Exploring Dedifferentiation Across the Adult Lifespan

Claire La Fleur

B.A., St. Mary's College of Maryland, 2013
M.A., University of Virginia, 2015

A Dissertation Presented by the Graduate Faculty of the University of Virginia
Department of Psychology

May 2017

Committee Member
Timothy Salthouse (Chair)
Chad Dodson
Angeline Lillard
Carol Manning

Table of Contents

Introduction

One of the most robust findings in psychometric research is that cognitive abilities are interrelated and that these relations between abilities can be organized in terms of a hierarchical structure (Carroll, 1993; Gustafson, 1984; Salthouse & Ferrer-Caja, 2003). For example, in the Cattell-Horn-Carroll (CHC) structure, the lowest level is composed of observed measures (Carroll, 1993; Carroll, 2003; Keith & Reynolds, 2010). At the second level are latent factors like mental transformation (for example reasoning), crystallized intelligence (ability to use learned knowledge and experience), visual perception, auditory perception, memory, and processing speed that are composed of the variance shared across the tests designed to measure them. The highest level of the structure is commonly known as general intelligence, or *g*, and represents the broadest type of cognitive functioning. It is measured by the variance the latent factors have in common. These ability structures may be impacted by certain individual differences (Tucker-Drob, 2008); the focus of this thesis will primarily be on one such difference, aging.

The age dedifferentiation hypothesis posits that in older adulthood, cognitive abilities begin to share more variance (Anstey, Hofer, & Luszcz, 2003; Baltes, Cornelius, Spiro, Nesselroade, Willis, 1980; Hertzog & Bleckley, 2001; Hulur, Ram, Willis, Schaie, & Gerstorf, 2015; Lindenberger & Baltes, 1997a; Tucker-Drob, 2009; Zelinski & Lewis, 2003). For example, Baltes & Lindenberger, (1997a) found that sensory abilities explained only 11% of variance in cognitive abilities for adults aged 25-69 compared with 31% in adults aged 70-103. Cognitive dedifferentiation is hypothesized to occur as senescence places progressively greater constraints on cognitive function, leading to increased correlations between abilities. (Ghisletta & Ribaupierre, 2005; Li et al. 2004). These constraints could be biological, for example structural or functional declines in the brain (Lindenberger & Baltes, 1997b; Hulur et al., 2015), or cognitive, for example decreased processing speed (Li et al., 2004, Ghisletta & Ribaupierre, 2005).

Understanding whether dedifferentiation occurs is crucial both theoretically and methodologically. Theoretically, it can teach us about how changes in the brain with age or pathology impact expression of cognitive performance. Identifying moderators to the cognitive structure can also provide more insight into why cognitive abilities are interrelated and about the nature of *g*. Methodologically, it is important because a difference in the structure of cognition at different ages implies that the measures of cognition vary in meaning across the adult lifespan. Such differences mean that quantitative comparisons made between age groups are confounded by qualitative differences, thus limiting the conclusions that can be made.

The dedifferentiation hypothesis can be tested in three ways. The first option is to evaluate how many cognitive factors are present (e.g. Hedden & Yoon, 2006, Schaie, Willis, Jay, & Chipuer, 1989). If dedifferentiation exists there should be fewer factors in older age. However, this method is somewhat weak in demonstrating the lack of dedifferentiation, because correlations could still increase even if the factor structure does not change. The second option is to test how much variance is explained by g (e.g. Juan-Espinosa et al., 2002). In this case, the expectation is that more variance should be shared with age. A final option is to examine how strong the correlations are between measures or latent factors representing various abilities (Hulur et al., 2015; Sims, Allaire, Gamaldo, Edwards, & Whitfield, 2009). In this case, I would expect higher correlations with age. These last two options are a more sensitive test of dedifferentiation as long as statistical tests are used to determine that the correlations are significantly higher for older adults.

Unfortunately, the research on dedifferentiation using all these approaches has been inconsistent. Both longitudinal (Cunningham & Birren, 1980; Deary, Whiteman, Starr, Whalley, & Fox, 2004; de Frias, Lovden, Lindenberger, & Nilsson, 2007; Ghisletta & Lindenberger, 2003; Hertzog, Dixon, Hultsch, & MacDonald, 2003; Hulur et al., 2015) and cross-sectional (Adrover-Roig, Sese, Barcelo, & Palmer, 2012; Babcock, Laguna, & Roesch, 1997; Balsamo & Romanelli, 2010; Cunningham & Birren 1980; de Frias et al., 2007; Hedden & Yoon, 2006; Hertzog, 1989; Hertzog & Bleckley, 2001; Li et al., 2004; Lindenberger & Baltes, 1997b; Nyberg et al., 2003; Salthouse & Saklofske, 2010; Schaie et al., 1989; Schultz, Kaye, & Hoyer, 1980) studies have supported dedifferentiation; for example, Li et al. (2004) found that *g* explained more variance in late adulthood than in young or middle adulthood. Similarly, Hulur et al. (2015) found that over a 30-year period, change in knowledge and mental transformation as well as between tests of mental transformation were increasingly coupled over time.

