differences in the neurocognitive and blood pressure predictors (e.g., Dahle et al., 2009; Salthouse, 2010) we present correlations separately by age group. These correlations provided an indication of associations between the neurocognitive predictors and PP on ToM without accounting for demographics. For correlation matrices between ToM and additional variables of interest see Appendix F, Table F4.

Table 5.5 Correlation Matrix for Neurocognitive Variables

	KBIT VK	*CVLT15*	*CVLT SDFR*	*CVLT LDFR*	*WAIS LN*	*WAIS DS*
			Full Sample			
KBIT VK	-					
CVLT 15	.30***	-				
CVLT SDFR	.27***	.81***	-			
CVLT LDFR	.30***	.78***	.87***	-		
WAIS LN	.26***	.49***	.40***	.40***	-	
WAIS DS	.22*	.22**	.27***	.31***	.41***	-
DKEFS CW	.20**	.35***	.26**	.28***	.32***	.26***

Note: KBIT VK = KBIT-2 Verbal Knowledge; CVLT 15 = CVLT-II Trials 1-5; CVLT SDFR = CVLT-II Short Delay Free Recall; CVLT LDFR = CVLT-II Long Delay Free Recall; WAIS LN= WAIS-III Letter Number Sequencing; WAIS DS = WAIS-III Backwards Digit Span; DKEFS CW = DKEFS Color-Word Interference Contrast Score. All measures are coded so that higher scores represent better performance. All reported associations represent Pearson correlation coefficients.

*p < .05, **p < .01, ***p < .001.

Table 5.6 Correlation Matrices for Theory of Mind, Neurocognitive Variables, and Pulse Pressure by Age Group

	Cognitive ToM	*Affective ToM*	*Age*	*Sex*[a]	*Memory*	*Executive Functions*
			Young Adults			
Cognitive ToM	-					
Affective ToM	.43***	-				
Age	-.04	0	-			
***Sex* (M/F)**[a]	-.15	.01	-.12	-		
Memory	.21*	.11	.17	-.10	-	
Executive Functions	.28**	.23*	.17	-.06	.42***	-
PP	-.13	-.13	-.05	-.56***	-.06	-.02
			Older Adults			
Cognitive ToM	-					
Affective ToM	.51***	-				
Age	-.24*	-.29**	-			
***Sex* (M/F)**[a]	-.17	.12	-.02	-		
Memory	.28*	.37**	-.26*	.53***	-	
Executive Functions	.39***	.42***	-.24*	.16	.53***	-
PP	-.29**	-.14	.16	.28**	.05	-.21*

Note: *Cognitive ToM* and *Affective ToM* reflect composite z-score variables. Memory = Learning & Memory *z*-score composite; Executive Functions = Executive Functions *z*-score composite. For ToM and all neurocognitive variables, higher scores represent better performance. We report age in years and sex as 0 = male, 1 = female.

[a] All reported associations are presented as Pearson correlation coefficients; except for associations with sex (M/F), which reflect Spearman's rank correlation coefficients.

*$p < .05$, **$p < .01$, ***$p < .001$.

5.2. Primary Research Objectives

5.2.1. Age Differences in Theory of Mind

Our first objective was to characterize the presence and magnitude of age differences in cognitive and affective ToM. As predicted, young adults outperformed older adults across all individual ToM measures. Figure 5.1 depicts the percentage of accurate responses by age group and the associated 95% confidence intervals (error bars) on each of the STORIES, RMET, YONI C-TOM, and YONI A-TOM variables. In Table 5.7 we present mean age differences for each measure and the corresponding effect sizes. As predicted, effect sizes for cognitive ToM ranged from medium to large (STORIES *d* = .75; YONI C-TOM *d* = 1.10), and effect sizes for affective ToM were consistently large (RMET *d* = .86; YONI A-TOM *d* = 1.16). We observed the same pattern across the composite ToM variables, indicating that age differences in cognitive and affective ToM were robust and relatively equal in magnitude.

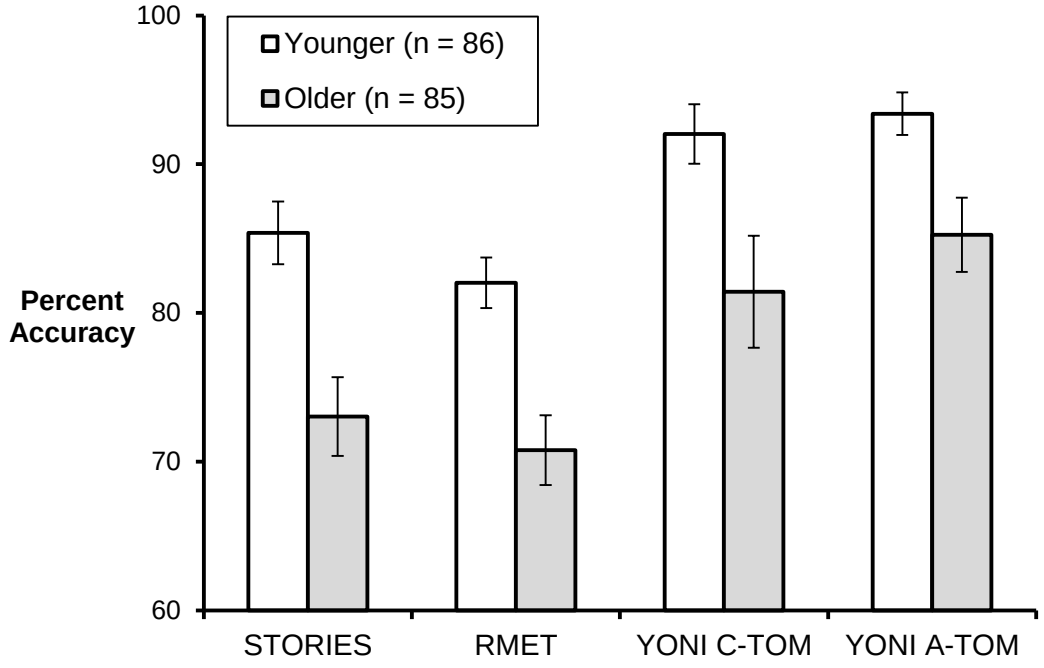


Figure 5.1 Percentage of accurate responses on ToM tests and 95% confidence intervals for mean standard error (error bars)

Note: Higher % accuracy indicates better theory of mind.

Table 5.7 Mean Theory of Mind Performance by Age Group

ToM Measure	Young (*n* = 86)			Older (*n* = 85)			Age Difference	
	Min	Max	Mean (SD)	Min	Max	Mean (SD)	*t*[a]	Effect Size *d*
STORIES	7	14	11.95 (1.41)	5	13	10.22 (1.75)	7.12***	1.10
RMET	17	25	21.33 (2.10)	9	24	18.40 (2.88)	7.60***	1.16
YONI C-TOM	12	28	25.76 (2.66)	10	28	22.80 (4.97)	4.88***	.75
YONI A-TOM	19	28	26.15 (1.89)	16	28	23.87 (3.29)	5.57***	.86

Note: The possible scores for each ToM measure are as follows: STORIES (range = 0 to 14), RMET (range = 0 to 26), YONI C-TOM (range = 0 to 28), and YONI A-TOM (range = 0 to 28). Min and Max refer to the actual minimum and maximum values obtained by sample participants.

[a] Unequal variances assumed. *p* value and Cohen's *d* derived from *t*-test (continuous data; small ES ≤ .20; medium ES ≥ .50; large ES ≥ .80; very large ES ≥ 1.30; Cohen, 1988).

* $p < .05$, ** $p < .01$, *** $p < .001$.

5.2.2. Identifying Predictors of Theory of Mind

Our second objective was to identify predictors of cognitive and affective ToM, including whether predictors differed by age. For both groups, the hypothesized model in Figure 2.1 was initially a poor fit to the data. For young adults, EC and MI statistics revealed that four modifications substantially improved the fit between our model and the data: The deletion of Learning & Memory as a predictor of cognitive ToM, the deletion of sex (M/F) as a predictor of affective ToM, the addition of sex (M/F) as a predictor of PP, and the addition of PP and sex (M/F) as predictors of cognitive ToM. For older adults EC and MI statistics revealed that five modifications substantially improved the fit between our model and the data: The deletion of Executive Functions as a predictor of cognitive ToM, the deletion of sex (M/F) as a predictor of affective ToM, the addition of sex (M/F) as a predictor of PP and Learning & Memory, the addition of sex (M/F) as a predictor of cognitive ToM, and modeling an error covariance between PP and Executive Functions. We conducted all analyses with these modifications in place.

Evaluation of Model Fit

Figure 5.2 presents the final path model representing predictors of ToM for young adults. Goodness-of-fit indices supported the hypothesized model structure with the above-noted modifications, χ^2 [*df* = 8] = 7.10, *p* = .53. Both incremental and parsimony fit indices fell within recommended thresholds and revealed good model fit, CFI = 1.00 and RMSEA = 0 (90% CI [0, .12]). The absolute fit index SRMR = .06 was slightly higher than recommended, but still within accepted ranges (Byrne, 2013). Together these factors suggested that the final model was a good fit to the young adult data.

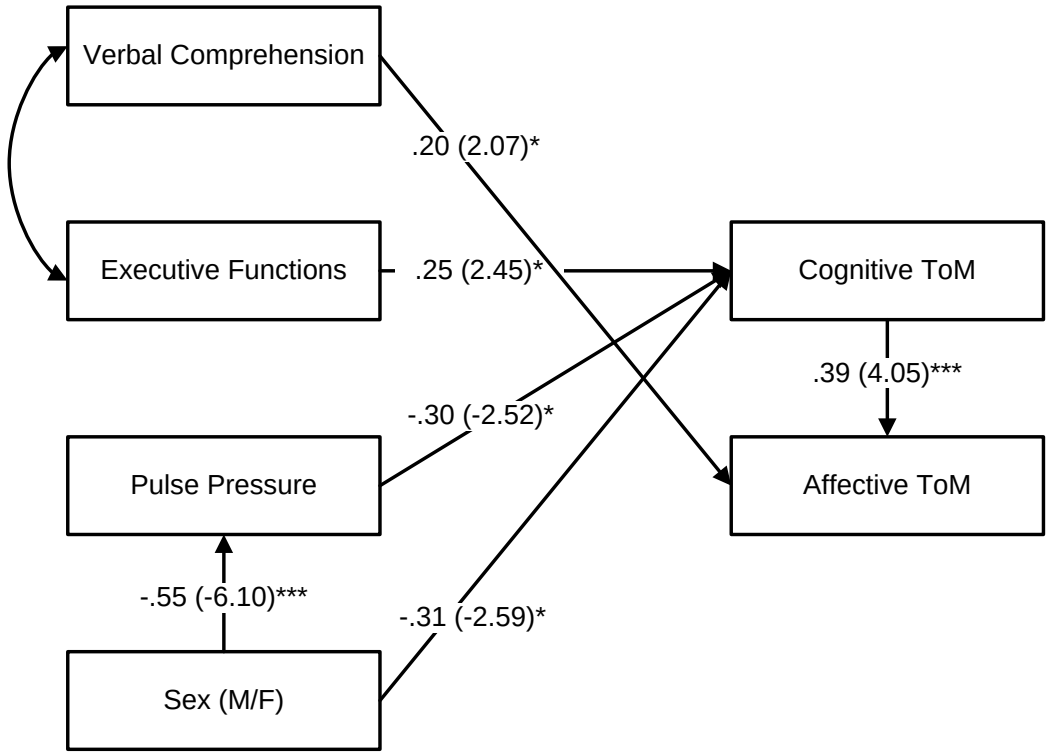


Figure 5.2 Path model for predictors of ToM in young adults

Note: χ^2 [*df* = 8] = 7.10, *p* = .53; CFI = 1.00; RMSEA = 0 (90% CI [0, .12]); SRMR = .06. Parameter values are expressed as maximum likelihood estimates (standardized solution). Numbers in parentheses indicate *t* values for parameter estimates (statistically significant *t* values > |1.96|). We coded sex (M/F) as 0 = male and 1 = female.

*p < .05, **p < .01, ***p < .001.

In Figure 5.3 we present the final path model representing predictors of ToM in older adults. Goodness-of-fit indices supported the hypothesized model structure with

the above-noted modifications, χ^2 [*df* = 8] = 7.65, *p* = .47. All incremental, parsimony, and absolute fit indices fell within recommended thresholds, CFI = 1.00, RMSEA = 0 (90% CI [0, .12]), SRMR = .06. Together these factors suggest that the model in Figure 5.3 was an excellent fit to our older adult data, with the exception of the SRMR = .06 that was slightly higher than recommended, but still generally acceptable (Byrne, 2013).

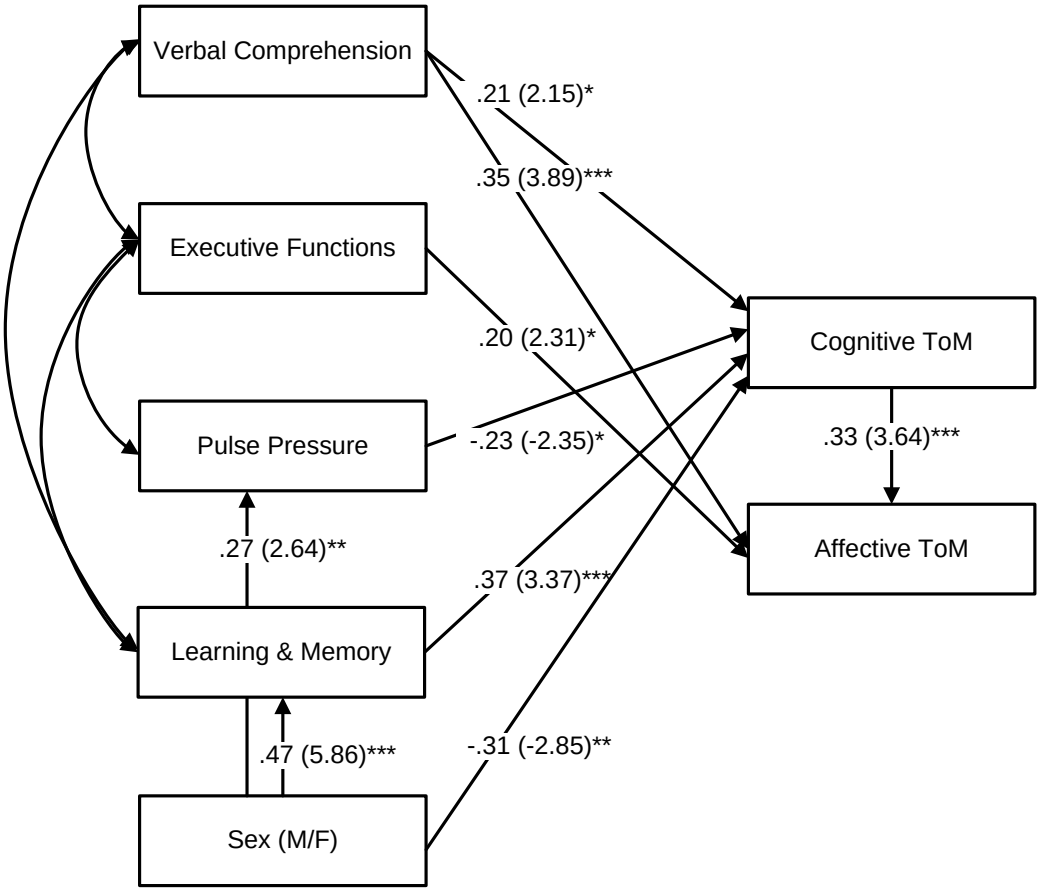


Figure 5.3 Path model for predictors of ToM in older adults

Note: χ^2 [*df* = 8] = 7.65, *p* = .47; CFI = 1.00; RMSEA = 0 (90% CI [0, .12]); SRMR = .06. Parameter values are expressed as maximum likelihood estimates (standardized solution). Numbers in parentheses indicate *t* values for parameter estimates (statistically significant *t* values > |1.96|). We coded sex (M/F) as 0 = male and 1 = female.

*p < .05, **p < .01, ***p < .001.

Shared Predictors of ToM between Age Groups

As hypothesized, several variables were significant predictors of ToM performance in both age groups: (1) higher verbal comprehension predicted better affective ToM for young (*ß* = .20, *p* = .04) and older adults (*ß* = .35, *p* < .01); (2) high PP predicted worse cognitive ToM for young (*ß* = -.30, *p* = .01) and older adults (*ß* = -.23, *p* = .02), (3) male sex predicted better cognitive ToM for young (*ß* = -.31, *p* = .01) and older adults (*ß* = -.31, *p* < .01); and (4) biological sex was significantly associated with PP, although the direction of this relationship differed by age. In young adults men had higher PP (*ß* = -.55, *p* < .01), whereas in older adults, women had higher PP (*ß* = .27, *p* < .01). Counter to our predictions, we did not observe an association between biological sex and affective ToM in either group.

Also consistent with our hypotheses, analyses revealed a moderate-sized, positive association between cognitive ToM and affective ToM for young (*ß* = .39, *p* < .01) and older adults (*ß* = .33, *p* < .01). Given existing theory from developmental and neuroimaging research (Shamay-Tsoory et al., 2009; Vetter et al., 2013), we suspected that the direction of this association was that of cognitive ToM predicting affective ToM. To explore directionality, we tested an alternative model with one modification: We changed the direction of association to reflect affective ToM predicting cognitive ToM. For both young and older adults, the alternative model revealed a similarly good fit to the sample data (young: $\Delta\chi^2$ = -.23, Δdf = 0; older: $\Delta\chi^2$ = 2.40, Δdf = 0). However, several empirically supported predictors of ToM in the original models dropped to non-significance. Specifically, for young adults, executive functions did no predict cognitive ToM in the alternative model, and for older adults verbal comprehension did not predict cognitive ToM and PP did not predict cognitive ToM in the alternative model. Thus, we retained the original models because they were the most consistent with supporting theory. We present the alternative models for young and older adults and associated goodness-of-fit statistics in Appendix H, Figures H1 and H2.

Non-shared Predictors of ToM in Older Adults

Three predictors of ToM were also significant for older adults only: (1) higher verbal comprehension predicted better cognitive ToM (*ß* = .21, *p* = .03); (2) better

executive functions predicted better affective ToM (*ß* = .20, *p* = .02); and (3) better Learning & Memory predicted better cognitive ToM (*ß* = .37, *p* < .01). In older adults only, women also displayed better Learning & Memory (*ß* = .47, *p* < .01).

Non-shared Predictors of ToM in Young Adults

In young adults only, better executive functions predicted better cognitive ToM (*ß* = .25, *p* = .01), suggesting that young adults with better executive abilities formed more accurate mental state inferences about cognitive content.

5.2.3. Age Differences in Predictors of Theory of Mind

Our third objective was to determine whether any age differences existed between predictors of cognitive and affective ToM. Using multi-group invariance procedures outlined by Byrne (2004), we tested for the equivalence of shared predictors of ToM between age groups. Our analyses represented *partial* invariance analyses because not all predictors were shared between age groups. The baseline model comparison revealed an optimal fit to the data, χ^2 [*df* = 15] = 13.82, *p* = .54; CFI = 1.00; RMSEA = 0, 90% CI [0, .07]; SRMR = .05. This indicated that when no cross-group constraints were imposed, our final models (Figures 5.2 & 5.3) represented the sample data well. We used the chi-squared value from this test (χ^2 [*df* = 15] = 13.82) as a baseline comparative value for all subsequent tests of invariance. The predictors we tested in the invariance analyses are summarized in Table 5.12

Table 5.8 Predictors Tested for Invariance

Label	Predictive Association Tested
a	Verbal Comprehension → Affective ToM
b	PP → Cognitive ToM
c	Sex (M/F) → Cognitive ToM
d	Cognitive ToM → Affective ToM
e	Sex (M/F) → PP

Because it is possible (although unlikely) that all predictors outlined in Table 5.12 might be equivalent between age groups, we first tested the validity of the invariance model with predictive associations held equal (Table 5.13, line 2). Comparison of Models

1 and 2 yielded a $\Delta\chi^2 = 49.41$ and $\Delta df = 5$, which is statistically significant. This indicated that one or more predictors were not equal between age groups and provided sufficient rationale to individually test each predictor for invariance. We present results from this series of tests in Table 5.13.

After testing for the invariance of all shared paths, only one predictor differed between age groups: Biological sex as a predictor of PP (predictor *e* in Tables 5.9 & 5.10). Inspection of the standardized regression coefficients reported in Figures 5.2 and 5.3 confirmed that this association was qualitatively different with age: in young adults male sex predicted high PP (*ß* = -.55), whereas in older adults female sex predicted high PP (*ß* = .27). Importantly, no statistically meaningful age differences existed between verbal comprehension and affective ToM (predictor *a*), PP and cognitive ToM (predictor *b*), biological sex and cognitive ToM (predictor *c*), or between cognitive ToM and affective ToM (predictor *d*). Irrespective of age, high PP and male sex predicted better cognitive ToM, whereas better verbal comprehension and better cognitive ToM predicted better affective ToM, thus indicating that the relationship between ToM and fundamental supporting variables did not differ between young and older adults.

Table 5.9 Invariance Analyses for Age Differences in Predictors of ToM

Model Description	χ^2	*df*	$\Delta\chi^2$	Δdf	χ^2_{crit} [Δdf, α = .05]	*p*	Interpretation
1. Baseline model (unconstrained)	13.82	15	—	—	—	—	—
2. All predictors (*a* to *f*) held equal	63.23	20	49.41	5	11.07	< .01	—
3. Predictor *a* held equal	14.94	16	1.12	1	3.84	*ns*	*Invariant*
4. Predictors *a* & *b* held equal	15.14	17	1.32	2	5.99	*ns*	*Invariant*
5. Predictors *a, b & c* held equal	15.15	18	1.33	3	7.81	*ns*	*Invariant*
6. Predictors *a, b, c, & d* held equal	15.29	19	1.47	4	9.49	*ns*	*Invariant*
7. Predictors *a, b, c, d, & e* held equal	47.03	20	33.21	5	11.07	< .01	*Age difference*

Note: $\Delta\chi^2$ = chi-square difference value between models; Δdf = difference in degrees of freedom between models; χ^2_{crit} = critical ratio from χ^2distribution associated with Δdf at a pre-defined alpha level of.05. Significance values correspond to $\Delta\chi^2$. We compared each of the Models 2-8 to Model 1 (baseline model)

5.3. Post-hoc Analyses

The association between high PP and worse cognitive ToM was both shared by young and older adults and equivalent in magnitude. This was surprising to us given the physiological widening of PP after age 60 and its strong relationship with late-life cardiovascular and cerebrovascular outcomes (Gorelick et al., 2011). Some researchers have suggested that modeling blood pressure, including PP, as a non-linear variable may assist in interpreting blood pressure effects on cognition (Waldstein, Giggey, Thayer, & Zondervan, 2005; see Appendix I for further discussion). To investigate this possibility in an alternative model we tested a quadratic PP function in place of the linear variable reported above. Findings revealed that use of the non-linear PP variable did not offer any statistical advantage to the fit of our final path models (young adults: χ^2 [*df* = 8] = 6.53 [$\Delta\chi^2$ from original model = -.57], *p* = .59; CFI = 1.00; RMSEA = 0 (90% CI [0, .11]); SRMR = .06; older adults: χ^2 [*df* = 8] = 15.29 [$\Delta\chi^2$ from original model = 7.64], *p* = .05; CFI = .94; RMSEA = .10 (90% CI [0, .18]); SRMR = .07). In additional, individual regression for the relationship between PP and cognitive ToM were similar in direction and magnitude when compared to the linear model. These results are consistent with other results from our lab (i.e., Fischer et al., 2014; Yeung & Thornton, 2011), and suggest that in non-clinical, community-residing samples, the application of nonlinear transformations to blood pressure data may not confer any advantage to characterizing associations with neurocognitive performance.

Chapter 6. Discussion

We began this study with the framework of ToM as a multidimensional construct wherein cognitive and affective components were distinguished on the basis of developmental, neuropsychological, and neurobiological theory and evidence. Building on questions posed in past research we comprehensively examined key predictors of ToM and assessed whether the strength of predictive relationships differed by age. Our findings revealed that while cognitive and affective ToM each showed robust, moderate-to-large age effects; key differences with underlying predictors exist. This dissertation makes three unique contributions to the understanding of ToM as a multidimensional construct. First, extending our previous work (Fischer et al., 2014), we provide the first evidence that poor vascular health, as assessed by PP, is directly associated with lower cognitive ToM. PP was a key predictor of cognitive ToM in young and older adults alongside executive functions, verbal comprehension, learning and memory, and biological sex. Second, in agreement with child development and cognitive neuroscience theory, we present the first neuropsychological evidence suggesting an organizational structure within ToM, such that cognitive ToM may be fundamental to affective ToM. Finally, we demonstrated that while certain neurocognitive predictors of ToM are salient only in later life, most predictors are shared across adulthood and are statistically equivalent. Together our results promote cognitive and affective ToM as integrated components that share a common basis in adulthood. This notion is key to clarifying how ToM changes with age and to advancing knowledge regarding the relevance of such changes in the context of underlying neurocognitive, health, and demographic factors.

6.1. Pulse Pressure Predicts Cognitive ToM

A unique finding was that cardiovascular inefficiencies were adversely associated with ToM performance. Qualifying our previous work (Fischer et al., 2014), high PP directly predicted worse cognitive ToM. Importantly, this association was observed for

both age groups and was equivalent in magnitude. Counter to our predictions, PP was not directly associated with affective ToM in either group. The equally robust relationship between PP and cognitive ToM between groups was surprising, as the physiological widening of PP after age 60 and its relation to late-life cerebrovascular outcomes (Gorelick et al., 2011; Nagai et al., 2010) led us to hypothesize a more salient relationship in older age. Given the potential for blood pressure-cognition associations to be age-dependent, we explored whether modeling PP as a non-linear variable may better capture the relationship between PP and ToM (Thorvaldsson et al., 2012; Waldstein et al., 2005). Similar to previous research from our lab (e.g., Fischer et al., 2014; Yeung & Thornton, 2011), we found no statistical advantage to this approach, suggesting that PP-ToM associations are most likely linear by nature. We note that blood pressure is a complex parameter and it is possible that associations in young adults may in part, reflect diverse influences including medical history, genetics, ethnic background, state anxiety, alcohol and/or caffeine consumption, or physiological arousal (Waldstein et al., 1995). In the current study adjusting for the variables ethnicity, depressive symptoms, and anxiety symptoms did not alter the results.

The presence of CVRFs in early adulthood is increasingly recognized as a risk factor for neurocognitive health (Elias, Elias, Robbins, & Budge, 2004; Yaffe et al., 2014). Aine and colleagues (2014) suggested that cardiovascular risk in early adulthood is frequently underestimated because individual CVRFs may fail to meet clinical diagnostic thresholds. This may lead clinicians and researchers to erroneously dismiss the potential risks to cognitive integrity at the time of measurement. For instance, Joosten and colleagues (2013) demonstrated that associations between high cardiovascular risk and worse neurocognitive performance were statistically equivalent between adults aged 35 to 44 years, 45 to 64 years, and 65+ years. This risk appears to apply to current *and* future neurocognitive health. In a prospective study Elias and colleagues (2004) demonstrated that young (18 to 47 years) and older adults (48 to 83 years) displayed equally strong relationships between baseline blood pressure (SBP, DBP, & PP) and incipient cognitive decline 20 years later. Considered alongside such evidence, our results support the dampening effect of poor cardiovascular health on neurocognitive performance, including cognitive ToM, as a phenomenon that may affect both young and older adults.

A secondary finding regarding PP pertains to sex-specific relationships: In young adults men had higher PP, whereas in older adults women had higher PP. Sex differences in vascular health are relatively unstudied in early adulthood; however, there is relatively strong evidence that older women are particularly susceptible to blood pressure difficulties. In Canada, treatment rates are roughly equal between men and women, yet women are less likely to achieve blood pressure control (Wilkins, Gee, & Campbell, 2012). Biomedical research supports key differences in the neurophysiological pathways underlying blood pressure control that may explain these trends. Specifically, physiological differences in the production and stimulation of sex hormones, the renin-angiotensin and oxidative stress pathways, and in the tone and function of central vasculature and the sympathetic nervous system each contribute to blood pressure regulation (Zimmerman & Sullivan, 2013). Presently, sex-specific treatment approaches are not commonplace but there is increasing recognition of their importance (Wilkins et al., 2012). The potential value of this type of treatment practice becomes even more important when considering downstream age- and sex-dependent effects of hypertension on neurocognitive performance reported in this dissertation and other recent work (e.g., Yasar et al., 2011; Yeung & Thornton, 2011).

Proposed mechanisms linking PP to neurocognitive performance, including ToM, relate to the effects of arterial stiffening on the compliance and integrity of vulnerable subcortical circuitry, leading to small vessel disease and increased risk of ischemic cognitive impairment (Birns & Kalra, 2009; Raz, Rodrigue, & Acker, 2003). Reduced arterial elasticity caused by high PP can lead to impaired cerebral blood flow regulation and hemodynamic inefficiencies that can cause brain tissue to be more susceptible to injury (Scuteri et al., 2011). Structural and functional integrity may be altered by white matter infarcts, cerebral atrophy, or *β*-amyloid angiopathy, leading to inefficiencies in cellular metabolism and neuronal transmission and, ultimately, disturbances in neurocognitive function (Scuteri et al., 2011; Waldstein et al., 2008). Neuroimaging studies confirm that high PP is associated with reduced volume and white matter integrity, particularly in frontal regions that are responsible for executive functions and ToM (O'Brien et al., 2002; Raz et al., 2003). Of note, this literature is based on studies of healthy aging populations, and mechanisms may differ in young samples due to possible interactive influences with age or other previously discussed variables (Waldstein, 1995).

6.2. Dissociable Predictors of Cognitive and Affective ToM

Consistent with a multidimensional framework, cognitive and affective ToM held differential relationships with supporting predictors. For cognitive ToM, high PP and male sex were key predictors across age groups. Age-specific relationships were also observed: Executive functions predicted cognitive ToM in young adults, and verbal comprehension and learning and memory predicted cognitive ToM in older adults. For affective ToM, verbal comprehension was a key predictor across age groups, with executive functions emerging as relevant in older age only. In light of past work (e.g., Mahy et al., 2014b; Rakoczy et al., 2012), we were surprised that executive functions did not globally predict cognitive and affective ToM. The pattern of association we observed is also mismatched with developmental theories stating that at critical periods of neurocognitive development and decline, cognitive and affective ToM should be sensitive to executive influences (Carlson & Moses, 2001; Ahmed & Miller, 2013). One possibility is that the theorized trends may pertain to specific executive abilities, such as response inhibition, working memory, or attention (Carlson & Moses, 2001). That is, our use of a composite executive measure may have masked potentially important relationships between ToM and individual processes. Further, in pursuing a comprehensive examination of cognitive and non-cognitive predictors, our inclusion of multiple concurrent predictors and/or controlling for shared measurement error between neurocognitive measures may have attenuated expected relationships. We also did not specifically examine indirect relationships between these variables and ToM due to limitations in statistical power (see Appendix J for further information).

Our findings also point toward key differences in how young and older adults may use verbal resources to succeed on tasks of ToM. While verbal abilities predicted affective ToM across groups, only older adults showed links between verbal comprehension and cognitive ToM. We suspect this finding captures our sample of well-educated older adults who were less susceptible to linguistic interference and therefore better able to ignore incorrect or misleading wording on our ToM measures. This may be particularly true for cognitive ToM, wherein one of our measures (Strange Stories) required participants to use rich expressive language to convey mental state inferences. Young adults with still-developing verbal abilities may rely on other strategies such as guessing or rapid executive selection processes to respond to ToM questions (Mahy et

al., 2014a). It is also possible that other verbal skills not tested in this study may explain additional variance in ToM (e.g., grammatical understanding, verbal fluency, reasoning: Rakoczy et al., 2012). However, Peterson and Miller (2012) found that verbal intelligence predicted affective ToM beyond other language elements such as figurative abilities (e.g., understanding and use of concepts such as relationship, emphasis, and figures of sound), thereby suggesting that it is the semantic knowledge component of language that directly influences ToM. We considered whether our findings might reflect the greater proportion of our young adult sample that identified a native language other than English (27.9% vs. 11.8% of older adults). Negligible correlations between cognitive ToM, affective ToM, and native language did not provide impetus to explore this further (see Appendix F, Table F4). Indeed, research suggests that at least in early life, bilingualism may be advantageous to the development and use of ToM (Goetz, 2003).

The finding that learning and memory predicted cognitive ToM only among older adults raises the important question of whether memory is truly fundamental to ToM. Our composite measure of cognitive ToM equally represented the Strange Stories test, which had built-in memory demands, and the Yoni Cognitive ToM task, which did not. Inclusion of this story-based measure may have increased demands on new learning efficiency, encoding, and short-term recall, thereby enhancing the likelihood of a significant association with our learning and memory composite. For older adults with weakened memory the saliency of this association was likely magnified (see also Baglio et al. 2012; Castelli et al., 2011; Fischer et al., 2014). Conversely, on our affective ToM measures participants viewed test stimuli continually while making responses, thereby mitigating encoding demands in favour of working memory and attention (Maylor et al., 2002). Perhaps the remaining question for researchers is not whether memory predicts ToM, but rather what type of memory is important? In a recent paper Moreau, Viallet, and Champagne-Lavau (2013) showed that generating mental state inferences required individuals to recall autobiographical memories of similar first-hand experiences as a template for anticipating and interpreting behaviour. This notion is further supported by neuroimaging findings that brain regions activated during ToM inferences overlap with those underlying autobiographical recollection (bilateral medial prefrontal cortex, medial parietal cortex, Rabin & Rosenbaum, 2012), as well as research with individuals with

amnestic MCI (i.e., characterized by early loss of autobiographical memories), who also show early and significant impairments in ToM (Baglio et al., 2012; Moreau et al., 2015).