However, there is also a great deal of evidence, both longitudinally (Anstey et al., 2003; Batterham, Christensen, & Mackinnon, 2011; Finkel, Reynolds, McArdle, & Pederson, 2007; Schaie, Maitland, Willis, & Interieri, 1998; Tucker-Drob, 2009; Zelinski & Stewart, 1998; Zelinski & Lewis, 2003) and cross-sectionally (Escorial, Juan-Espinosa, Garcia, Rebollo, & Colom, 2003; Hale et al., 2011; Hildebrandt, Wilhelm, Schmiedek, Herzmann, & Sommer, 2011; Hull, Martin, Beir, Lane, & Hamilton, 2011; Johnson, Logie, & Brockmole, 2010; Park, Lautenschlager, Hedden, Davidson, & Smith, 2002; Sims et al., 2009; Singer, Verhaeghen, Ghisletta, Lindenberger, & Baltes, 2003; Vaughn & Giovanello, 2010), that is not consistent with dedifferentiation. For example, Anstey et al. (2003) found no consistent patterns of dedifferentiation cross-sectionally (comparing individuals from 60-84 years at T1) or longitudinally (over a period of up to 8 years) using measures of memory, processing speed, and knowledge. Meanwhile, while Tucker-Drob (2008) found some changes in the correlations with older age, those changes actually suggested that *greater* differentiation was occurring.

The inconsistent findings from the dedifferentiation literature can be complemented by studies evaluating cognitive testing batteries for measurement invariance, a series of statistical techniques that test whether the same construct is being measured at different times or in different groups (Benson, Hulac, & Kranzler, 2010; Molenaar, Dolan, Wicherts, van der Maas, 2001; Niileksela, Reynolds, & Kaufman, 2013; Parker, 1983; Taub, McGrew, & Witta, 2004; Waller & Waldman, 1990; Ward, Axelrod, & Ryan, 2000). These studies test for two types of invariance that are potentially relevant to dedifferentiation- configural invariance and metric invariance.

Configural invariance evaluates whether the same model has a good fit across groups. A good fit means that the same organization of variables can be used. A failure to find configural invariance could provide evidence of dedifferentiation if older adults have fewer cognitive ability factors (signifying a compression of the structure). However, configural invariance by itself cannot prove a lack of dedifferentiation, as there could still be an increase in the relations between cognitive abilities that is not enough to change the configural structure.

Metric invariance can be a more relevant test of dedifferentiation when it evaluates whether both the factor loadings and correlations within a model can be constrained to be the same across groups. A failure to find metric invariance due to higher correlations between factors or higher loadings between cognitive abilities and g would be evidence for dedifferentiation. If metric invariance is found between the factor loadings with *g* and there is enough power to detect differences, this would suggest a lack of dedifferentiation, as the relations between cognitive abilities and/or the amount of variance explained by g is stable.

A number of studies do show measurement invariance (Bowden, Weiss, Holdnack, & Lloyd, 2006; Burton, Ryan, Paolo, & Mittenberg, 1994; Cohen, 1957; Parker, 1983; Taub et al., 2004; Niileksela et al., 2013). For example, Taub et al. (2004) found that the factor structure and the loadings were similar for the Weschler Adult Intelligence Scale III across all 13 age groups used in the study. However, evidence of invariance is not found in all studies (Benson et al., 2010; Balsamo & Romanelli, 2010; Fox, Berry, & Freeman, 2014; Waller & Waldman, 1990; Ward, Axelrod, & Ryan, 2000). For example, Waller & Waldman (1990) found that a two-factor model fit better for the 70-74 year-old group than the three-factor model that provided the best fit for age groups that ranged from 16-69. So overall, even the measurement invariance literature has been inconsistent.

A review of studies supporting or failing to support dedifferentiation are listed in Supplementary Table 1. The first important pattern to notice in this table is the large number of different cognitive abilities that are used to test for dedifferentiation. This could be an issue if only certain cognitive abilities dedifferentiate across the adult lifespan (Hertzog & Bleckley, 2001; Li et al., 2004). Secondly, there is a wide variety of age ranges between studies. This is an issue if dedifferentiation only appears at a certain point in the lifespan (de Frias et al., 2007). A third concern is the strength of the evidence that is presented in the literature. For the purposes of this review, weak evidence is defined as analyses that do not directly test whether correlations are larger. Strong evidence statistically tests for differences in correlations with age, although not all include information regarding power. What is clear is that much of the evidence testing for dedifferentiation is weak, which may explain the lack of consistency in the literature. Finally, almost none of the studies provide information about effect sizes (but see Salthouse & Saklofske, 2010) and there is a wide range of sample sizes. If the effect is small, only some of the studies may have the necessary power to detect dedifferentiation. In summary, there are a number of differences in the previous literature that may be responsible for the inconsistent findings.

For my dissertation, I propose a cross-sectional and a longitudinal study that will advance our knowledge of age dedifferentiation. These studies are valuable because they will account for some of the major theoretical and methodological issues that have limited previous work from establishing the nature of dedifferentiation and the conditions in which it occurs. These issues are described in greater detail below.

Different Combinations of Variables

The first theoretical issue that may explain the inconsistences in the literature is the type of cognitive abilities used to test for dedifferentiation. Different combinations of abilities have been examined across studies, such as (a) relations between test of mental transformation, processing speed and memory (Adrover-Roig et al., 2012; Hale et al., 2011; Hulur et al., 2015); (b) between those fluid abilities and crystallized intelligence (Finkel et al., 2007; Singer et al., 2003; Ghisletta & Lindenberger, 2003); and (c) even between measures of cognitive performance and visual or auditory acuity (Lindenberger & Baltes, 1997a). It is also common for studies to use tests that measure a variety of cognitive abilities but not explore if certain combinations of abilities are differentially impacted by age (Balsamo & Romanelli, 2010; Escorial et al., 2003; Sims et al., 2009). This is an issue because these different combinations may be more or less likely to exhibit dedifferentiation, or to differ in when dedifferentiation begins to occur. For the sake of clarity, I will discuss cognitive abilities in four major groups: memory, speed, mental transformation (e.g. working memory, reasoning or measures identified as fluid intelligence), and knowledge.