We were surprised that male sex strongly predicted cognitive ToM in both age groups while no relationship was observed between sex and affective ToM. This stands in contrast to research suggesting a female advantage on tests of affective ToM. However, on close examination of this literature, the bulk of evidence favouring females on tasks of affective ToM is reported in children between infancy and adolescence (Charman et al., 2002; Ibanez et al., 2013), with cautious support into early adulthood (Kirkland et al., 2013). Thus it is possible that sex differences in affective ToM are more salient in early development, or perhaps our strategy of accounting for neurocognitive variables on which individuals may differ in performance by sex (e.g., verbal comprehension; Heaton et al., 2009) may have attenuated variance previous attributed to sex in past studies. We are aware of one other study where men outperformed women on a cognitive ToM task. Russell, Tchanturia, Rahman, and Schmidt (2007) suggested that enhanced cognitive ToM among men might reflect sex differences in systematizing information and analytical reasoning. Indeed, social neuroscience research suggests that women tend to engage brain regions underlying emotion during social cognitive tasks, whereas men recruit areas preferentially involved in cognitive control and deliberative thinking (Christov-Moore et al., 2014). We specifically examined biological sex rather than identified gender due to important sex-related neurobiological differences in cognition and blood pressure (Heaton et al., 2009; Yasar et al., 2011; Zimmerman & Sullivan, 2013). However, it is possible that gender roles and predisposition to empathy may have influenced our results. There is preliminary evidence that asking participants to adopt 'enhanced empathy' can improve accuracy on measures of ToM, regardless of sex (Thomas & Maio, 2008). Importantly, our results call into question the extent to which sex differences in cognitive and affective ToM can be generalized across age and measures used to assess them.

6.3. Inter-relationship of Cognitive and Affective ToM

Our results support cognitive and affective ToM as distinct, but overlapping conceptual entities. We demonstrated a moderate-sized, positive association such that

cognitive ToM predicted affective ToM in both age groups. Changing the direction of this relationship to affective ToM predicting cognitive ToM resulted in a substantially weaker fit to underlying theory. Earlier we reviewed multi-disciplinary evidence suggesting that cognitive ToM may provide key organizational scaffolding for the emergence and subsequent function of affective ToM (e.g., Shamay-Tsoory et al., 2009; Stone & Gerrans, 2006; Vetter et al., 2013). This hypothesis is supported by differences in the developmental timing of the two ToM components and neuroimaging observations that, in response to affective ToM inferences, associated neural correlates (i.e., ventromedial prefrontal cortex, inferior frontal gyrus) are activated in addition to the broader ToM network (Schlaffke et al., 2015; Shamay-Tsoory et al., 2009). Thus, in addition on-line support from perceptual skills and neurocognitive influences, cognitive ToM appears to be an important part of the organizational structure supporting affective ToM. For much of this work we have argued that the cognitive and affective components of ToM are dissociable and dependent on different supporting processes, yet in natural environments, our capacity to make inferences about other people may routinely involve both components (Dvash & Shamay-Tsoory, 2014). It is their interaction with one another, and with the array of supporting neurocognitive, health, and demographic resources, that provides the hallmark for sophisticated social reasoning. Moving forward, more research is needed to unravel the theoretical distinction and integration between cognitive and affective ToM. This work likely necessitates a multi-disciplinary approach wherein analysis can be conducted at multiple levels of organization (e.g., brain physiology and morphology, developmental timing, neurocognitive resources) and examination of both direct and indirect relationships in order to elucidate the causative mechanisms at play (see Kraemer, Stice, Kazdin, Offord, & Kupfer, 2001). Nevertheless, we believe our findings advance current perspectives on regarding the interplay between cognitive and affective ToM.

6.4. Age Differences in Predictors of ToM

Our use of invariance analyses to address whether the strength of association for shared predictors of ToM represents a novel and important advantage over past work. We found that only one shared relationship differed by age and this relationship was not directly pertinent to our hypotheses regarding ToM. No statistically meaningful age

existed between verbal comprehension and affective ToM, PP and cognitive ToM, biological sex and cognitive ToM, or between cognitive ToM and affective ToM. Irrespective of age, high PP and male sex predicted better cognitive ToM, whereas better verbal comprehension and better cognitive ToM predicted better affective ToM. These results have important implications for understanding how the ToM construct may differ with age. It is clear that longitudinal research is needed to clarify whether these predictors are useful markers of future changes in ToM performance. Preliminary cross-sectional data point toward an inverted U-shaped pattern of age differences in ToM performance across childhood, adolescence, young adulthood, and later life (Bernstein et al., in preparation); however, some authors contend that age differences in social cognition may reflect cohort differences rather than true cognitive change (O'Brien, Konrath, Grühn, & Hagen, 2013).

6.5. Challenges & Limitations

A primary concern of the ToM literature is score reliability. In this dissertation, estimates of score reliability ranged from inadequate to good, where scores on affective ToM measures were the least reliable. These estimates are consistent with the few studies that have previously reported on the psychometric properties of ToM measures (e.g., Fischer et al., 2014; Söderstrand & Almkvist, 2012; Vellante et al., 2013). Low score reliability may be related to poor item representation of the broader construct, mismatches between the population under study, and the types of ToM questions asked. An additional concern is that few ToM measures exist that clearly delineate between cognitive and affective ToM, and between first- and second-order question types. When administered to cognitively healthy, community-residing adults and older adults, first-order questions may be too easy—eliciting near ceiling performance, while responses on harder items may fall closer to chance. The combination of these problems may leave few items with good score variation, resulting in truncated variance and skewed distributions. We investigated the influence of the distributional properties of our ToM measures in a series of post-hoc analyses. We ran new models wherein we removed participants who performed at ceiling (i.e., 100%) on (a) at least one ToM measure, and (b) on both measures comprising a composite. Forty-five percent of the young sample and 19% of the older sample had ceiling performance on one or more ToM measures

(Strange Stories: young = 9.3% older = 0%; Yoni Cognitive ToM: young = 30.2% and older = 12.9%; Yoni Affective ToM: young = 27.9% and older = 12.9%; RMET: young and older = 0%). A total of 5.8% of young adults performed at ceiling on both measures comprising the Cognitive ToM composite. Age differences were maintained across groups when we conducted analyses with these participants removed. Further, goodness-of-fit indices suggested that the alternative path models fit the data well and coefficients for individual predictors were similar in direction and magnitude to the original models. Given these results, our findings appear relatively robust to ceiling effects. We elected to retain the original models to maximize power, avoid the cost of losing data from an aging population, and increase the generalizability of our findings to the broader ToM literature. However, we strongly contend that new ToM instruments with strong psychometric properties be developed, particularly prior to the translation of ToM measures into mainstream clinical use. At the very least, researchers should carefully examine the psychometric properties of ToM measures and report on these findings, as we have done. This may assist in clarifying mixed results regarding the fundamental predictors of cognitive and affective ToM.

It can also be argued that our use PP represents a rudimentary assessment of vascular health. Using a simple in-office assessment we successfully identified a strong relationship between PP and ToM; however, we acknowledge that more rigorous physiological methods are needed to directly address the hypothesis that decreased arterial compliance underlies worse cognitive ToM (e.g., carotid–femoral pulse wave velocity or systolic pulse contour analysis; Waldstein et al., 2008). It is also difficult to assess whether PP is an independent risk factor for reduced cognitive ToM or rather a marker for general cardiovascular health. It is well established that blood pressure interacts with other CVRFs to exacerbate neurocognitive difficulties (e.g., type 2 diabetes, Hassing et al., 2004), and may also be influenced by treatment effects. Approximately one-third of our older adult participants reported treatment with anti-hypertensive medication. Thus it is possible that treatment effects may have influenced results either by diluting observed associations with cognitive ToM and potential associations with affective ToM, or via the direct impact of medication on neurocognitive performance (Gorelick et al., 2011). Future research comparing adults with and without anti-hypertensive treatment may be beneficial in addressing these areas.

Finally, we note that our young and older samples were relatively small for the single-group path analyses and only adequate to detect large effect sizes. Our final models comprised 86 young and 85 older adults, respectively; it is recommended that for accurate parameter estimation, path models should ideally comprise 100 or more participants per model tested (O'Rourke & Hatcher, 2013). Byrne (2004, 2013) describes several consequences of small sample size, including reduced precision of parameter estimates, greater standardized errors, and increased probability of obtaining improper solutions. Certain indices of absolute fit (i.e., RMSEA, SRMR) may also be positively biased in smaller samples (Hoyle, 2012). Testing a continuous age range rather than categorical age group predictor would allow for more meaningful conclusions the nature of age differences in ToM and how these may relate to fundamental predictors. We also note that measurement error can lead to lowered ability of statistical tests to provide accurate estimation. Despite these issues, we obtained robust findings that were in line with guiding theory. Given the novelty of research in this area, we felt our approach was appropriate for the sample size, and for our interests in identifying theoretically supported predictors of cognitive and affective ToM to follow-up with future research.

6.6. Future Directions

Research on ToM and aging is at an exciting stage where studies are emerging with the capability to combine neuropsychological, neuroscience, imaging, and developmental perspectives. Focus on development of novel psychometrically sound ToM measures as well as the cautious improvement of existing measures will be essential to the meaningfulness of research in this area. Key questions remain regarding the practical relevance of examining ToM in clinical settings and whether current instruments of ToM may be valuable markers of daily functioning (e.g., social engagement, decision-making, quality of life; Bailey et al., 2008; Yeh, 2013). In naturalistic environments our inferences about others' mental states in moment-to-moment interactions are not always neatly defined and may proceed at a lower level of awareness than implied by literature. Face valid tests should be developed that can account for other aspects of the social experience—including voice intonation, person or context familiarity, or judgments made in the presence of social groups. As evidenced by recent developments in social neuroscience and psychology, there is keen interest in

adopting ToM measures as clinical tools, particularly when one considers the large number of clinical disorders in which ToM difficulties are documented. Interventions for ToM difficulties may benefit from research such as that conducted in this dissertation to clarify the multidimensional nature of ToM and to identify what factors (e.g., linguistic skills, executive functions) may be targeted to enhance rehabilitation and development of ToM skills. In clinical neuropsychological evaluations, incorporation of psychometrically sound social cognitive measures, including cognitive and affective ToM, to standardized assessment batteries may aid in characterizing cognitive strengths and weaknesses, and assist in differential diagnosis (e.g., AD versus bvFTD; Le Bouc et al., 2012; see also Adenzato & Poletti, 2013). It is our hope that the current results provide a valuable point of departure for answering many of these questions.

References

Aboulafia-Brakha, T., Christe, B., Martory, M.-D., & Annoni, J-M. (2011). Theory of mind tasks and executive functions: A systematic review of group studies in neurology. *Journal of Neuropsychology, 5,* 39-55. doi:10.1348/174866410x533660

Adenzato & Poletti (2013). Theory of mind abilities in neurodegenerative diseases: An update and a call to introduce mentalizing tasks in standard neuropsychological assessments. *Clinical Neuropsychiatry, 10,* 226-234.

Adolphs, R. (2003). Cognitive neuroscience of human social behaviour. *Nature Reviews Neuroscience, 4,* 165-178. doi: 10.1038/nrn1056

Ahmed, F.S. & Miller, L.S. (2011). Executive function mechanisms of theory of mind. *Journal of Autism and Developmental Disorders, 41,* 667-678. doi:10.1007/s10803-010-1087-7

Ahmed, F.S. & Miller, L.S. (2013). Relationship between theory of mind and functional independence is mediated by executive functions. *Psychology & Aging, 28,* 293-303. doi:10.1037/a0031365

Aine, C.J., Sanfratello, L., Adair, J.C., Knoefel, J.E., Qualls, C., Lundy, S.L., Caprihan, A., . . . Stephen, J.M. (2014). Characterization of a normal control group: Are they healthy? *NeuroImage, 84,* 796-809. doi:10.1016/j.neuroimage.2013.09.025

Allain, P., Havet-Thomassin, V., Verny, C., Gohier, B., Lancelot, C., Besnard, J., Fasotti, L., & Le Gall, D. (2011). Evidence for deficits on different components of theory of mind in Huntington's disease. *Neuropsychology, 25,* 741-751. doi:10.1037/a0024408

American Psychiatric Association. (2013). *Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition, DSM-5*. Arlington, VA: American Psychiatric Association.

Anderson, K.M., Odell, P.M., Wilson, P.W.F., & Kannel, W.B. (1991). Cardiovascular disease risk profiles. *American Heart Journal, 121,* 293-298. doi10.1016/0002-8703(91)90861-b

Apperly, I.A., Samson, D., & Humphreys, G.W. (2005). Domain-specificity and theory of mind: Evaluating neuropsychological evidence. *Trends in Cognitive Sciences, 9,* 572-577. doi:10.1016/j.tics.2005.10.004

Astington, J.W. & Jenkins, J.M. (1999). A longitudinal study of the relation between language and theory-of-mind development. *Developmental Psychology, 35,* 1311-1320. doi: 10.1037/0012-1649.35.5.1311

Baglio, F., Castelli, I., Alberoni, M., Blasi, V., Griffanti, L., Falini, A., Nemni, R., & Marchetti, A. (2012). Theory of mind in amnestic mild cognitive impairment: An fMRI study. *Journal of Alzheimer's Disease, 29,* 25-37. doi:10.3233/JAD-2011-111256

Bailey, P.E. & Henry, J.D. (2008). Growing less empathic with age: Disinhibition of the self-perspective. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 63,* P219-P226. doi:10.1093/geronb/63.4.p219

Bailey, P.E., Henry, J.D., & von Hippel, W. (2008). Empathy and social functioning in late adulthood. *Aging & Mental Health, 12,* 499-503. doi:10.1080/13607860802224243

Baker, C.A., Peterson, E., Pulos, S., & Kirkland, R.A. (2014). Eyes and IQ: A meta-analysis of the relationship between intelligence and "Reading the Mind in the Eyes." *Intelligence, 44,* 78-92. doi:10.1016/j.intell.2014.03.001

Baron-Cohen, S. (2000). *Theory of mind and autism: A fifteen-year review*. In S. Baron-Cohen, H. Tager-Flusberg, and D.J. Cohen (Eds.), Understanding other minds: Perspectives from developmental cognitive neuroscience (2nd ed.). New York, NY: Oxford University Press.

Baron-Cohen, S., Wheelright, S., Hill, J., Raste, Y., & Plumb, I. (2001). The "Reading the Mind in the Eyes" Test revised version: A study with normal adults, and adults with Asperger syndrome or high-functioning autism. *Journal of Child Psychology and Psychiatry, 42,* 241-251. doi:10.1111/1469-7610.00715

Bernstein, D.M., Coolin, A., Fischer, A.L., Thornton, W.L., & Sommerville, J.A. (2015). *False belief reasoning from 3 to 92 years of age*. Manuscript in preparation.

Bernstein, D.M., Thornton, W.L., & Sommerville, J.A. (2011). Theory of mind through the ages: Older and middle-aged adults exhibit more errors than do younger adults on a continuous false belief task. *Experimental Aging Research, 37,* 481-502. doi:10.1080/0361073x.2011.619466

Bibby, H. & McDonald, S. (2005). Theory of mind after traumatic brain injury. *Neuropsychologia, 43,* 99-114. doi:10.1016/j.neuropsychologia.2004.04.027

Birns, J. & Kalra, L. (2009). Cognitive function and hypertension. *Journal of Human Hypertension, 23,* 86-96. doi:10.1038/jhh.2008.80

Bollen, K. A., & Stine, R. (1990). Direct and indirect effects: Classical and bootstrap estimates of variability. *Sociological Methodology*, *20*, 15-140.doi:10.2307/271084

Bosco, F.M., Gabbatore, I., & Tirassa. (2014). A broad assessment of theory of mind in adolescence: The complexity of mindreading. *Consciousness and Cognition, 24,* 84-97. doi:10.1016/j.concog.2014.01.003

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In: K. A. Bollen & J. S. Long (Eds.), Testing structural equation models (pp. 136-162). Newbury Park, CA: Sage Publications, Inc.

Burnett, A.C., Reutens, D.C., & Wood, A.G. (2010). Social cognition in Turner's syndrome. *Journal of Clinical Neuroscience, 17,* 283-286. doi:10.1016/j.jocn.2009.09.006

Byrne, B.M. (2004). Testing for multigroup invariance using AMOS Graphics: A road less traveled. *Structural Equation Modeling, 11,* 272-300. doi:10.1207/s15328007sem1102_8

Byrne, B.M. (2013). *Structural equation modeling with AMOS: Basic concepts, applications, and programming* (2nd ed.). New York, NY: Routledge.

Campbell, L.E., Stevens, A.F., McCabe, K., Cruikshank, L., Morris, R.G., Murphy, D.G.M., & Murphy, K.C. (2011). Is theory of mind related to social dysfunction and emotional problems in 22q11.2 deletion syndrome (velo-cardio-facial syndrome)? *Journal of Neurodevelopmental Disorders, 3,* 152-161. doi:10.1007/s11689-011-9082-7

Campbell, N.R.C., Joffres, M.R., & McKay, D.W. (2005). Hypertension surveillance in Canada. *Canadian Journal of Public Health, 96,* 217-220.

Carlson, S.M. & Moses, L.J. (2001). Individual differences in inhibitory control and children's theory of mind. *Child Development, 72,* 1032-1053. doi:10.1111/1467-8624.00333

Carruthers, P. & Smith, P.K. (1996). *Theories of theories of mind.* Cambridge, UK: Cambridge University Press.

Castelli, I., Pini, A., Alberoni, M., Liverto-Sempio, O., Baglio, F., Massaro, D., . . . Nemni, R. (2011). Mapping levels of theory of mind in Alzheimer's disease: A preliminary study. *Aging & Mental Health, 15,* 157-168. doi:10.1080/13607863.2010.513038

Castaneda, A.E., Tuulio-Henriksson, A., Marttunen, M., Suvisaar, J., & Lönnqvist, J. (2008). A review on cognitive impairments in depressive and anxiety disorders with a focus on young adults. *Journal of Affective Disorders, 106,* 1-27. doi: 10.1016/j.jad.2007.06.006

Cavallini, E., Lecce, S., Bottiroli, S., Palladino, P., & Pagnin, A. (2013). Beyond false belief: Theory of mind in young, young-old, and old-old adults. *International Journal of Aging and Human Development, 76,* 181-198. doi: 10.2190/AG.76.3.a

Charlton, R.A., Barrick, T.R., Markus, H.S., & Morris, R.G. (2009). Theory of mind associations with other cognitive functions and brain imaging in normal aging. *Psychology and Aging, 24,* 338-348. doi:10.1037/a0015225

Charman, T., Ruffman, T., & Clements, W. (2002). Is there a gender difference in false belief development? *Social Development, 11,* 1-10. doi:10.1111/1467-9507.00183

Chobanian, A.V., Bakris, G.L., Black, H.R., Cushman, W.C., Green, L.A., Izzo, J.L. Jr., Jones, D.W., . . . and the National High Blood Pressure Education Program Coordinating Committee. (2003). Seventh report of the Joint National Committee on Prevention, Detection, Evaluation, and Treatment of High Blood Pressure: The JNC 7 report. *Journal of the American Medical Association, 289,* 2560-2572. doi:10.1161/01.hyp.0000107251.49515.c2

Christov-Moore, L., Simpson, E.A., Coudé, G., Grigaityte, K., Iacoboni, M., & Ferrari, P.F. (2014). Empathy: Gender effects in brain and behaviour. *Neuroscience and Biobehavioral Reviews.* Advance online publication. doi:10.1016/j.neubiorev.2014.09.001

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences*. (2nd ed.). Hillsdale, NJ: Lawrence Earlbaum Associates.

Cornish, K., Burack, J.A., Rahman, A., Munir, F., Russo, N., & Grant, C. (2005). Theory of mind deficits in children with fragile X syndrome. *Journal of Intellectual Disability Research, 49,* 372-378. doi: 10.1111/j.1365-2788.2005.00678.x

Crawford, J.R., Sutherland, D., & Garthwaite, P.H. (2008). On the reliability and standard errors of measurement of contrast measures from the D-KEFS. *Journal of the International Neuropsychological Society, 14,* 1069-1073. doi:10.1017/s1355617708081228

Dahle, C.L., Jacobs, B.S., & Raz, N. (2009). Aging, vascular risk, and cognition: Blood glucose, pulse pressure, and cognitive performance in healthy adults. *Psychology and Aging, 24,* 154-162. doi: 10.1037/a0014283

Decety, J. & Jackson, P.L. (2004). The functional architecture of human empathy. *Behavioral and Cognitive Neuroscience Reviews, 3,* 71-100. doi:10.1177/1534582304267187

Delis, D.C., Kaplan, E., & Kramer, J.H. (2001). *Delis-Kaplan Executive Function System (DKEFS) Technical Manual.* San Antonio, TX, USA: The Psychological Corporation.

Delis, D., Kramer, J. A., Kaplan, E., & Ober, B. (2000). *The California Verbal Learning Test – 2nd Edition - Adult Version*. San Antonio, TX, USA: The Psychological Corporation.

Dolan, M. & Fullam, R. (2004). Theory of mind and mentalizing ability in antisocial personality disorders with and without psychopathy. *Psychological Medicine, 34,* 1093-1102. doi:10.1017/s0033291704002028

Drag, L.L. & Bieliauskas, L.A. (2010). Contemporary review 2009: Cognitive aging. *Journal of Geriatric Psychiatry and Neurology, 23,* 75-93. doi: 10.1177/0891988709358590

Duval, C., Bejanin, A., Piolino, P., Laisney, M., de La Sayette, V., Belliard, S., Eustache, F., & Desgranges, B. (2012). Theory of mind impairments in patients with semantic dementia. *Brain, 135,* 228-241. doi:10.1093/brain/awr309

Duval, C., Piolino, P., Bejanin, A., Eustache, F., & Desgranges, B. (2011). Age effects on different components of theory of mind. *Consciousness and Cognition, 20,* 627-642. doi:10.1016/j.concog.2010.10.025

Dvash, J. & Shamay-Tsoory, S.G. (2014). Theory of mind and empathy as multidimensional constructs: Neurological foundations. *Topics in Language Disorders, 34,* 282-295. doi: 10.1097/TLD.0000000000000040

Edgington, E.S. (1995). *Randomization tests: 3rd edition revised and expanded* (p.183). New York, NY: Marcel Dekker, Inc.

Elias, M.F., Elias, P.K., Sullivan, L.M., Wolf, P.A., & D'Agostino, R.B. (2003). Lower cognitive function in the presence of obesity and hypertension: The Framingham Heart Study. *International Journal of Obesity, 27,* 260-268. doi:10.1038/sj.ijo.802225

Elias, P.K., Elias, M.F., Robbins, M.A. & Budge, M.M. (2004). Blood pressure-related cognitive decline: Does age make a difference? *Hypertension, 44,* 631-636. doi:10.1161/01.HYP.0000145858.07252.99

Fernández-Abascal, E.G., Cabello, R., Fernández-Berrocal, P., & Baron-Cohen, S. (2013). Test-retest reliability of the 'Reading the Mind in the Eyes' test: A one-year follow-up study. *Molecular Autism, 4,* 1-6. doi:10.1186/2040-2392-4-33

Fett, A.-K.J., Viechtbauer, W., Dominguez, M.-d.-G., Penn, D.L., van Os, J., & Krabbendam, L. (2011). The relationship between neurocognition and social cognition with functional outcomes in schizophrenia: A meta-analysis. *Neuroscience and Biobehavioral Reviews, 35,* 573-588. doi:10.1016/j.neubiorev.2010.07.001

Finney, J.M. (1972). Indirect effects in path analysis. *Sociological Methods & Research, 1,* 175-186. doi:10.1177/004912417200100202

Fischer, A.L., Bernstein, D.M., & Thornton, W.L. (2014). Vascular health modifies theory of mind performance in older adults. *Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 69,* 219-227. doi:10.1093/geronb/gbs120

Folstein, M.F., Folstein, S.E., & McHugh, P.R. (1975). Mini-mental state. A practical method for grading the cognitive state of patients for the clinician. *Journal of* Psychiatric Research, 12, 189-198. doi:10.1016/0022-3956(75)90026-6

Franco, M.G. & Smith, P.K. (2014). Theory of mind, old age, and educational opportunity in Columbia. International Journal of Aging & Human Development, 77, 107-125. doi: 10.2190/AG.77.2.b

Franklin, S.S. (2004). Pulse pressure as a risk factor. *Clinical and Experimental Hypertension,* 26, 645-652. doi:10.1081/ceh-200031962

Gallese, V. & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences, 2,* 493-501. doi: 10.1016/S1364-6613(98)01262-5

German, T.P. & Hehman, J.A. (2006). Representational and executive selection resources in 'theory of mind': Evidence from compromised belief-desire reasoning in old age. *Cognition, 101,* 129-152. doi:10.1016/j.cognition.2005.05.007

Gerrard, P., Martin, N.H., Giunti, P., & Cipolotti, L. (2008). Cognitive and social cognitive functioning in spinocerebellar ataxia. *Journal of Neurology, 255,* 398-405.doi:10.1007/s00415-008-0680-6

Giaouri, S., Alevriadou, A., & Tsakiridou, E. (2010). Theory of mind abilities in children with Down syndrome and non-specific intellectual disabilities: An empirical study with some educational implications. *Procedia – Social and Behavioral Sciences, 2,* 3883-3887. doi:10.1016/j.sbspro.2010.03.609

Gifford, K.A., Badaracco, M., Liu, D., Tripodis, Y., Gentile, A., Lu, Z., Palmisano, J., & Jefferson, A.L. (2013). Blood pressure and cognition among older adults: A meta-analysis. *Archives of Clinical Neuropsychology, 28,* 649-664. doi:10.1093/arclin/act046

Gillberg, C., Cederlund, M., Lamberg, K., & Zeijlon, L. (2006). Brief report: "The Autism epidemic". The registered prevalence of autism in a Swedish urban area. *Journal of Autism and Developmental Disorders, 36,* 429-435. doi:10.1007/s10803-006-0081-6

Goetz, P.J. (2003). The effects of bilingualism on theory of mind development. *Bilingualism: Language and Cognition, 6,* 1-15. doi:10.1017/s1366728903001007

Gorelick, P.B., Scuteri, A., Black, S.E., DeCarli, C., Greenberg, S.M., Iadecola, C., Launer, L.J., . . .on behalf of the American Heart Association Stroke Council, Council on Epidemiology and Prevention, Council on Cardiovascular Nursing, Council on Cardiovascular Radiology and Intervention, and Council on Cardiovascular Surgery and Anesthesia. (2011). Vascular contributions to cognitive impairment and dementia: A statement for healthcare professionals from the American Heart Association/American Stroke Association. *Stroke, 42,* 2672-2713. doi:10.1161/str.0b013e3182299496

Gunstad, J., Paul, R.H., Cohen, R.A., Tate, D.F., Spitznagel, M.B., & Gordon, E. (2007). Elevated body mass index is associated with executive dysfunction in otherwise healthy adults. *Comprehensive Psychiatry, 48,* 57-61. doi:10.1016/j.comppsych.2006.05.001

Hallerbäck, M.U., Lunegård, T., Hjärthag, F., & Gillberg, C. (2009). The Reading the Mind in the Eyes Test: Test-retest reliability of a Swedish version. *Cognitive Neuropsychiatry, 14,* 127-143. doi:10.1080/13546800902901518

Happé, F.G.E. (1994). An advanced test of theory of mind: Understanding of story characters' thought and feelings by able autistic, mentally handicapped, and normal children and adults. *Journal of Autism and Developmental Disorders, 24,* 129-154. doi:10.1007/bf02172093

Happé, F.G.E., Winner, E., & Brownell, H. (1998). The getting of wisdom: Theory of mind in old age. *Developmental Psychology, 34,* 358-362. doi:10.1037//0012-1649.34.2.358

Harkness, K.L., Sabbagh, M.A., Jacobson, J.A., Chowdrey, N.K., & Chen, T. (2005). Enhanced accuracy of mental state decoding in dysphoric college students. *Cognition and Emotion, 19,* 999-1025. doi:10.1080/02699930541000110

Hasher, L., & Zacks, R.T. (1988). Working memory, comprehension, and aging: A review and a new view. In G.H. Bower (Ed.), *The psychology of learning and motivation* (Vol. 22) (pp. 193-225). New York, USA: Academic Press.

Hassing, L.B., Hofer, S.M., Nilsson, S.E., Berg, S., Pedersen, N.L., McClearn, G., & Johansson, B. (2004). Comorbid type 2 diabetes mellitus and hypertension exacerbates cognitive decline: Evidence from a longitudinal study. *Age and Ageing, 33,* 355-61. doi:10.1093/ageing/afh100

Heaton, R.K., Ryan, L., & Grant, I. (2009). Demographic influences and use of demographically corrected norms in neuropsychological assessment. In I. Grant & K.M. Adams (Eds.), *Neuropsychological assessment of neuropsychiatric and neuromedical disorders* (3rd Ed.) (pp. 127-155). New York, USA: Oxford University Press.

Henry, J.D., MacLeod, M.S., Phillips, L.H., & Crawford, J.R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging, 19,* 27-39. doi:10.1037/0882-7974.19.1.27

Henry, J.D., Phillips, L.H., Ruffman, T., & Bailey, P.E. (2013). A meta-analytic review of age differences in theory of mind. *Psychology and Aging, 28,* 826-839. doi:10.1037/a0030677

Henry, J.D., Phillips, L.H., & von Hippel, C. (2014). A meta-analytic review of theory of mind difficulties in behavioural-variant frontotemporal dementia. *Neuropsychologia, 56,* 53-62. doi:10.1016/j.neuropsychologia.2013.12.024

Henson, R. K. (2001). Understanding internal consistency estimates: A conceptual primer on coefficient alpha. *Measurement and Evaluation in Counseling and Development, 34,* 177-189.

Hoyle, R.H. (2012). *Handbook of structural equation modelling*. New York, NY: The Guilford Press.

Ibanez, A., Huepe, D., Gempp, R., Gutiérrez, V., Rivera-Rei, A., & Toledo, M.I. (2013). Empathy, sex and fluid intelligence as predictors of theory of mind. *Personality and Individual Differences, 54,* 616-621. doi:10.1016/j.paid.2012.11.022

IBM Corp. (2012). IBM SPSS Amos, Version 21.0 [Computer Program]. Armonk, NY: IBM Corp.

IBM Corp. (2013). IBM SPSS Statistics for Windows and MacIntosh, Version 22.0 [Computer Program]. Armonk, NY: IBM Corp.

Joosten, H., van Eersel, M.E.A., Gansevoort, R.T., Bilo, H.J.G., Slaets, J.P.J., & Izaks, G.J. (2013). Cardiovascular risk profile and cognitive function in young, middle-aged, and elderly subjects. *Stroke, 44,* 1543-1549. doi:10.1161/STROKEAHA.111.000496

Jorm, A.F. (2000). Does old age reduce the risk of anxiety and depression? A review of epidemiological studies across the adult lifespan. *Psychological Medicine, 30,* 11-22. doi:10.1017/s0033291799001452

Kaufman, A. S., & Kaufman, N. L. (2004). Kaufman Brief Intelligence Test, Second Edition. Bloomington, MN: Pearson, Inc.

Kearney, P.M., Whelton, M., Reynolds, K., Munter, P., Whelton, P.K., & He, J. (2005). Global burden of hypertension: Analysis of worldwide data. *The Lancet, 365,* 217-223. doi:10.1016/s0140-6736(05)70151-3

Keightley, M.L., Winocur, G., Burianova, H., Hongwanishkul, D., & Grady, C.L. (2006). Age effects on social cognition: Faces tell a different story. *Psychology and Aging, 21,* 558-572.doi:10.1037/0882-7974.21.3.558

Kemp, J., Després, O., Sellal, F., &Dufour, A. (2012). Theory of mind in normal ageing and neurodegenerative pathologies. *Ageing Research Reviews, 11,* 199-219. doi:10.1016/j.arr.2011.12.001

Kerr, N., Dunbar, R.I.M., & Bentall, R.P. (2003). Theory of mind deficits in bipolar affective disorder. *Journal of Affective Disorders, 73,* 253-259. doi:10.1016/s0165-0327(02)00008-3

Khattar, R.S., Swales, J.D., Dore, C., Senior, R., &Lahiri, A. (2001). Effect of aging on the prognostic significance of ambulatory systolic, diastolic, and pulse pressure in essential hypertension. *Circulation, 104,* 783-789. doi:10.1161/hc3201.094227

Kirkland, R.A., Peterson, E., Baker, C.A., Miller, S., & Pulos, S. (2013). Meta-analysis reveals adult female superiority in "Reading the Mind in the Eyes Test." *North American Journal of Psychology, 15,* 121-146.