One of the most important abilities to consider when looking for dedifferentiation is processing speed. Based on findings that processing speed mediates a majority of age differences in other cognitive abilities (Salthouse, 1995; Salthouse, 1996; Salthouse, 2000), particularly for memory and mental transformation abilities compared with knowledge abilities (Li et al., 2004), the general slowing account was proposed and can be extended to provide an explanation for dedifferentiation. According to this account, increasing age is associated with a loss of processing speed, which places constraints on other cognitive domains. These constraints mean that performance on tests of abilities that rely on processing speed, like memory and mental transformation, becomes increasingly determined by speed. This results in dedifferentiation.

Consistent with this explanation, a number of findings show dedifferentiation when a measure of processing speed is included (Babcock et al., 1997; Ghisletta et al., 2003; Ghisletta et al., 2005; Hertzog & Bleckley, 2001). These patterns exist of measures mental transformation, but also for general knowledge. For example, Ghisletta & de Ribaupierre, (2005) showed that over 5 years, performance on a cross-out test (a measure of speed of processing) was related to change on a knowledge test but the opposite pattern was not found.

However, the patterns of dedifferentiation with general knowledge are likely weaker or occur later in the adult lifespan. This is because knowledge is presumed to accumulate with life experiences like education or occupation whereas memory, and mental transformation tend to be primarily impacted by biological influences (Ghisletta & Lindenberger, 2003; Ghisletta & de Ribaupierre, 2005). Unfortunately, besides Ghisletta & de Ribaupierre (2005) and Hertzog & Bleckley (2001), only a few studies (e.g. Li et al., 2004) have formally tested whether correlations involving speed of processing show greater dedifferentiation or whether there are different patterns of dedifferentiation for correlations between certain cognitive abilities.

In summary, some of the conflicting findings in the literature may be due to the different types of cognitive abilities that were used to test for dedifferentiation. However, this is an issue that could easily be resolved by using a variety of cognitive abilities and explicitly examining whether the magnitude and/or trajectory of dedifferentiation are moderated by the particular types of cognitive abilities.

Different Sample Characteristics

A second source for the inconsistencies in the literature could be the characteristics of the samples used. There are a variety of factors, including the sample sizes, the representativeness of the samples, what age groups were used/how long they were followed, and the proportion of individuals with undiagnosed pathological issues that could influence whether dedifferentiation was found.

Firstly, there is a lot of variation in the size of the studies. As shown in Table S1, sample sizes range from 74 to over 95,000, with the average being around 200-500 participants. Although these samples may appear to be appropriately sized for investigating correlational patterns, the effect size of dedifferentiation is likely small based on the number of studies where no results are found (effect sizes are rarely reported) and therefore may be difficult to consistently detect even with smaller samples. It is likely then that at least some of the inconsistency may be explained by differences in the sample size.

Another issue is the representativeness of the samples. In general, the samples used to investigate dedifferentiation tend to be well educated (for example, younger adults are often college students, Cunningham & Birren, 1980; Hertzog, 1989), and white (but see Sims et al., 2009). This means that even if the results from this work were consistent, they may not extend to people of lower socioeconomic status, of different races, or with physical health conditions.

Representativeness can also be an issue if the samples within a study are not comparable. For example, dedifferentiation is also found when comparing low- and high-ability individuals (Abad, Colom, Juan-Espinosa, & Garcia, 2003; Batterham et al., 2011; Tucker-Drob, 2008). Therefore, if the samples are not comparable in terms of IQ or education, results may incorrectly support age dedifferentiation where none exists or obscure age dedifferentiation that is present.

A third concern is the differences between the age ranges that are being used across studies. Cross-sectional studies have compared young adults with older adults (Babcock et al., 1997; Hertzog, 1989; Schultz et al., 1980), middle-aged adults with older adults (Hertzog & Bleckley, 2001; Hull et al., 2008), and even young-old adults with old-old adults (Hedden & Yoon, 2006; Singer et al., 2003; Vaughn & Giovanello, 2010). Similarly, longitudinal studies have evaluated individuals who can be in their 20's at baseline (Cunningham et al., 1980; Deary et al., 2004), middle aged (de Frias et al., 2007; Finkel et al., 2007), or 60+ (Anstey et al., 2003; Batterham et al., 2011). Finally, these studies also vary in the amount of time the individuals are followed, from less than 10 years (Anstey et al., 2003; de Frias et al., 2007; Ghisletta et al., 2003a; Ghisletta et al., 2005) to over 30 years (Cunningham et al., 1980; Ghisletta et al., 2003; Hulur et al., 2015).

These between-study age differences are a notable concern for two reasons; firstly, the age groups may be too close to one another and or the amount of time elapsed may be too short to detect dedifferentiation. In addition, dedifferentiation may only be found in older adulthood. For example, differences/changes in the correlations may not begin to appear until a certain level of cognitive loss or brain atrophy has already occurred (de Frias et al., 2007). Much of the previous research is consistent with non-linearity (de Frias et al., 2007; Li et al., 2004; Schaie, 1989; Ward, 2000). For example, Waller & Waldman (1990) found that the three-factor model that provided the best fit for all age groups except for the oldest (70-74) where a two-model (a dedifferentiated model) provided the best fit for individuals.

This non-linear pattern could be due to something at the end of the normal adult lifespan. However, it could also be a results of the increasing presence of individuals with pathological or terminal issues primarily associated with older adulthood (Batterham et al., 2011; de Frias et al., 2007; Wilson, Segawa, Hizel, Boyle, & Bennett, 2012). This raises the question of how carefully the older participants were screened for pathological decline.

Conditions like dementia and terminal decline, the period before death, are characterized by increased declines in cognitive performance (Laukka, Macdonald, & Backman, 2006; Lovden, Bergman, Adolfsson, Lindenberger, & Nilsson, 2005; Sliwinski, Hofer, Hall, Buschke, & Lipton, 2003; Wilson et al., 2010) and in the brain (Schneider, Aggarwal, Barnes, Boyle, & Bennett, 2009; Uylings & de Brabander, 2002). For Alzheimer's disease, these impacts appear to be general, with most of the differences between healthy and mild AD patients attributable to a global factor (Salthouse & Becker, 1998). In fact, recent work evaluating dedifferentiation in terminally declining and pathological individuals shows that closeness to death and pathology tended to be better predictors of dedifferentiation than actual age (de Frias, Dixon, & Strauss, 2009; Batterham et al., 2011; Wilson et al., 2012, Sliwinski, Hofer, & Hall, 2003). For example, Sliwinski et al. (2003) found that the correlations between within-person change were stronger for older individuals who had been diagnosed with preclinical dementia compared with those who had not. Similarly, Batterham et al. (2011) found that for 687 adults 70 years and over who were followed for 17 years, a) a time-to-death metric exhibited significant dedifferentiation on four tests whereas age only showed dedifferentiation for two tests, and b) evidence of any dedifferentiation was attenuated when individuals with cognitive impairment were excluded (even though they were only 11.8% of the sample).

Research evaluating configural and metric invariance between healthy individuals and those with pathological conditions is mixed. For example, Siedlecki, Honig, & Stern (2008) found that although the same five-factor structure provided a good configural fit in healthy older adults, individuals with questionable dementia, and probable dementia, metric invariance was not possible. This could be due to larger correlations between measures, although that is not certain. An evaluation of the ADNI neuropsychological battery at baseline found that the covariances between factors could be constrained to the same between a group of less cognitively impaired and more cognitively impaired individuals while still providing an excellent model fit (Park et al., 2012). This suggests that the relations between cognitive abilities are stable between impairment groups and dedifferentiation is not occurring. In Study 2, I will retest this using a model more consistent with the CHC to determine whether similar properties are found. Unfortunately, none of these studies evaluated model fit in multiple conditions across time to determine whether within-person changes are seen with disease progression.

These findings are important because not all studies report whether participants were asked whether they had any pathological conditions, whether they were given a screening measure like the Mini-Mental Status Examination (MMSE; Folstein, Folstein, & McHugh, 1975; e.g. Hertzog & Bleckley, 2001, Sims et al., 2009), or whether any other strictures were implemented to ensure that the older sample was aging normally. The proportions of individuals with some kind of pathological decline could therefore vary between studies, which in turn results in some studies finding and others failing to find evidence of dedifferentiation.

In summary, there are a lot of differences between these studies in regards to important factors like sample size, representativeness, age, and proportion of pathological individuals that could lead to the inconsistent findings. A better approach would be to use a large, representative adult lifespan sample where participants were screened for any health factors that may affect the results. In addition, the role of pathological/terminal changes in older adulthood in explaining dedifferentiation needs to be more extensively examined.

Methodological Factors

In addition to the issues with different cognitive abilities and sample characteristics, there are a variety of methodological differences between the studies that could be responsible for the inconsistencies in the literature. In particular, differences in the reliability of the measures, the variability of the scores, and the analytical methods could produce inconsistencies.

A study is only as good as the measures that are used. Reliability refers to the degree to which a measurement is vulnerable to random influences, and can be measured by having the same participants take a test multiple times (test-retest reliability), or to test whether a participant will perform similarly on one part of a test as they do on another (internal consistency). When reliability is low, there may not be enough systematic variance in the variables to show the true underlying patterns in the data.

Unfortunately, the reliabilities of the tests are not always reported (for example, de Frias, Dixon, & Strauss, 2006 noted that a major problem with executive function measures is their questionable reliability and yet they do not report the reliabilities of their own measures). There is also sometimes variability in reliability within studies for the individual age groups (Babcock et al., 1997; Li, 2004; Lindenberger & Baltes, 1997); for example, Babcock et al. (1997), reported Cronbach alpha scores (a measure of internal consistency) ranging from .53 to .93 in the older sample versus .73 to .94 in the younger sample. Unfortunately, they did not test whether these differences were significant. These differences in reliability both within and between studies could result in the appearance of age patterns that support or refute age dedifferentiation that are not trustworthy. It is therefore important to use only the most reliable measures.

Differences in variability between age groups could also result in misleading results. As shown in Goodwin & Leech (2006), a correlation will generally appear to be stronger if there is greater variability than when there is less variability, a problem known as restriction of range. This is an issue for dedifferentiation because variability can differ between age groups. In addition, because some studies used younger samples that consist of only college students (a homogenous, well-educated group) and older adults that are gathered from the community (a sample that could be more heterogeneous), the groups may differ in variability because of sample selection. Finally, some studies have used various formulas for correcting for restriction of range in the magnitude of correlations in different groups (Deary et al., 2004; Escorial et al., 2003; Juan-Espinosa et al., 2002). However, many do not do this (Adrover-Roig et al., 2012; Ghisletta & Lindenberger, 2003), which could explain some of the discrepancies between studies.

Finally, previous studies have used different types of analytical methods to test for dedifferentiation. The choice of analysis is important because different approaches allow researchers to test different hypotheses. For example, an analysis looking just for increasing variance explained by g doesn't evaluate whether those increases are primarily found for certain types of cognitive abilities, or whether the same number of abilities constructs are maintained. Equally importantly, certain methods can produce more certainty in ones' results than others. For example, visually comparing correlations does not provide information about whether there is a statistical difference or about the size of that difference. Similarly, comparing correlations at the latent level as opposed to the observed level is preferable because latent factors are better representations of their constructs.