Koelkebeck, K., Hirao, K., Kawada, R., Miyata, J., Saze, T., Ubukata, A., . . . Murai, T. (2011). Transcultural differences in brain activation patterns during theory of mind (ToM) task performance in Japanese and Caucasian participants. *Social Neuroscience, 6,* 615-626. doi:10.1080/17470919.2011.620763

Koenig, K., Klin, A., & Schultz, R. (2004) Deficits in social attribution ability in Prader-Willi syndrome. *Journal of Autism and Developmental Disorders, 34,* 573-582. doi:10.1007/s10803-004-2551-z

Kraemer, H.C., Stice, E., Kazdin, A., Offord, D., & Kupfer, D. (2001). How do risk factors work together? Mediators, moderators, and independent, overlapping, and proxy risk factors. *American Journal of Psychiatry, 158,* 848-856. doi:10.1176/appi.ajp.158.6.848

Laisney, M., Bon, L., Guiziou, C., Daluzeau, N., Eustache, F., & Desgranges, B. (2013). Cognitive and affective theory of mind in mild to moderate Alzheimer's disease. *Journal of Neuropsychology, 7,* 107-120. doi:10.1111/j.1748-6653.2012.02038.x

Le Bouc, R., Lenfant, P., Delbeuck, X., Ravasi, L., Lebert, F., Semah, F., & Pasquier, F. (2012). My belief or yours? Differential theory of mind deficits in frontotemporal dementia and Alzheimer's disease. *Brain, 135,* 3026-3038. doi:10.1093/brain/aws237

Lemay, S., Bédard, M.-A., Rouleau, I., & Tremblay, P.-L.G. (2004). Practice effect and test-retest reliability of attention and executive tests in middle-aged to elderly subjects. *The Clinical Neuropsychologist, 18,* 284-302. doi:10.1080/13854040490501718

Li, X., Wang, K., Wang, F., Tao, Q., Xie, Y., & Cheng, Q. (2013). Aging of theory of mind: The influence of educational level and cognitive processing. *International Journal of Psychology, 48,* 715-727. doi:10.1080/00207594.2012.673724

MacPherson, S.E., Phillips, L.H., & Della Sala, S. (2002). Age, executive function, and social decision making: A dorsolateral prefrontal theory of cognitive aging. *Psychology and Aging, 17,* 598-609. doi:10.1037//0882-7974.17.4.598

Mahy, C.E.V., Moses, L.J., & Pfeifer, J.H. (2014a). How and where: Theory-of-mind in the brain. *Developmental Cognitive Neuroscience, 9,* 68-81. doi: 10.1016/j.dcn.2014.01.002

Mahy, C.E.V., Vetter, N., Kühn-Popp, N., Löcher, C., Krautschuk, S., & Kliegel, M. (2014b). The influence of inhibitory processes on affective theory of mind in young and old adults. *Aging, Neuropsychology, and Cognition, 21,* 129-145. doi:10.1080/13825585.2013.789096

Maylor, E.A., Moulson, J.M., Muncer, A.-M., & Taylor, L.A. (2002). Does performance on theory of mind tasks decline in old age? *British Journal of Psychology, 93,* 465-485. doi:10.1348/000712602761381358

McFall, G.P., Wiebe, S.A., Vergote, D., Jhamandas, J., Westaway, D., & Dixon, R.A. (2014). IDE (rs6853817) polymorphism and pulse pressure are independently and interactively associated with level and change and executive function in older adults. *Psychology and Aging, 29,* 418-430. doi:10.1037/a0034656

McKinnon, M.C. & Moscovitch, M. (2007). Domain-general contributions to social reasoning: Theory of mind and deontic reasoning explored. *Cognition, 102,* 179-218. doi:10.1016/j.cognition.2005.12.011

Meyers, L.S., Gamst, G.C., & Guarino, A.J. (2013). *Performing data analysis using IBM SPSS.* Hoboken, NJ, USA: Wiley.

Moreau, N., Viallet, F., & Champagne-Lavau, M. (2013). Using memories to understand others: The role of episodic memory in theory of mind impairment in Alzheimer disease. *Ageing Research Reviews, 12,* 833-839. doi:10.1016/j.arr.2013.06.005

Moreau, N., Rauzy, S., Bonnefoi, B., Renié, L., Martinez-Almoyna, L., Viallet, F., & Champagne-Lavau, M. (2015). Different patterns of theory of mind impairment in mild cognitive impairment. *Journal of Alzheimer's Disease*. Advanced online publication. doi: 10.3233/JAD-143021

Moran, J. M. (2013). Lifespan development: The effects of typical aging on theory of mind. *Behavioural Brain Research, 237,* 32-40. doi:10.1016/j.bbr.2012.09.020

Nagai, M., Hoshide, S., & Kario, K. (2010). Hypertension and dementia. *American Journal of Hypertension, 23,* 116-124. doi:10.1038/ajh.2009.212

Narme, P., Mouras, H., Roussel, M., Devendeville, A., & Godefroy, O. (2013a). Assessment of socioemotional processes facilitates the distinction between frontotemporal lobar disease and Alzheimer's disease. *Journal of Clinical and Experimental Neuropsychology, 35,* 728-744. doi:10.1080/13803395.2013.823911

Narme, P., Mouras, H., Roussel, M., Duru, C., Krystkowiak, P., & Godefroy, O. (2013b). Emotional and cognitive social processes are impaired in Parkinson's disease and are related to behavioral disorders. *Neuropsychology, 27,* 182-192. doi: 10.1037/a0031522

Nichols, W. W., O'Rourke, M. F., & Vlachopouolos, C. (2011). McDonald's blood flow in arteries: Theoretical, experimental and clinical principles (6th ed.). Oxford, UK: Oxford University Press.

O'Brien, E., Konrath, S.H., Grühn, D., & Hagen, A.L. (2013). Empathic concern and perspective taking: Linear and quadratic effects of age across the adult life span. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 68,* 168-175. doi:10.1093/geronb/gbs055

O'Brien, J.T., Wiseman, R., Burton, E.J., Barber, B., Wesnes, K., Saxby, B., & Ford, G.A. (2002). Cognitive associations of subcortical white matter lesions in older people. *Annals of the New York Academy of Sciences, 977,* 436-44. doi:10.1111/j.1749-6632.2002.tb04849.x

O'Rourke, N., & Hatcher, L. (2013). *A Step-by-Step Approach to Using SAS for Factor Analysis and Structural Equation Modeling* (2nd Ed.). Cary, NC: SAS Institute.

O'Rourke, N., Hatcher, L., & Stepanski, E. J. (2005). *A Step-by-Step Approach to Using SAS for Univariate and Multivariate Statistics* (2nd Ed.). New York, NY: John Wiley & Sons.

Pardini, M. & Nichelli, P.F. (2009). Age-related decline in mentalizing skills across adult life span. *Experimental Aging Research*, *35*, 98–106. doi:10.1080/03610730802545259

Paterson, T.S.E., O'Rourke, N., Elmer, E., Shapiro, R.J., & Thornton, W.L. (2011). The composition and structure of depressive symptomatology in renal disease. *Canadian Journal of Behavioural Science, 43,* 318-327. doi: 10.1037/a0025104

Peterson, E. & Miller, S.F. (2012). The eyes test as a measure of individual differences: How much of the variance reflects verbal IQ? *Frontiers in Psychology, 3,* 1-6. doi: 10.3389/fpsyg.2012.00220

Pfaltz, M.C., McAleese, S., Saladin, A., Meyer, A.H., Stoecklin, M., Opwis, K., Dammann, G., & Martin-Soelch, C. (2013). The Reading the Mind in the Eyes Test: Test-retest reliability and preliminary psychometric properties of the German version. *International Journal of Advances in Psychology, 2,* 1-9.

Phillips, L.H., Bull, R., Allen, R., Insch, P., Burr, K. & Ogg, W. (2011). Lifespan aging and belief reasoning: Influences of executive function and social cue decoding. *Cognition, 120,* 236-247. doi:10.1016/j.cognition.2011.05.003

Phillips, L.H., MacLean, R.D.J., & Allen, R. (2002). Age and the understanding of emotions: Neuropsychological and sociocognitive perspectives. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences, 57,* P526-P530. doi:10.1093/geronb/57.6.p526

Plehn, K., Marcopulos, B.A., & McLain, C.A. (2004). The relationship between neuropsychological performance, social functioning, and activities of daily living in a sample of rural older adults. *The Clinical Neuropsychologist, 18,* 101-113. doi:10.1080/13854040490507190

Pöttgen, J., Dziobek, I., Reh, S., Heesen, C., & Gold., S.M. (2013). Impaired social cognition in multiple sclerosis. *Journal of Neurology, Neurosurgery, & Psychiatry, 84,* 529-536. doi:10.1136/jnnp-2012-304157

Poletti, M., Enrici, I., Bonuccelli, U., & Adenzato, M. (2011). Theory of mind in Parkinson's disease. *Behavioural Brain Research, 219,* 342-350. doi:10.1016/j.bbr.2011.01.010

Premack, D. & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? *Behavioral and Brain Sciences, 1,* 515–526. doi:10.1017/s0140525x00076512

Public Health Agency of Canada. (2010). Report from the Canadian Chronic Disease Surveillance System: Hypertension in Canada, 2010. Retrieved June 19, 2014 from http://www.phac-aspc.gc.ca/cd-mc/cvd-mcv/ccdss-snsmc-2010/pdf/CCDSS_HTN_Report_FINAL_EN_20100513.pdf. doi:10.1037/e614992012-001

Qiu, C., Winblad, B., & Fratiglioni, L. (2005). The age-dependent relation of blood pressure to cognitive function and dementia. *The Lancet Neurology, 4,* 487-499. doi:10.1016/s1474-4422(05)70141-1

Rabin, J.S. & Rosenbaum, R.S. (2012). Familiarity modulates the functional relationship between theory of mind and autobiographical memory. *NeuroImage, 62,* 520-529. doi:10.1016/j.neuroimage.2012.05.002

Radloff, L.S. (1977). The CES-D scale: A self-report depression scale for research in the general population. *Applied Psychological Measurement, 1,* 385-401. doi:10.1177/014662167700100306

Rakoczy, H., Harder-Kasten, A., & Sturm, L. (2012). The decline of theory of mind in older age is (partly) mediated by developmental changes in domain-general abilities. *British Journal of Psychology, 103,* 58-72 .doi:10.1111/j.2044-8295.2011.02040.x

Raz, N., Rodrigue, K. M., & Acker, J. D. (2003). Hypertension and the brain: Vulnerability of the prefrontal regions and executive functions. *Behavioral Neuroscience, 117,* 1169–1180. doi:10.1037/0735-7044.117.6.1169

Reynolds, M. W. (1999) *Multidimensional Anxiety Questionnaire Professional Manual.* Odessa, FL, USA: Psychological Assessment Resources.

Roper, B.L., Bieliauskas, L.A., & Peterson, M.R. (1996). Validity of the Mini-Mental State Examination and the Neurobehavioral Cognitive Status Examination in cognitive screening. *Neuropsychiatry, Neuropsychology, & Behavioral Neurology, 9,* 54-57.

Russell, T.A., Schmidt, U., Doherty, L., Young, V., & Tchanturia, K. (2009). Aspects of social cognition in anorexia nervosa: Affective and cognitive theory of mind. *Psychiatry Research, 168,* 181-185. doi:10.1016/j.psychres.2008.10.028

Russell, T.A., Tchanturia, K., Rahman, Q., & Schmidt, U. (2007). Sex differences in theory of mind: A male advantage on Happé's "cartoon" task. *Cognition and Emotion, 21,* 1554-1564. doi:10.1080/02699930601117096

Salthouse, T.A. (2009). Decomposing age correlations on neuropsychological and cognitive variables. *Journal of the International Neuropsychological Society, 15,* 650-661. doi: 10.1017/S1355617709990385

Salthouse, T.A. (2010). Selective review of cognitive aging. *Journal of the International Neuropsychological Society, 16,* 745-760. doi: 10.1017/S1355617710000706

Saltzman, J., Strauss, E., Hunter, M., & Archibald, S. (2000). Theory of mind and executive functions in normal human aging and Parkinson's disease. *Journal of the International Neuropsychological Society, 6,* 781-788. doi:10.1017/s1355617700677056

Sandoz, M., Démonet, J.F., & Fossard, M. (2014). Theory of mind and cognitive processes in aging and Alzheimer type dementia: A systematic review. *Aging & Mental Health, 18,* 815-827. doi:10.1080/13607863.2014.899974.

Saris, W.E., Satorra, A., & Sörbom, D. (1987). The detection and correction of specification errors in structural equation models. *Sociological Methodology, 17,* 105-129. doi:10.2307/271030

Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., Schmidt-Wilcke, T., & Brüne, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping, 36,* 29-39. doi: 10.1002/hbm.22610

Schneider, W., Eschman, A., & Zuccolotto, A. (2002). *E-Prime Reference Guide*. Pittsburgh: Psychology Software Tools Inc.

Scuteri, A., Nilsson, P.M., Tzurio, C., Redon, J., & Laurent, S. (2011). Microvascular brain damage with aging and hypertension: Pathophysiological consideration and clinical implications. *Journal of Hypertension, 29,* 1469-1477. doi:10.1097/hjh.0b013e328347cc17

Sesso, H.D., Cook, N.R., Buring, J.E., Manson, J.E., & Gaziano, J.M. (2008). Alcohol consumption and the risk of hypertension in women and men. *Hypertension, 51,* 1080-1087. doi:10.1161/HYPERTENSIONAHA.107.104968

Shamay-Tsoory, S.G. & Aharon-Peretz, J. (2007). Dissociable prefrontal networks for cognitive and affective theory of mind: A lesion study. *Neuropsychologia, 45,* 3054-3067. doi:10.1016/j.neuropsychologia.2007.05.021

Shamay-Tsoory, S.G., Aharon-Peretz, J., & Perry, D. (2009). Two systems for empathy: A double dissociation between emotional and cognitive empathy in inferior frontal gyrus versus ventromedial prefrontal lesions. *Brain, 132,* 617-627. doi: 10.1093/brain/awn279

Shamay-Tsoory, S.G., Tibi-Elhanany, Y., & Aharon-Peretz, J. (2007). The green-eyed monster and malicious joy: The neuroanatomical basis of envy and gloating (schadenfreude). *Brain, 130,* 1663-1678. doi: 10.1093/brain/awm093

Shamay-Tsoory, S.G., Tomer, R., Berger, B.D., Goldsher, D., & Aharon-Peretz, J. (2005). Impaired “affective theory of mind” is associated with right ventromedial prefrontal damage. *Cognitive & Behavioral Neurology, 18,* 55-67. doi: 10.1097/01.wnn.0000152228.90129.99

Shany-Ur, T. & Rankin, K.P. (2011). Personality and social cognition in neurodegenerative disease. *Current Opinion in Neurology, 24,* 550-555. doi:10.1097/wco.0b013e32834cd42a

Shrout, P.E. & Fleiss, J.L (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin, 86,* 420-428. doi: 10.1037/0033-2909.86.2.420

Singer, J., Trollor, J.N., Baune, B.T., Sachdev, P.S., & Smith, E. (2014). Arterial stiffness, the brain, and cognition: A systematic review. *Ageing Research Reviews, 15,* 16-27. doi:10.1016/j.arr.2014.02.002

Slessor, G., Phillips, L.H., & Bull, R. (2007). Exploring the specificity of age-related differences in theory of mind tasks. *Psychology and Aging, 22,* 639-643. doi:10.1037/0882-7974.22.3.639

Söderstrand, P. & Almkvist, O. (2012). Psychometric data on the Eyes Test, the Faux Pas Test, and the Dewey Social Stories Test in a population-based Swedish sample. *Nordic Psychology, 64,* 30-43. doi:10.1080/19012276.2012.693729

Stone, V.E., Baron-Cohen, S., & Knight, R.T. (1998). Frontal lobe contributions to theory of mind. *Journal of Cognitive Neuroscience, 10,* 640-656. doi:10.1162/089892998562942

Stone, V.E. & Gerrans, P. (2006). What's domain-specific about theory of mind? Social Neuroscience, 1, 309-319. doi: 10.1080/17470910601029221

Sullivan, S., & Ruffman, T. (2004). Social understanding: How does it fare with advancing years? *British Journal of Psychology, 95,* 1–18. doi:10.1348/000712604322779424

Tabachnick, B.G. & Fidell, L.S. (2013). Using multivariate statistics (6th Edition). Boston, MA: Pearson, Inc.