In conclusion, there are notable differences in the methodological choices made across studies that could potentially influence whether dedifferentiation is found. In order to conclusively test for dedifferentiation, highly reliable measures need to be used, differences in variability need to be considered, and analyses that are best suited for answering important theoretical issues should be conducted.

Summary of Literature

Understanding whether cognitive dedifferentiation occurs is important both methodologically and theoretically. However, the research on this topic is notably inconsistent. This is likely due to a variety of factors, most importantly the differences in the cognitive abilities included, the samples collected, and the methodological approaches used.

In regards to the cognitive abilities included, several explanations for dedifferentiation and previous work on dedifferentiation suggest that processing speed may have an important role (Li et al., 2004; Ghisletta & Lindenberger, 2003). Based on these findings, I hypothesize that dedifferentiation will be stronger for correlations that include processing speed. I also suspect the dedifferentiation will be weaker or appear later for correlations that include knowledge. It is therefore important to include correlation type as a moderator or to examine individual loadings to determine where changes in correlation strength are occurring.

Studies also differ on a number of important sample characteristics. For example, studies often vary in both the number and representativeness of participants, thus limiting the generalizability of their results. There are also differences between studies in age groups used and/or in the amount of change measured. This matters because dedifferentiation may be non-linear (de Frias et al., 2007), with the changes in correlation magnitude beginning only in older adulthood and not when cognitive abilities first begin to decline.

Finally, the proportion of individuals with pathology is important because it appears that pathology is linked to dedifferentiation. Therefore, a large number of pathological participants in an otherwise healthy sample could lead to the appearance of dedifferentiation in normal aging (Batterham et al., 2011). To examine dedifferentiation in healthy adults, it is necessary to use a large representative adult lifespan sample that has been screened for conditions that could influence cognitive function. To examine the relationship between dedifferentiation and pathology, more research is needed on (a) individuals with different types of pathology and (b) the progression of those pathologies using more measures and more sophisticated analytic techniques.

In regards to methodological issues, unreliable measures, differences in variability between age groups, and inappropriate analytic techniques limit our ability to test for dedifferentiation. Measures should be reliable and differences in variability at least be examined; analyses should be chosen that can answer theory-driven questions and provide conclusive results.

Based on these conclusions regarding the dedifferentiation literature, my thesis consists of two studies. The first study is a meta-analysis where I tested whether dedifferentiation was found cross-sectionally in a series of large, representative adult lifespan samples that were screened for health conditions using a self report. The correlations were collected from normed cognitive tests where the tests are reliable and information is available about the variability in performance at different ages. I also examined the trajectory of dedifferentiation across the adult lifespan, and whether it is affected by the types of cognitive abilities included in the correlation.

In the second study I tested for increased correlations between latent factors using longitudinal data. The data for this study were collected over three years, included healthy and pathological participants, and used reliable tests. In this study, the goal was to explore how MCI and Alzheimer's disease were related to correlation magnitude compared with healthy adults, and evaluate whether the progression of these diseases resulted in additional dedifferentiation.

Study 1: Cross-Sectional Dedifferentiation in Healthy Adults

The goal of the first study was to test whether cognitive dedifferentiation occurs in the adult lifespan using a variety of datasets collected as part of the norming process for cognitive batteries like the Wechsler Adult Intelligence Scale (WAIS; Wechsler, 1955; 1981; 1997; 2008b), Wechsler Abbreviated Scale of Intelligence (WASI; Wechsler, 1999), Wechsler Memory Scale (WMS, 1987; 1997; 2009), Kaufman Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1992), Kaufman Brief Intelligence Test (KBIT; Kaufman & Kaufman, 2004), and Woodcock-Johnson (WJ; Woodcock & Johnson, 1989).

To develop the standards that are used to determine how an individual performs in comparison to a hypothetical "average peer", also known as norms, a large, representative sample needs to be collected. These samples were designed to include a quota of individuals from specific regions in the country, educational levels, and racial/ethnic backgrounds. Different test batteries include other factors, such as religious affiliation, consumption of alcohol, etc. In addition, many of these test batteries have exclusionary criteria for individuals with conditions that impair cognition function (e.g. WAIS, WASI, and WMS batteries), although this is not always the case (e.g. Woodcock Johnson batteries).

Another benefit of using normed cognitive tests is that the correlations between tests, means, standard deviations, and reliabilities are available for several age groups. Although the test batteries do differ in the number of age groups (some include only three adult samples where others have up to eleven), all tests included in these batteries have been shown to have acceptable reliability across the age groups (> .60) and the differences in variability between age groups were small. This meant that I could have confidence that any differences found were unlikely to be due to reliability or variability issues. Therefore, using these samples resolves some of the methodological issues present in earlier work.

Hypotheses

H1a. That dedifferentiation will be found if it occurs in healthy adulthood. I expected that if dedifferentiation occurs in normally aging adults, whether due to the changes in the brain and cognitive abilities that accompany senescence (Ghisletta & Lindenberger, 2003; Li et al., 2004), age would positively predict the size of the correlations. I further explored these positive relations when they occurred by comparing the average correlation size in a young (less than 48 years) and old group (more than 48 years). However, as recent work has suggested that pathology may be a primary factor in dedifferentiation (de Frias et al., 2009; Batterham et al., 2011; Wilson et al., 2012, Sliwinski et al., 2003a), it is possible that dedifferentiation will not be found in these primarily healthy samples.

H1b. *Certain combinations of cognitive abilities (i.e. memory and knowledge vs. speed and memory) may be more or less likely to exhibit dedifferentiation.* There is already evidence that certain combinations of abilities may be more likely to dedifferentiate with age. For example, it has been argued that declines in speed of processing compress performance on other abilities, which results in dedifferentiation (Li et al., 2004). Therefore, I predicted dedifferentiation would be strongest between speed of processing and abilities that heavily rely on it, like memory and mental transformation.

I expected dedifferentiation for correlations between memory and mental transformation tests as both they are dependent on speed. However, since speed was not directly assessed in these correlations I expected that the patterns might be weaker. I also hypothesized that dedifferentiation for correlations between speed and knowledge would be weaker, as increasing biological constraints in speed should impact knowledge less than abilities like memory and mental transformation. Dedifferentiation should be weakest for correlations that include knowledge and do not include speed, for the reasons outlined above.

To test whether certain combinations of abilities, hereafter referred to as types, show more dedifferentiation, I tested multiple contrasts where I set one type (e.g. speed with mental transformation) as a baseline or reference and compared the age patterns in the others (e.g. speed with memory, mental transformation with memory) to this reference. I further explored significant deviations when they occurred by conducting a second set of models where age predicted correlation size for each type individually. I also evaluated the average correlation size in a young (less than 48) and old group (more than 48) for each type when age was a significant predictor.

H2a. Patterns of dedifferentiation are non-linear, with differences appearing primarily in older adulthood. Previous work suggests that dedifferentiation may not occur until older age (Balsamo & Romanelli, 2010; de Frias et al., 2007; Li et al., 2004; Waller & Waldman, 1990), and not in the linear pattern seen with cognitive decline (Salthouse, 2010). If this is the case, then correlations between abilities should not begin to increase until older adulthood, beginning around age 70. I chose age 70 because most findings of non-linear dedifferentiation appear at that age (de Frias et al., 2007; Ward et al., 2000; Waller & Waldman, 1990). If dedifferentiation is non-linear because a certain threshold of decline is needed, I hypothesized that quadratic age (age^2) would predict correlation size. However, if the non-linear patterns found in previous work were due to the increasing presence of pathology in older adults (de Frias et al., 2009), I did not expect to find it in this generally healthy sample. I explored significant age^2 patterns present in conjunction with a significant age effect (suggesting non-linear differentiation/dedifferentiation) by examining average correlation size in three age groups (18-39,40-69,70+).

H2b. *Correlations between different types of cognitive abilities (i.e. memory and knowledge vs. speed and memory) may vary in when dedifferentiation begins.* Dedifferentiation may be expected to occur later for knowledge than for other abilities because it relies less heavily on speed than memory or mental transformation. This would mean stability in younger and middle age groups and larger correlations in older groups. To test whether certain combinations show more dedifferentiation, I conducted multiple contrasts where I set one type (e.g. speed with mental transformation) as a reference and compared the age^2 patterns in the additional types (e.g. speed with memory, mental transformation with memory) to that reference. I explored a significant age^2 predictor present in conjunction with a significant age predictor (suggesting non-linear differentiation/dedifferentiation) for each specific type by conducting a second set of models where age and age^2 predicted correlation size for each type individually and examining average correlation size in three age groups (18-39,40-69,70+).

Method

Inclusion Criteria

I collected a convenience sample of all the manuals containing information on the standardized test batteries with norms that could be easily obtained within the University of

Virginia community. This search yielded a total of nineteen different batteries or versions of batteries (i.e. WAIS IV versus WAIS III) that were then evaluated for the project. Batteries were included if they had correlation coefficients between tests for at least three adult age groups, the tests had acceptable validity and reliabilities over .6, and the batteries were normed using a nationally representative sample. Eleven of the batteries fit these conditions, and are described in greater detail in Table S2. For all test batteries, the samples were representative in terms of race/ethnicity, geographic region, and education level based on the most recent census data available at norming.

Classification of Tests & Correlations

The tests contributing to each correlation were categorized based on previous factor analytic work (Benson, Hulac, Kranzler, 2010; Bowden, Cook, Bardenhagen, Shores, & Carstairs, 2004; Flanagan & Mcgrew, 1998; Ortiz, Flanagan, & Alfonso, 2013; Hoelzle, Nelson, & Smith, 2011; Tulsky & Price, 2003), with some additions based on my own judgment and the ratings of another individual with extensive knowledge of the various tests. Details for each test, including name, battery, description (from the relevant technical manual), and classification are in Table S3. Below I explain in greater detail how I made the classifications for each battery. As the classification studies often used different names for their factors, I included the Cattell-Horn-Carroll (CHC) classification in parentheses for all abilities.

WAIS Batteries

The WAIS battery was evaluated using a cross-battery assessment with the Luria-Nebraska Battery (Shelly & Goldstein, 1982). These four factors can be interpreted as representing general intelligence (g), a combination of reading/writing (Grw), quantitative (Gq) and short term memory (Gsm), comprehension-knowledge (Gc), and a combination of visual abilities (Gv) and fluid intelligence (Gf), respectively. Picture completion loaded onto Gsm and Gc and was categorized as primarily mental transformation with general knowledge. To maintain consistency across the WAIS batteries I additionally classified arithmetic and similarities as general knowledge and mental transformation, respectively even though these cross loadings were not significant in Shelly and Goldstein's (1982) analysis.