Tager-Flusberg, H. & Sullivan, K. (2000). A componential view of theory of mind: Evidence from Williams syndrome. *Cognition, 76,* 59-90. doi:10.1016/s0010-0277(00)00069-x

Thomas, G., & Maio, G. R. (2008). Man, I feel like a woman: when and how gender-role motivation helps mind-reading. *Journal of Personality and Social Psychology*, *95*, 1165-1179. doi:10.1037/a0013067

Thornton, W.L., Deria, S., Gelb, S., Shapiro, R.J., & Hill, A. (2007). Neuropsychological mediators of the links among age, chronic illness, and everyday problem solving. *Psychology and Aging, 22,* 470-481. doi:10.1037/0882-7974.22.3.470

Thorvaldsson, V., Skoog, I., Hofer, S.M., Börjesson-Hanson, A., Östling, S., Sacuiu, S., & Johansson, B. (2012). Nonlinear blood pressure effects on cognition in old age: Separating between-person and within-person associations. *Psychology and Aging, 27,* 375-383. doi:10.1037/a0025631

Vellante, M., Baron-Cohen, C., Melis, M., Marrone, M., Petretto, D.R., Masala, C., & Preti, A. (2013). The "Reading the Mind in the Eyes" test: Systematic review of psychometric properties and a validation study in Italy. *Cognitive Neuropsychiatry, 18,* 326-354. doi:10.1080/13546805.2012.721728

Vetter, N.C., Altgassen, M., Phillips, L., Mahy, C.E.V., & Kliegel, M. (2013). Development of affective theory of mind across adolescence: Disentangling the role of executive functions. *Developmental Neuropsychology, 38,* 114-125. doi:10.1080/87565641.2012.733786

Waldstein, S.R. (1995). Hypertension and neuropsychological function: A lifespan perspective. *Experimental Aging Research, 21,* 321-252. doi: 10.1080/03610739508253989

Waldstein, S.R. (2003). The relation of hypertension to cognitive function. *Current Directions in Psychological Science, 12,* 9-12. doi:10.1111/1467-8721.01212

Waldstein, S.R., Giggey, P.P., Thayer, J.F., & Zonderman, A.B. (2005). Nonlinear relations of blood pressure to cognitive function: The Baltimore Longitudinal Study of Aging. *Hypertension, 45,* 374-379. doi:10.1161/01.HYP.0000156744.44218.74

Waldstein, S. R., Rice, S. C., Thayer, J. F., Najjar, S. S., Scuteri, A., & Zonderman, A. B. (2008). Pulse pressure and pulse wave velocity are related to cognitive decline in the Baltimore Longitudinal Study of Aging. *Hypertension, 51,* 99–104. doi:10.1161/hypertensionaha.107.093674

Wang, Z. & Su, Y. (2013). Age-related differences in the performance of theory of mind in older adults: A dissociation of cognitive and affective components. *Psychology and Aging, 28,* 284–291. doi:10.1037/a0030876

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale–3rd Edition.* San Antonio, TX: Psychological Corporation.

Wellman, H.M., & Lagattuta, K.H. (2000). Developing understandings of mind. In S. Baron-Cohen, H. Tager-Flusberg, & D.J. Cohen (Eds.) *Understanding Other Minds: Perspectives from Developmental Cognitive Neuroscience* (2nd ed.) (pp. 21-49). New York, USA: Oxford University Press.

Wilkins, K., Campbell, N.R.C., Joffres, M.R., McAlister, F.A., Nichol, M., Quach, S., Johansen, H.L., & Tremblay, M.S. (2010). Blood pressure in Canadian adults. *Statistics Canada Health Reports, 21,* 1-10. Catalogue no. 82-003-XPE

Wilkins, K., Gee, M., & Campbell, N.R.C. (2012). The difference in hypertension control between older men and women. *Statistics Canada Health Reports, 23,* 1-10. Catalogue no. 82-003-XPE

Wilson, R.S., Beckett, L.A., Barnes, L.L., Schneider, J.A., Back, J., Evans, D.A., & Bennett, D.A. (2002). Individual differences in rates of change in cognitive abilities of older persons. *Psychology & Aging, 17,* 179-193. doi: 10.1037/0882-7974.17.2.179

Wolkenstein, L., Schönenberg, M., Schirm, E., & Hautzinger, M. (2011). I can see what you feel, but I can't deal with it: Impaired theory of mind in depression. *Journal of Affective Disorders, 132,* 104-111. doi: 10.1016/j.jad.2011.02.010

Woods, S.P., Delis, D.C., Scott, J.C., Kramer, J.H., & Holdnack, J.A. (2006). The California Verbal Learning Test – second edition: Test-retest reliability, practice effects, and reliable change indices for the standard and alternate forms. *Archives of Clinical Neuropsychology, 21,* 413-420. doi:10.1016/j.acn.2006.06.002

World Health Organization. (2009). Global health risks: Mortality and burden of disease attributable to selected major risks. Geneva, Switzerland: World Health Organization Press.

Yaffe, K., Barrett-Connor, E., Lin, F., & Grady, D. (2002). Serum lipoprotein levels, statin use, and cognitive function in older women. *Archives of Neurology, 59,* 378-384. doi:10.1001/archneur.59.3.378

Yaffe, K., Vittinghoff, E., Pletcher, M.J., Hoang, T.D., Launer, L.J., Whitmer, R.A., . . .Sidney, S. (2014). Early adult to midlife cardiovascular risk factors and cognitive function. *Circulation, 129,* 1560-1567. doi:10.1161/circulationaha.113.004798

Yasar, S., Ko, J.Y., Nothelle, S., Mielke, M.M., & Carlson, M.C. (2011). Evaluation of the effect of systolic blood pressure and pulse pressure on cognitive function: The Women's Health and Aging Study II. *PLoS ONE, 6,* e27976. doi:10.1001/archneur.59.3.378

Yeh, Z.-T. (2013). Role of theory of mind and executive function in explaining social intelligence: A structural equation modeling approach. *Aging and Mental Health, 17,* 527-534. doi:10.1080/13607863.2012.758235

Yeung, S.E., Fischer, A.L., & Dixon, R.A. (2009). Exploring effects of type 2 diabetes on cognitive functioning in older adults. *Neuropsychology, 23,* 1-9. doi:10.1037/a0013849

Yeung, S.E. & Thornton, W.L. (2011). Age-related effects of blood pressure on everyday cognitive function in community-dwelling women. *Aging, Neuropsychology, and Cognition, 18,* 733-755. doi:10.1080/13825585.2011.609882

Ze, O., Thoma, P., & Suchan, B. (2014). Cognitive and affective empathy in younger and older individuals. *Aging & Mental Health, 7,* 929-935. doi10.1080/13607863.2014.899973

Zhu, J., Tulsky, D.S., Price, L., & Chen, H.-Y. (2001). WAIS-III reliability data for clinical groups. *Journal of the International Neuropsychological Society, 7,* 862-866. doi:10.1093/arclin/14.8.724

Zimmerman, M.A. & Sullivan, J.C. (2013). Hypertension: What's sex got to do with it? *Physiology, 28,* 234-244. doi:10.1152/physiol.00013.2013

Appendices

Appendix A.

Clinical Relevance of ToM

Table A 1 Clinical Disorders Associated with Alterations in ToM

Clinical Group	Select reference
Neurological disorders	
Alzheimer's disease	Laisney et al. (2013)
Behavioural variant frontotemporal dementia	Henry, Phillips, & von Hippel (2014)
Huntington's disease	Allain et al. (2011)
Mild cognitive impairment (amnestic type)	Moreau et al. (2015)
Multiple sclerosis	Pöttgen, Dziobek, Reh, Heesen, & Gold (2013)
Parkinson's disease	Poletti, Enrici, Bonuccelli, & Adenzato (2011)
Semantic dementia	Duval et al. (2012)
Traumatic brain injury (acquired)	Bibby & McDonald (2005)
Psychiatric and personality disorders	
Anorexia nervosa	Russell, Schmidt, Doherty, Young, & Tchanturia (2009)
Antisocial personality disorder & psychopathy	Dolan & Fullam (2004)
Autism spectrum disorders*	Baron-Cohen (2000)
Bipolar disorder	Kerr, Dunbar, & Bentall (2003)
Major depressive disorder	Wolkenstein et al. (2011)
Schizophrenia	Fett et al. (2011)
Genetic disorders	
22q11.2 deletion (velo-cardio-facial) syndrome	Campbell et al. (2011)
Down's syndrome	Giaouri, Alevriadou, & Tsakiridou (2010)
Fragile X syndrome	Cornish et al. (2005)
Prader-Willi syndrome	Koenig, Klin, & Schultz (2004)
Spinocerebellar ataxia	Garrard, Martin, Giunti, & Cipolotti (2008)
Turner syndrome	Burnett, Reutens, & Wood (2010)
Williams syndrome	Tager-Flusberg & Sullivan (2000)

Note: We provide full citations for each reference in the reference list. *Includes DSM-IV Asperger's Syndrome.

Appendix B.

Reliability of ToM Tests in Aging Research

Table B 1 Summary of ToM Studies in Healthy Aging and Reported Reliability

Study	ToM Test	Cognitive/Affective	Reliability
Ahmed & Miller (2013)	STORIES	Cognitive	Unreported
	Other: FP	Cognitive	
Bailey & Henry (2008)	RMET	Affective	Unreported
	Other: FB	Cognitive	
Bernstein et al. (2011)	Other: FB	Cognitive	Unreported
Castelli et al. (2010)	STORIES	Cognitive	Unreported
	RMET	Affective	
	Other: FB	Cognitive	
Castelli et al. (2011)	STORIES	Cognitive	Unreported
	RMET	Affective	
	Other: FB	Cognitive	
	Other: ER	Affective	
Cavallini et al. (2013)	STORIES (Italian vers.)	Cognitive	IRR = .93
Charlton et al. (2009)	STORIES	Cognitive	IRR > .71
Duval et al. (2011)	STORIES	Cognitive	Unreported
	RMET	Affective	
	Other: IA	Cognitive	
	Other: FB	Cognitive	
	Other: Composite	*na*	
Fischer et al. (2014)	STORIES	Cognitive	IRR = .94
German & Hehman (2006)	Other: FB	Cognitive	Unreported
Happé et al. (1998)	STORIES	Cognitive	Unreported
Keightley et al. (2006)	Other: Stories	Cognitive	Unreported
	Other: ER	Affective	
MacPherson et al. (2002)	Other: ER	Affective	Unreported
	Other: FP	Cognitive	
Mahy et al. (2014b)	Other: Voice	*na*	Unreported
Maylor et al. (2002)	STORIES	Cognitive	Unreported
McKinnon & Moscovitch (2007)	Other: Stories	Cognitive	Unreported

Pardini & Nichelli (2009)	RMET	Affective	Unreported
Phillips et al. (2002)	Other: Stories	Cognitive	α = .85*
	Other: Video/cartoon	*na*	α = .49*
Phillips et al. (2011)	Other: Stories	Cognitive	Unreported
	Other: Video/Cartoon	*na*	
Rakoczy et al. (2012)	STORIES	Cognitive	Unreported
	Other: Video/Cartoon	*na*	
Saltzman et al. (2000)	STORIES	Cognitive	Unreported
	Other: FB	Cognitive	
Slessor et al. (2007)	STORIES	Cognitive	Unreported
	RMET	Affective	
	Other: Video/Cartoon	*na*	
Sullivan & Ruffman (2004)	STORIES	Cognitive	Unreported
	Other: Video/Cartoon	*na*	
Wang & Su (2013)	STORIES	Cognitive	IRR = .89
Yeh (2013)	STORIES (Chinese ver.)	Cognitive	α = .93
	Other: FP	Cognitive	α = .91
	Other: Video/Cartoon	*na*	α = .93

Note: STORIES = Strange Stories Test. RMET = Reading the Mind in the Eyes Test. "*na*" denotes when insufficient information was provided about a test to accurately judge representation of cognitive/affective content. Tests listed under "Other" include all ToM tests labeled into the following categories by type of test: FP = faux pas, FB = false belief, EI = emotion recognition, IA = intention attribution, MJ = moral judgment and Video/Cartoon = test that used videos, cartoons, or picture series to assess ToM.

*Cited reliability data from a previously published study.

Appendix C.

Associations between ToM and Neurocognitive Performance

Table C1. Reported correlation coefficients between ToM and neurocognitive performance in healthy adults

	Sample size	ToM		Neurocognitive domain							
Study	*n*	C-ToM	A-ToM	EF-Comp	Inhibition	WM/Attn	Flexibility	Fluency	Fluid Rs	Verbal	Memory
Ahmed & Miller (2011)	123 YA (*M* = 19.04)	S-Stories			-.00	.18*	-.01	.22*	.14		
		FP			.06	.04	.07	.22*	.18*		
			RMET		.22*	-.03	.03	.23*	.21*		
Ahmed & Miller (2013)	*46 OA* (*M* = *77.0)*	S-Stories			.38**		.47**				
		FP			.25		.37**				
Bailey & Henry (2008)	32 YA (*M* = 19.5)	FB videos			.27* to.35**			.18	.34**		.23
	33 OA (*M* = 72.2)		RMET		.34** to .52**			.33**	.38**		.28*
Bernstein et al. (2011)	38 YA (*M* = 19.2)	SB		-.23		-.11				.24	
	20 MA (*M* = 57.3)										
	37 OA (*M* = 67.6)										
Cavallini et al. (2013)	30 YA (*M* = 23.6)	S-Stories			-.30**	.65**				.30**	
	30 MA (*M* = 64.9)										
	29 OA (*M* = 75.0)										
Charlton et al. (2009)	106 OA (*M* = 69.0)	S-Stories		.38**	.33**	.22 - .26**	.34**	.26**		.46**	
Duval et al. (2011)	25 YA (*M* = 23.8)	1st FB		.50**							.13
	20 MA (*M* = 52.6)	2nd FB		.27**							.09
	25 OA (*M* = 70.1)	Attr.Intent		.46**							.15
			RMET	.20							.29*
Fischer et al. (2014)	66 OA (*M* = 73.5)	S-Stories			-.00	.32*	-.18				.35**

	Sample size	ToM		Neurocognitive domain						
Study	*n*	C-ToM	A-ToM	EF-Comp	Inhibition	WM/Attn	Flexibility	Fluency	Fluid Rs	Verbal	Memory
German & Hehman (2006)	27 YA (*M* = 20.0) 20 OA (*M* = 78.0)	FB stories			-.42**	.35 to .49**				.28	
Keightley et al. (2006)	30 YA (*M* = 25.7) 30 OA (*M* = 72.5)	ToM stories									.47**
Li et al. (2013)	28 YA (*M* = 20.5) 52 OA (*M* = 73.5)	FB stories FP			-.34* -.36*	.51** .55**	-.30* -.33*				
			RMET		-.31*	-.05	-.24				
Maylor et al. (2002)	30 YA (*M* = 21.2) 30 OA (*M* = 80.6)	S-Stories						.27*	-.27*	.25*	
Peterson et al. (2012)	45 YA (*M* = 19.0)		RMET						.18	.49*	
Phillips et al. (2002)	30 YA (*M* = 29.9)	FB stories							.19	.43*	
	30 OA (*M* = 69.2)		RMET						.10	.04	
			Blends						.20	.45*	
Phillips et al. (2011)	52 YA (*M* = 25.8)	FB stories			-.09 to -.26**	.31**					
	41 MA (*M* = 51.8)	FB videos			-.09 to .01	.25**					
	36 OA (*M* = 73.4)										
Rakoczy et al. (2012)	27 YA (*M* = 22.7)	S-Stories		-.33*	-.29*		-.28*			-.32*	
	20 OA (*M* = 73.3)	FB videos		-.57**	-.48**		-.52**			-.52**	
Saltzman et al. (2000)	9 YA (*M* = 20.9)	S-Stories						.29			
	8 OA (*M* = 71.6)	FB: other						.36 - .60**			
	11 PD (*M* = 71.0)										
Sullivan & Ruffman (2004)	24 YA (*M* = 30.0)	FB stories							.08 - .44*		
	24 OA (*M* = 73.0)	FB videos							.07 - .21		
Wang & Su (2013)		FB stories			.12 to .33*		.35*				
			FB stories		-.05 to .07		.02				
					-.22 to .27*		.21*				

	Sample size	ToM			Neurocognitive domain						
Study	*n*	C-ToM	A-ToM	EF-Comp	Inhibition	WM/Attn	Flexibility	Fluency	Fluid Rs	Verbal	Memory
Yeh et al. (2013)	184 OA (*M* = 71.6)	S-Stories					-.16*	.10			
		FP					-.28**	.16*			
		FB pictures					-.29**	.12			

Note: Where available, values represent bivariate correlations specific to each study sample. EF-Comp: Executive functions composite *z*-score or principle component analysis score. Inhibition: Hayling test; Stroop Golden test; DKEFS CW; Go-no-Go tasks. WM/Attn: WAIS LNS or DS; Spatial span. Flexibility: DKEFS Trail Making test; Trail Making Test A & B; Wisconsin Card Sorting Task categories. Fluency: FAS phonemic fluency; Animals semantic fluency; design fluency. Fluid Rs: Fluid reasoning measures such as Raven's Matrices or WAIS PIQ. Verbal: Language measures such as Mill Hill vocabulary, ETS vocabulary, WAIS VIQ. Memory: short- and long-term retention measures from CVLT-II, HVLT-R. S-Stories: Strange stories test from Happé et al. (1998). FP: Faux pas test from Stone et al. (1999). FB videos: 1st and 2nd order false-belief videos from Sullivan & Ruffman (2004). SB: Sandbox test from Bernstein et al. (2011). RMET: Reading the Mind in the Eyes test from Baron-Cohen et al. (2001). 1st FB, 2nd FB, and FB stories: mix of false-belief stories to assess cognitive ToM and in one case affective ToM (Wang & Su, 2013).

*$p < .05$, **$p < .01$

Appendix D.

Participant Recruitment and Data Preparation

Data Preparation

Prior to the main analyses, we checked the dependent and independent variables of interest for accuracy of data entry, missing values, and fit between variable distributions and the assumptions our initial analyses and those of path analysis (outlined in footnote 7 of the main text). After excluding the five participants (4 young, 1 older) we identified as exhibiting low effort during neurocognitive testing, we examined variables separately for the remaining n = 89 young and n = 89 older adults.