The WAIS-R battery was evaluated in a cross-battery analysis with WMS-R (Bowden et al., 2004). Although their factor structure was not developed with the CHC in mind, I believe their six-factor structure can be interpreted as follows: verbal comprehension (Gc), perceptual organization (Gv/Gf), processing speed (Gs), working memory (Gsm), verbal memory (Glr), and visual memory, (Glr). Tests that loaded onto Gc were categorized as General Knowledge. Tests that loaded on Gv/Gf and Gsm were categorized as Mental Transformation. Tests that loaded on Gs were categorized as processing speed. No tests loaded onto Glr. Although no cross loadings were tested in Bowden et al. (2004), to maintain consistency across the batteries I additionally categorized arithmetic, similarities, and picture completion as general knowledge, mental transformation, and general knowledge, respectively.

The WAIS III battery was evaluated in the cross-battery analysis of the WAIS-III and the WMS-III batteries (Tulsky & Price, 2003). As with the WAIS-R analyses, I interpreted their factors in terms of CHC abilities: verbal comprehension (Gc), perceptual organization (Gv/Gf), auditory memory (Glr), working memory (Gsm), visual memory (Glr), and perceptual speed (Gs). Tests that loaded on Gs were categorized as speed. Tests that loaded on Gf or Gsm were categorized as mental transformation. Tests that loaded on Gc were categorized as general knowledge. None of the WAIS tests loaded onto Glr. Although picture arrangement loaded onto crystallized intelligence, to maintain consistency across other WAIS batteries (where it never loaded on Gc), it was categorized only as mental transformation.

The WAIS IV classifications were reviewed in Ortiz et al. (2013). They categorized the WAIS IV tests into six abilities: verbal comprehension (Gc), perceptual speed (Gs), fluid reasoning (Gf), short-term/working memory (Gsm), quantitative knowledge (Gq), and visual processing (Gv). Tests that were categorized as Gc or Gq were categorized as general knowledge. Tests that were categorized as Gf or Gsm were categorized as mental transformation. Tests that were categorized as Gs were categorized as speed. For tests that represented multiple abilities, the main factor as specified by Ortiz et al. (2013) was used as the primary categorization.

The WASI battery was evaluated in a cross-battery CFA with the Wide Range Intelligence Test (WRIT; Canivez, Konold, Collins, & Wilson, 2009). They evaluated a two-factor structure including crystallized intelligence (Gc) and fluid intelligence (Gf). Tests from the WASI battery loaded onto both factors. Tests that loaded onto the Gc factor were categorized as general knowledge. Tests that loaded onto the Gf factors were categorized as mental transformation. To maintain consistency across the WAIS batteries, similarities was additionally coded as mental transformation.

WMS Batteries

The WMS-R battery was evaluated in a cross-battery CFA with the WAIS-R in Bowden et al., (2004). As noted above, I interpreted their six factors in terms of CHC theory. Tests that loaded onto Glr were treated as memory. Tests that loaded on Gsm were treated as mental transformation. There were no cross-loadings in this analysis.

The WMS III tests were categorized used an EFA from Hoelzle et al. (2011) and the cross-battery analysis of the WAIS-III and WMS-III batteries from Tulsky and Price (2003). The Hoelzle et al. (2011) analysis showed that across the nine normative samples, all the tests loaded onto one memory factor. Tulsky and Price (2003) found that spatial span and letter number sequencing primarily loaded onto a working memory factor, and were categorized as mental transformation. Cross-loadings with the Gv/Gf factor were with the visual memory tests, and seem to be more about visual processing than mental transformation and these patterns were not seen for any of the other WMS batteries. Therefore, no additional types were assigned based on these relations.

The WMS IV was evaluated with exploratory principal component analyses in Hoelzle et al. (2011). They found a robust fit across age groups for a two-factor model that included visual and auditory memory abilities. However, spatial addition and symbol span were designed to measure working memory, and to maintain consistency across batteries these tests were classified as mental transformation. All other tasks were categorized as memory. There were no cross-loadings.

Kaufman Batteries

The KAIT battery was evaluated in a cross-battery CFA with the Woodcock Johnson-Revised in Flanagan and Mcgrew (1998). They evaluated a nine-factor structure including crystallized intelligence (Gc), fluid intelligence (Gf), memory span (Gsm), perceptual speed, (Gs), associative memory (Glr), visual memory (Gsm), closure speed (Gv), phonetic coding (Ga), and reading (Grw). Tests from the KAIT battery loaded onto the Gf, Glr, Gsm, and Gc factors. Tests that loaded onto the Gf and Gsm factors were categorized as mental transformation. Tests that loaded onto the Glr were categorized as memory, with the exception of memory for block designs. As noted in the technical manual (Kaufman & Kaufman, 1992), this test has also been used to assess Gf and was more consistent in content with the tests that were categorized as Gf. Therefore, I classified it as mental transformation. Although the auditory delayed recall loaded onto both Glr and Gc, it was categorized as memory as the main task is to recall information. Tests that loaded onto the Gc factor were categorized as general knowledge. When tests loaded onto multiple factors, the highest loading was used for categorization.

I could not find a factor analysis of the KBIT battery. However, the battery was designed to tap crystallized intelligence (Gc, Vocabulary) and fluid intelligence (Gf, Matrices) and therefore I classified the tests using these a priori assumptions (Kaufman & Kaufman, 2004).

Woodcock-Johnson Battery

The Woodcock-Johnson-R (WJ-R) battery was developed using the GF-GC theory (which was a partial basis for the CHC), and therefore I used the classifications included in the technical manual (Woodcock, & Johnson, 1989). The achievement tests were not included. In addition, the sound blending, incomplete words, picture recognition, and visual closure tests were not included because they only measured auditory and visual processing abilities and were specifically designed not to measure fluid intelligence (Gf) or speed (Gs). For tests that were categorized as multiple abilities, the primary factor was chosen as specified by the technical manual.