At the univariate level, we examined normality using *Q-Q* plots and histograms for the demographic, neurocognitive, and blood pressure variables. We used a pre-defined alpha of .001 to determine outlying cells on each variable. Outlying cells were defined as those with *z*-scores greater than |3.29| from the mean value of all other cells. In order to maintain rank order of the data we altered each score to a less extreme value by adding one unit to the highest non-outlying score (Tabachnick & Fidell, 2013). We 32 cells (~.001%) of the overall data by this procedure. Following outlier transformation, we compared mean values for relevant variables with the respective 5% trimmed means to ensure that the impact of influential data points was attenuated. Standardized estimates for skew and kurtosis were within acceptable ranges for all independent variables. However, our individual dependent ToM variables (STORIES, RMET, YONI C-TOM and YONI A-TOM) demonstrated negative skew and failed the Shapiro-Wilk normality test. To address non-normality we applied logarithmic and square root transformations, however, these did not significantly improve the data distributions, thus we transformed each dependent variable to z-scores and retained the original distribution properties. Visual inspection of the scatterplots for the z-score dependent variables against each independent variable using general and lowess fit lines suggested pairwise linearity for each dependent variable. Further, no issues with normality were identified once the variables were translated into composite indicators. Thus, we retained z-score dependent variables for all subsequent analyses.

At the multivariate level, we assessed for multivariate outliers in the linear set of independent variables using conservative probability estimates for Mahalanobis' D^2 (χ^2 = 27.87, p< .001) and centred leverage values (h_{ii} = 0.16, p < .001; Tabachnick & Fidell, 2013). No cases met these criteria. We compared externally studentized residuals for each dependent variable to an alpha of .05 corrected for multiple comparisons (t = 4.87, p< .0125) to assess for any cases with undue influence over the regression line for the set of predictors (i.e., discrepancy). No cases exceeded this value and no dependent variable exceeded recommended criterion that |E∗i| >2.00 represent no more than five percent of cases. Visual inspection of *Q-Q* plots and scatterplots for residuals using general fit lines and lowess fit lines also suggested that, for each dependent variable, the spread of residuals was relatively uniform across values of the predicted scores. We then examined Cook's D, standardized DFBETAS, and scatterplots of externally studentized residuals against centred leverage values to assess for any cases with extreme influence for each dependent variable. Across all dependent variables, seven

cases emerged as highly influential multivariate outliers: four young adults and three older adults. We ran the primary analyses with and without these outliers; however, the fit of our models was significantly decreased by their presence so we opted to run the final analyses without these outliers (i.e., final *N* = 171; 86 young adults and 85 older adults). Post-hoc analyses of the five influential cases revealed no discernable pattern to suggest the cases were associated with a specific demographic, cognitive, or clinical trait (examined = biological sex, ethnicity, English as an additional language status, born in North America, CES-D or MAQ total score, test examiner, test location, educational attainment, presence of a cardiovascular illness [yes/no]).

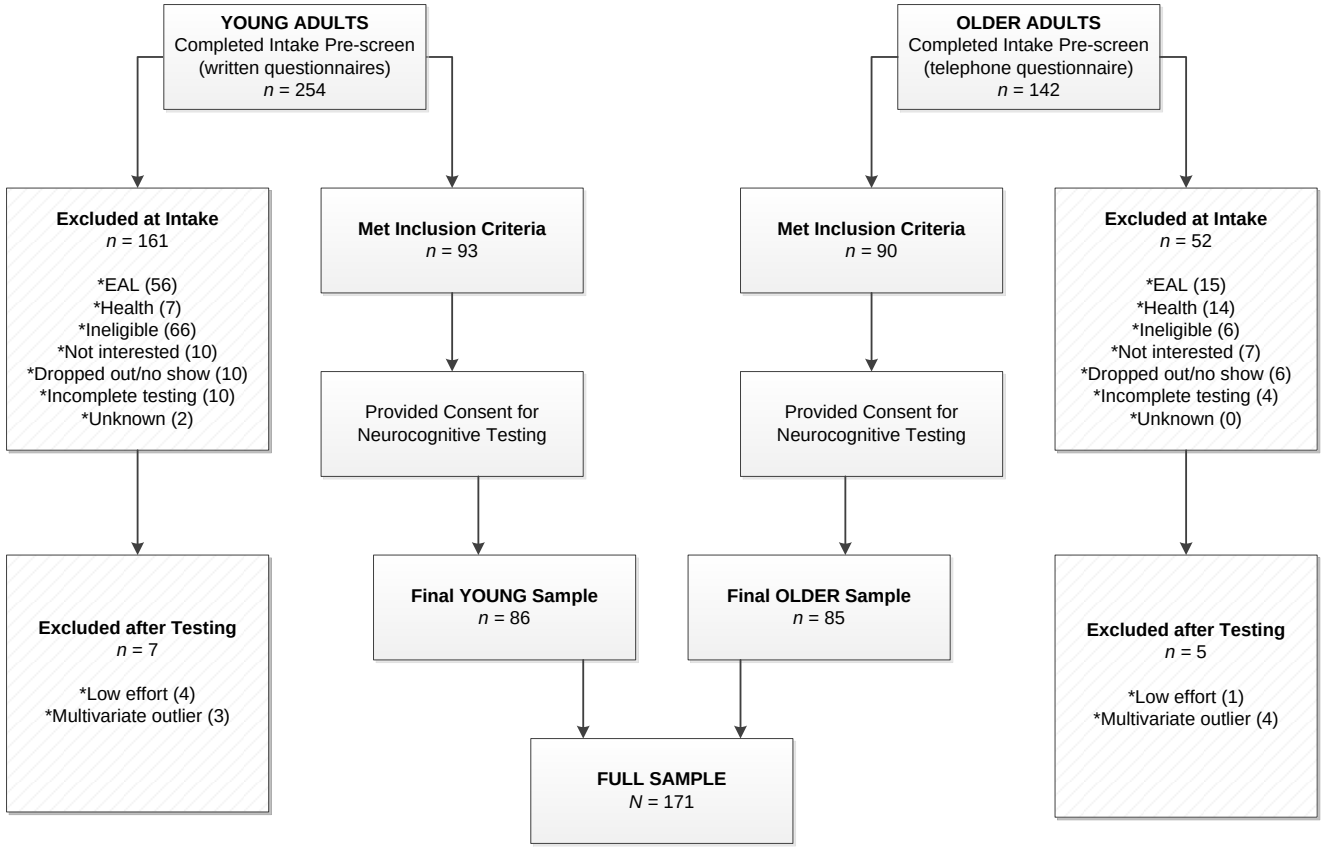


Figure D 1 Recruitment Flow Chart

Note: Regarding inclusion/exclusion criteria at intake, EAL = participants who indicated on an acculturation questionnaire less than 3 out 4 preferences as "English" for speaking, reading, writing, and thinking; Health = diagnosed cognitive impairment, colour-blindness, severe learning disability, traumatic brain injury with loss of consciousness >15 minutes, severe migraine affecting thinking abilities; Ineligible = for young adults, participants did not need the full number of course credits we were offering; Not interested = participants who declined study participation due to lack of interest; Dropped out/show = participants who signed up to complete study but dropped out prior to testing or did not show up for their appointment; Incomplete testing = participants who, for various reasons, did not complete the full testing battery.

Appendix E.

Sample Items from ToM Tests

Table E 1 Sample Items: Strange Stories Test (Cognitive ToM)

Example 1	Sample Responses
Brian is always hungry. Today at school it is his favourite meal—sausages and beans. He is a very greedy boy, and he would like to have more sausages than anybody else, even though his mother will have made him a lovely meal when he gets home! But everyone is allowed two sausages and no more. When it is Brian's turn to be served, he says, "Oh, please can I have four sausages, because I won't be having any dinner when I get home!"	**2 points:** *He's lying to make them feel sorry for him* *He's greedy and wants to persuade them to give him more* **1 point:** *He likes sausages and wants more than anyone else* *He's a very greedy boy who loves sausages*
Q: Why does Brian say this?	**0 points:** *Because his mom won't be cooking for him tonight* *He wants to take the sausages home with him*

Example 1	Sample Responses
Jill wanted to buy a kitten, so she went to see Mrs. Smith, who had lots of kittens that she didn't want. Now Mrs. Smith loved the kittens, and she wouldn't do anything to harm them, though she couldn't keep them all herself. When Jill visited she wasn't sure she wanted one of Mrs. Smith's kittens, since they were all males and she had wanted a female. But Mrs. Smith said, "If no one buys the kittens I'll just have to drown them!"	**2 points:** *So Jill will feel sorry for the kittens and want to buy one of them* *To manipulate or threaten her to take one* **1 point:** *Because she wanted Jill to buy one* *The old lady doesn't want to keep them all*
Q: Why did Mrs. Smith say that?	**0 points:** *Because otherwise she would have to drown them* *She's a terrible woman and would kill them otherwise*

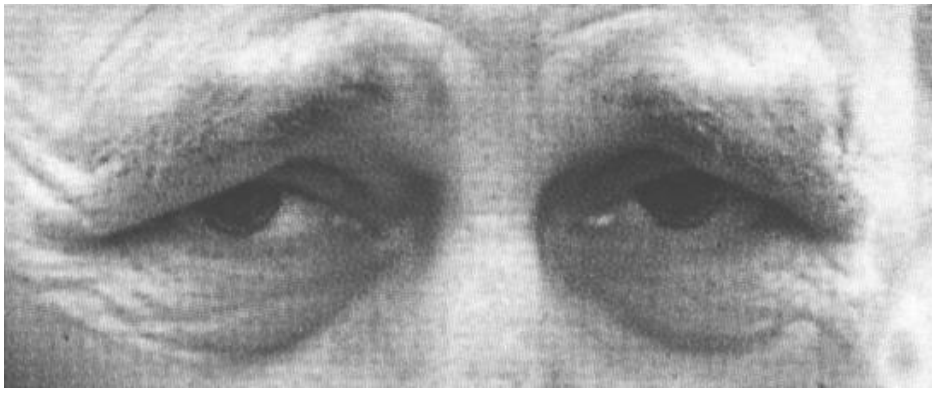
Example 1

playful *comforting*

irritated *bored*

Example 2

terrified *upset*

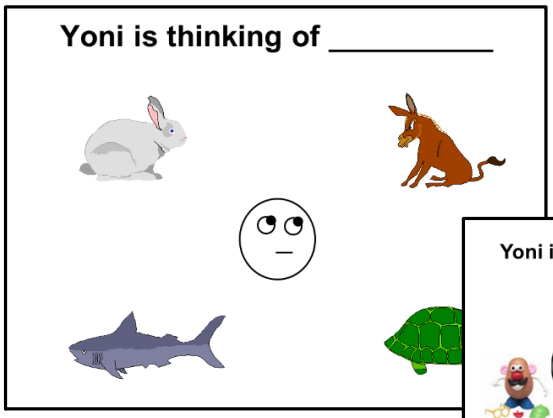


arrogant *annoyed*

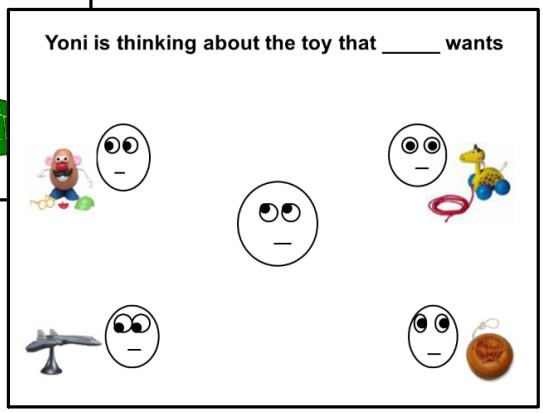
Figure E 1 Sample Items: Reading the Mind in the Eyes Test (Affective ToM)

Note: The correct answer for example 1 is *playful*. The correct example for sample 2 is *upset*.

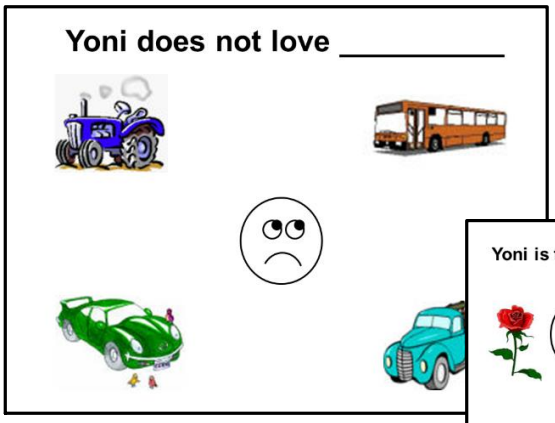
(a)First-order cognitive ToM



(b)Second-order cognitive ToM



(c)First-order affective ToM



(d)Second-order affective ToM

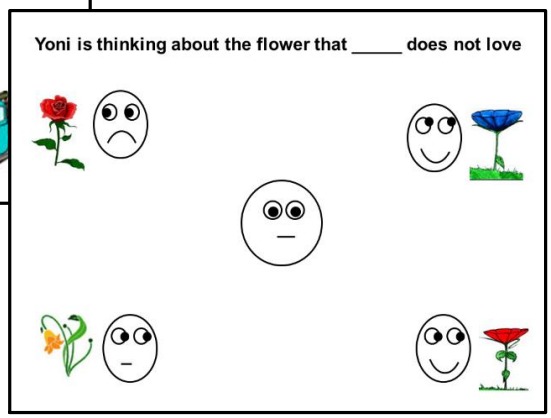


Figure E 2 Sample Items: Yoni Test (cognitive & affective ToM)

Note: The correct answers are (a) top right, (b) top left, (c) top right, and (d) top left.

Appendix F.

Supplementary Analysis Tables

Table F 1 Summary of Items Deleted for each Theory of Mind Test during Psychometric Analysis

Test	Items Deleted				Score Range	
	Total Deleted	Poor Response Variability	Poor Foil Items[a]	Low Item-Total Agreement	Original Version	Final Version
STORIES	1	6 (Hat)	-	none	0 - 16	0 - 14
RMET	10	none	23, 25	1, 4, 6, 7, 12, 14, 17, 29	0 - 36	0 - 26
YONI C-TOM	6	11, 17, 75	-	25, 93, 63	0 - 34	0 - 28
YONI A-TOM	5	12, 31	-	20, 35, 50	0 - 33	0 - 28

Note: 'Poor response variability' represents items that had zero variability in the responses made by participant, thereby suggesting inadequate item properties. 'Poor foil items' represents items in which one or more foil options was chosen by >25% of respondents (assessed for RMET only), thereby suggesting that the item poorly discriminated among foil and target response options. 'Low item-total agreement' represents items that had low correlations with the composite test score ($r < .10$), thereby indicating that these items were poorly associated with overall performance on the respective test.

[a] We followed criteria outlined in Baron-Cohen and colleagues' (2001) article during RMET development to determine whether individual items had adequate distribution properties. As per the authors, we considered items to be satisfactorily difficult if at least 50% of all participants in the sample selected the target response and no more than 25% of all participants selected any one of the foil responses. In our sample, no target word was selected by less than 50% of participants; however, two foil items were selected by greater than 25% of participants.

Table F 2 Distribution Properties for the Theory of Mind Measures

Measure	Skew (S.E.)	Kurtosis (S.E.)
Individual ToM variables		
STORIES	-.84 (.19)	.57 (.37)
RMET	-.46 (.19)	.01 (.37)
YONI C-TOM	-1.61 (.19)	..82 (.37)
YONI A-TOM	-1.01 (.19)	.68 (.37)
Composite ToM variables		
Cognitive ToM	-.83 (.19)	-.64 (.37)
Affective ToM	-.53 (.19)	-.08 (.37)

Note: *Cognitive ToM* and *Affective ToM* represent composite mean z-score variables. For all variables higher scores represent better performance.

Table F 3 Correlation Matrices for Theory of Mind and Individual Neurocognitive Variables by Age Group

	Cognitive ToM	*Affective ToM*	*KBIT VK*	*CVLT15*	*CVLT SDFR*	*CVLT LDFR*	*WAIS LN*	*WAIS DS*
				Young Adults				
Cognitive ToM	-							
Affective ToM	.41***	-						
KBIT VK	.24*	.29**	-					
CVLT 15	.27*	.09	.29**	-				
CVLT SDFR	.18	.11	.23*	.75***	-			
CVLT LDFR	.17	.11	.33**	.74***	.90***	-		
WAIS LN	.28**	.21*	.29**	.41**	.28**	.28*	-	
WAIS DS	.22*	.17	.24*	.25*	.31**	.30**	.50***	-
DKEFS CW	.15	.13	.25*	.40***	.18	.21*	.26*	.21*
				Older Adults				
Cognitive ToM	-							
Affective ToM	.56***	-						
KBIT VK	.33**	.51***	-					
CVLT 15	.31**	.37**	.32**	-				
CVLT SDFR	.23*	.33**	.30**	.86***	-			
CVLT LDFR	.25*	.34**	.28*	.83***	.84***	-		
WAIS LN	.27**	.35**	.23*	.58***	.53***	.52***	-	
WAIS DS	.36**	.41***	.20	.18	.22*	.32**	.32**	-
DKEFS CW	.25*	.21*	.15	.29**	.34**	.24**	.36***	.29**

Note: *Cognitive ToM* and *Affective ToM* represent composite mean z-score variables. KBIT VK = KBIT-2 Verbal Knowledge subtest; CVLT 15, SDFR & LDFR = CVLT-II Trials 1-5, Short Delay Free Recall, and Long Delay Free Recall; WAIS LN & DS = WAIS-III Letter Number Sequencing & Backwards Digit Span; DKEFS CW = DKEFS Color-Word Interference Contrast Score (timed contrast index). For all tests higher scores represent better performance, except for DKEFS CW, where lower (faster) scores represent better performance. *p < .05, **p < .01, ***p < .001.

Table F 4 Correlation Matrices for Theory of Mind and Demographic Variables by Age Group

	Cognitive ToM	*Affective ToM*	*Ethnicity*	*Birthplace*	*EAL*	*CES-D*	*MAQ*
				Young Adults			
Cognitive ToM	-						
Affective ToM	.41***	-					
Ethnicity	-.01	..06	-				
Birthplace	.09	-.03	.28**	-			
EAL	-.04	.06	-.31**	-.52***	-		
CES-D	-.11	-.13	.06	.03	0	-	
MAQ	-.13	-.14	.03	-.08	.02	.74***	-
Hypertension[a]	-	-	-	-	-	-	-
				Older Adults			
Cognitive ToM	-						
Affective ToM	.56***	-					
Ethnicity	.04	.11	-				
Birthplace	.07	.14	.20	-			
EAL	-.16	-.38***	-.44***	-.29**	-		
CES-D	-.19*	-.04	.15	.10	-.18	-	
MAQ	-.21*	-.09	.13	0	-.13	.80***	-
Hypertension	-.13	-.08	-.17	-.13	.17	0	.04

Note: *Cognitive ToM* and *Affective ToM* represent composite mean z-score variables. EAL = as an additional language; CES-D = Centre for Epidemiologic Studies Depression Scale; MAQ = Multidimensional Anxiety Questionnaire. For ethnicity 1 = Caucasian, 0 = non-Caucasian; for birthplace 1 = born in North America, 0 = born outside of North America; for EAL 1 = indicated English as their first language, 0 = indicated any language other than English as their first language; for hypertension, 1 = self-reported physician's diagnosis of hypertension and corresponding medication or lifestyle treatment, 0 = no hypertension.

[a] No participant in the young adult sample reported a physician's diagnosis of hypertension. *p< .05, **p< .01, ***p< .001

Appendix G.

Reliability of the Reading the Mind in the Eyes Test

It is worth mentioning the poor reliability of the RMET considering its use in over 250 studies across the lifespan (Kirkland et al., 2013). Most information regarding the reliability of the RMET (and ToM more broadly) is garnered from cross-cultural validation studies that have translated the test into different languages (Vellante et al., 2013). Recently reported internal consistency estimates for the RMET are generally quite poor (e.g., α = .64, Söderstrand & Almkvist, 2012; α = .61, Vellante et al., 2013), and test-retest reliability estimates range from poor to minimally acceptable across time ranges of (one-year interval: ICC = .63, Fernández-Abascal et al., 2013; three-week interval: r = .60, Hallerbäck, Lugnegård, Hjärthag, & Gillberg, 2009; average three-week interval: Pfaltz et al., 2013). These estimates are notably similar to what we observed in the current study. For the RMET, it appears that a number of test items may not be measuring the same construct across populations. Similar to others (Peterson & Miller, 2012; Pfaltz et al., 2013), we used item analysis to inform our decisions regarding questionable test items and conducted our primary analyses on a truncated 26-item version of the original test. Possible explanations for low inter-item convergence that warrant further consideration include sex differences in effort or performance (Kirkland et al., 2013), speed-accuracy trade-off in responding (Harkness et al., 2005), or ethnocultural differences in face or emotion processing (Koelkebeck et al., 2011). Other authors have found that examining the emotional valences of the mental state judgments required by each question in the RMET may yield useful information for construct validity (Harkness et al., 2005).

Appendix H.

Alternative Path Models

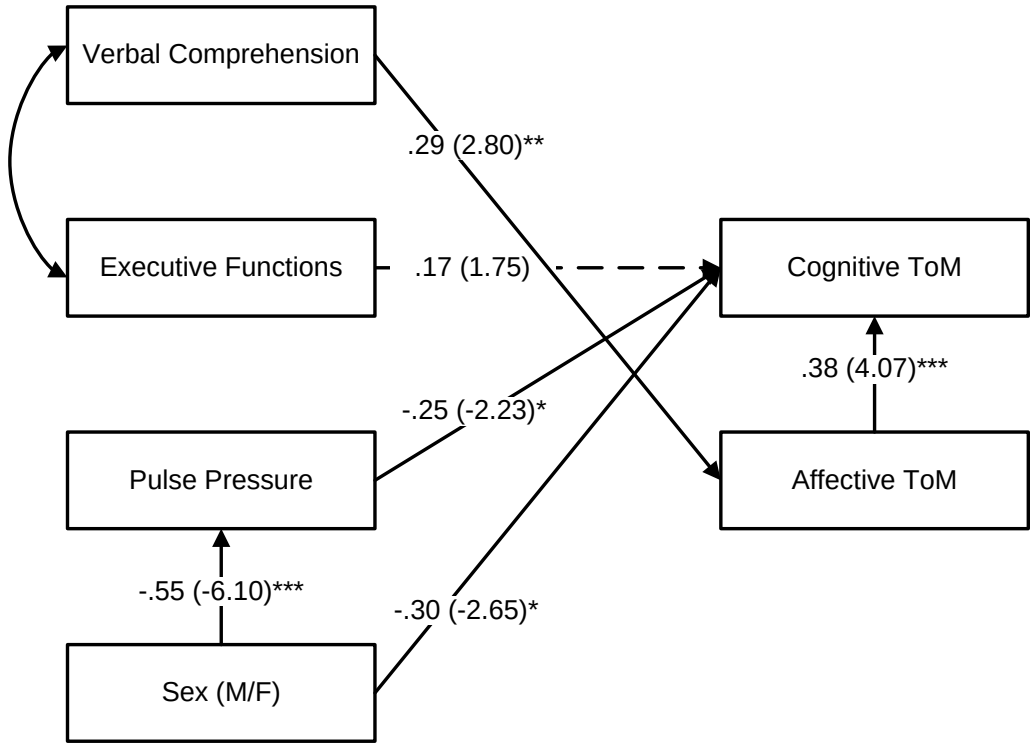


Figure G1 Alternative path model with affective ToM predicting cognitive ToM in young adults

Note1: χ^2 [*df* = 8] = 7.34, *p* = .50; CFI = 1.00; RMSEA = 0 (90% CI [0, .12]); SRMR = .07. Parameter values are expressed as maximum likelihood estimates (standardized solution). Numbers in parentheses indicate *t* values for parameter estimates (statistically significant *t* values > |1.96|). We coded sex (M/F) as 0 = male and 1 = female.

Note2: Dashed arrows represent the individual associations that dropped to non-significance in the alternative model.

*p < .05, **p < .01, ***p < .001.

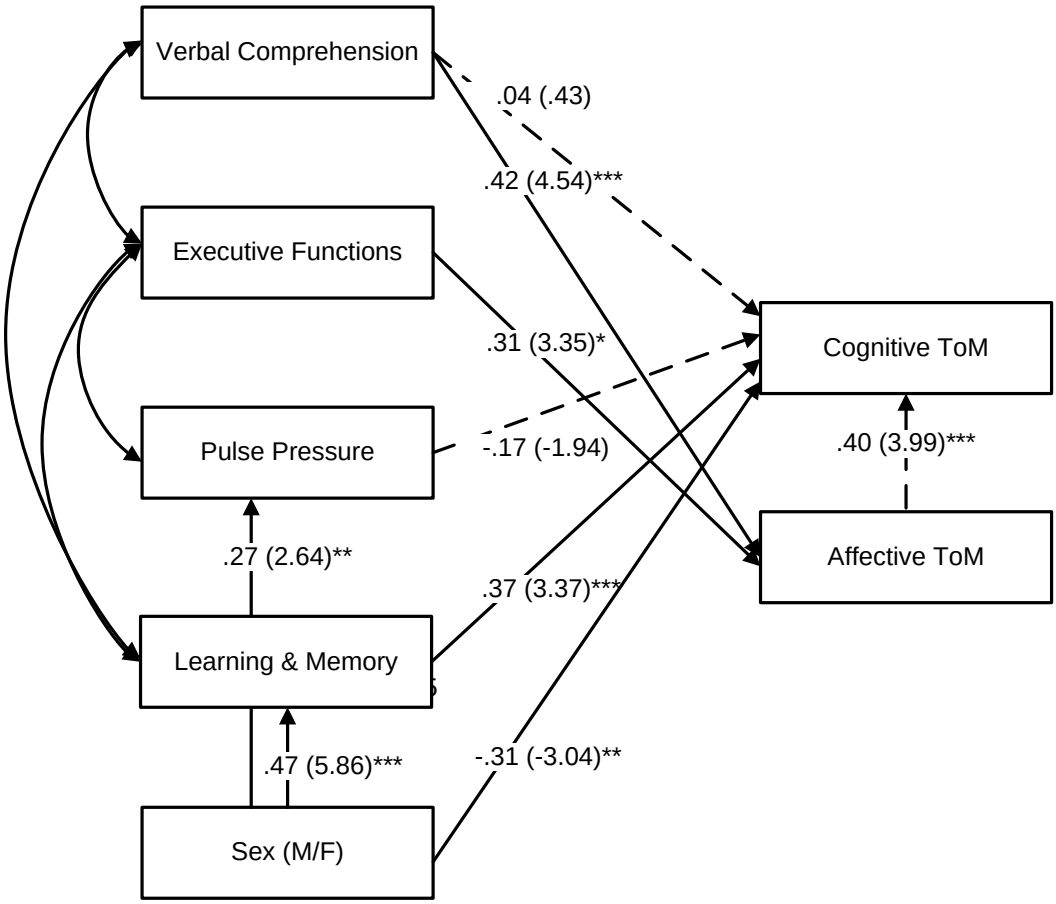


Figure G2 Alternative path model with affective ToM predicting cognitive ToM in older adults

Note1: χ^2 [*df* = 8] = 5.25, *p* = .73; CFI = 1.00; RMSEA = 0 (90% CI [0, .09]); SRMR = .06. Parameter values are expressed as maximum likelihood estimates (standardized solution). Numbers in parentheses indicate *t* values for parameter estimates (statistically significant *t* values > |1.96|). We coded sex (M/F) as 0 = male and 1 = female.

Note2: Dashed arrows represent the individual associations that dropped to non-significance in the alternative model.

*p < .05, **p < .01, ***p < .001.

Appendix I.

Non-linear Blood Pressure Effects

A number of variables in community-residing samples of young and older adults may contribute to the broad range of blood pressure that is typically observed in cross-sectional studies (e.g., use/non-use of anti-hypertensive medication and varying adherence rates, white coat hypertension, comorbid CVRFs, genetics, etc.). Where a wide range of blood pressure is available, it is possible that nonlinear (U-shaped or J-shaped) relations between blood pressure and neurocognitive performance may emerge: Very low blood pressure and very high blood pressure may be both detrimental to neurocognitive health when compared to mid-ranges of blood pressure (Waldstein et al., 2005). Thorvaldsson and colleagues (2012) found that in community-residing older adults, both extremely low (< 75 mmHg) and extremely high (> 95 mmHg) DBP were associated with cognitive decline over a 30-year period. Importantly, the use of a non-linear (quadratic) variable conferred a statistical advantage in predicting associations than the linear variable. Similar findings are reported for SBP (Waldstein et al., 2005). Contrary to this research, we did not find that non-linear relationships offered any meaningful advantage to interpreting data in the current study (see also Fischer et al., 2014; Yeung & Thornton, 2011). It is possible non-linear patterns may emerge in clinical samples with greater severity and burden of CVRFs and wider ranges of PP available (Thorvaldsson et al., 2012).

Appendix J.

Indirect Predictors of ToM

As a supplemental analysis we assessed whether any indirect or mediating relationships existed among the predictor variables and ToM using bootstrap procedures implemented in the AMOS software. Analysis of indirect effects was conducted based on 5000 bootstrapped samples using the separate path models for young and older adults (Figures 5.2 & 5.3). Table J1 summarizes the standardized estimates for the Indirect Effects and significance levels that relate to cognitive and/or affective ToM, along with 95% bias corrected and accelerated confidence intervals. An indirect mediating relationship was considered to be indicated when, for any effect, the 95% bootstrapped CI did not include the value zero, indicating that the indirect effect of the mediator differed from the value zero at $p < .05$ (two tailed).

Consistent with theoretical evidence presented in this dissertation for an organizational scaffolding within ToM itself (e.g., Shamay-Tsoory et al., 2009; Vetter et al., 2013), cognitive ToM mediated the association between executive functions and affective ToM in young adults (Indirect Effect = .10, $p < .01$, 95% CI [.03, .18]). This indicated that in early adulthood, better executive functions predicted better affective ToM indirectly via strong cognitive ToM. Lower cognitive ToM mediated the association between high PP and lower affective ToM across age groups (young: Indirect Effect = -.12, $p < .01$, 95% CI [-.24, -.03]; older: Indirect Effect = -.07, $p = .02$, 95% CI [-.18, -.01]). This indicated that across ages, high PP negatively impacted affective ToM via poor cognitive ToM. In older adults, cognitive ToM mediated links between better Learning & Memory and better affective ToM (Indirect Effect = .12, $p < .01$, 95% CI [.04, .23]).

Further, in line with cognitive aging literature (e.g., McFall et al., 2014; Yeung & Thornton, 2011), PP was an important mediator of ToM performance. High PP mediated the association between biological sex and cognitive ToM in young adults (Indirect Effect = .17, $p < .01$, 95% CI [.05, .33]). In older adults high PP jointly (along with Learning & Memory and cognitive ToM) mediated associations between biological sex and affective ToM (Indirect Effect = -.70, $p = .02$, 95% CI [-.18, -.01]).

Such indirect associations should be considered preliminary due to the smaller size of our sample and potential interactive influences in the model; however they do lend confidence to our findings regarding age differences in ToM. Variability in cognitive and affective ToM is influenced by an array of neurocognitive, health, and demographic variables and these variables operate both directly and indirectly. Our findings are in line with neuropsychological views of ToM as a multidimensional construct, rather than an innate, automatic process as assumed by historical perspectives. To truly understand how these indirect effects operate future work is needed to replicate and extend these findings within a larger sample and at multiple levels of analysis (e.g., neurocognitive versus morphological; Kraemer et al., 2001). We note that that bootstrapped indirect effects and their confidence intervals are less reliable when samples are small, such as in the current study (Bollen & Stine, 1990). In this dissertation interpretation of the direction and magnitude of indirect effects is limited due to interactive influences and our use of nested models for the age groups (Finney, 1972).

Table J 1 Summary of Indirect Effects for Predictors of Cognitive and Affective ToM

Association	Indirect Predictor(s)	Direct Effect	*p*	Indirect Effect	*p*	95% CI [lower, upper]	Indirect Effect indicated?
Young Adults							
Executive Functions → Affective ToM	Cognitive ToM	0	-	.10	< .01	[.03, .18]	**yes**
PP → Affective ToM	Cognitive ToM	0	-	-.12	< .01	[-.24, -.03]	**yes**
Sex (M/F) → Cognitive ToM	PP	-.31	< .01	.17	< .01	[.05, .31]	**yes**
Sex (M/F) → Affective ToM	Cognitive ToM	0	-	-.06	.07	[-.14, .01]	no
Older Adults							
Verbal Comprehension → Affective ToM	Cognitive ToM	.35	< .01	.07	.07	[-.01, .19]	no
PP → Affective ToM	Cognitive ToM	0	-	-.07	.02	[-.18, -.01]	**yes**
Learning & Memory → Affective ToM	Cognitive ToM	0	-	.12	< .01	[.04, .23]	**yes**
Sex (M/F) → Cognitive ToM	Learning & Memory, PP	-.31	.02	.11	.15	[-.04, .27]	no
Sex (M/F) → Affective ToM	Learning & Memory, PP, Cognitive ToM	0	-	-.07	.03	[-.15, -.01]	**yes**

Note: We present the standardized Direct Effects and standardized Indirect Effects for analyses conducted on 5000 bootstrapped samples. We considered the Indirect Effect significant if the associated 95% bias corrected and accelerated confidence interval (i.e., column four) did not contain the value zero.