These classifications resulted in ten correlation types (MT-MT, MT-S, MT-GK, MT-M, S-S, S-GK, S-M, M-M, M-GK, GK-GK). Most of these groups were over 200 and all the between ability types were above 200 with the exception of Speed with Memory (S-M) and Knowledge with Memory (K-M), see Supplementary Table 4. Therefore, I chose not to aggregate but to be cautious in interpreting the S-M, K-M, and small within-ability types.

Data Cleaning

I used R (V.3.30; R Development Core Team, 2008) for all the analyses in Study 1. The data included the following variables (Test Battery, Sample Size (N), Mean Age, Correlation, Type, Sample, Age^2 R-Z Correlations, Weight), the last four of which are described in greater detail below.

One assumption of a basic meta-analysis is that the outcome measures are independent of one another. This is not the case in this study because a number of correlations were drawn from each sample. To resolve this dependency issue, I included a random effect for sample because it allowed me to keep the complexity of the data without producing biased estimates from assuming independence where dependencies exist (Van den Noortgate, Lopez-Lopez, Marin-Martinez, Sanchez-Meca, 2013). The 'sample' variable identifies which correlations were drawn from which independent sample and can be used to control for this issue. The WAIS-3 and WMS-3 were coded with the same identifier as were the WAIS-4 and WMS-4 because they relied on the same samples.

I used the 'meta-for' package (Viechtbauer, 2010) to obtain Fisher's Z transformations of the original correlations (R-Z Correlations) and a measure of their relative sample size (Weights). Fisher's Z is a commonly used variance-stabilizing transformation that controls for the fact that the variances of untransformed r's are increasingly smaller as they approach the upper and lower bounds (-1,1) of their distribution. This transformation allows for unbiased comparisons between correlations. The correlations are weighted by the number of individuals included in the study. To test my hypotheses regarding non-linear patterns, I standardized my age variable and squared it to make an age^2 variable which was always entered simultaneously with age.

My hypotheses regarding dedifferentiation primarily focused on correlations across abilities and not within abilities. Therefore, I excluded within ability correlations (S-S, MT-MT, M-M, GK-GK) from the analyses described below.

Testing Assumptions

I first evaluated whether my data met the assumptions needed to conduct the meta-regression. First I examined whether the dependent variable, correlation size, was normally distributed. Unfortunately, no normality tests have been developed in R that control for the different weightings of the correlations and the dependencies between the correlations (multiple correlations being drawn from one sample). However, I can use a quantile-quantile figure to check visually for normality.

In a Q-Q plot, the correlations are organized into ascending order and assigned to quantiles based on frequency (these are the sample quantiles, y-axis). The correlations are then categorized into the quantiles they would in theory be in if the data was perfectly normal (these are the theoretical quantiles, x-axis). Evidence of normality is found when a majority of the points are between -1.96 and 1.96 and the data forms a diagonal line with the lowest point in the bottom left corner and the highest point in the upper right corner. This plot is presented in Figure 1. Based on the findings that 1) the majority of the points (particularly those based on larger sample sizes) are in the middle of the distribution and 2) the pattern is almost perfectly linear and diagonal, I determined that my dependent variable likely met the assumption of normality.

I then examined the variance present in the data in several different ways. First, I wanted to determine how much variability was present in my dependent variable. To do this, I constructed a random-effects model that controlled for the dependencies present in the data. As expected, the variance between the correlations, or τ^2, is significantly non-zero. I calculated a measure of heterogeneity, I^2, by subtracting the df from the chi statistic produced from this analysis (Cochran's Q) and dividing this difference by the Cochran's Q (Higgins & Thompson, 2002). I found that 72.80% of the total variation in the correlations was unexplained after controlling for the intercept.

I then tested for homogeneity of variance across age by type and age^2 by type. Unfortunately, as with normality, there is no test for heterogeneity in R that can control for differentially weighted items or dependencies. However, I can visually test for homogeneity of variance by plotting the predicted values from models that included either age*type or age^2*type as predictors and on the x-axis and the residuals for each correlation after controlling for these predictors on the y-axis. Homogeneity of variance is shown when the vertical spread of the points is similar across the x-axis. As shown in Figure 1, there is slightly less spread on the y-axis at certain places. However, these places appear to be primarily where there are fewer points altogether (and therefore less extreme values, which are predicted less well). Therefore, it appears that the homogeneity of variance assumption is likely met.

Analytic Plan

If aging is associated with increased dedifferentiation, I expected to find that age would significantly and positively predict the size of the correlations and that age patterns would be higher in the older versus younger groups. If a certain threshold of decline is needed before dedifferentiation occurs, or if dedifferentiation is primarily due to pathological conditions more common in older adults, I expected that age^2 would significantly predict correlation size and that age patterns would primarily be higher in the oldest versus the younger groups. I also expected that correlations involving processing speed (e.g. S-MT, S-M) would be more likely to show dedifferentiation than correlations between tasks that do not include processing speed.

I constructed a set of five models that included the intercept, age, and age^2 for one of the cognitive ability type (hereafter referred to as the reference) and the deviations from those intercept, age, and age^2 predictions for the other types of correlations. Setting a reference was necessary because ability type was a categorical variable. Comparing the other types to a reference also allowed me to test my hypotheses that dedifferentiation would be stronger for certain combination of variables than others. Finally, I examined age and age^2 as predictors when only the data for each of the types was included individually.

Significant results for the intercept, age, and age^2 variables in the reference show that those variables significantly predicted correlation size in that cognitive ability type. For example, if speed with mental transformation was the reference and age was significant, this would mean that age significantly predicted correlation size for correlations that include tests of speed and memory. A significant deviation would demonstrate that the intercept is higher/lower or that age/age^2 did not predict correlation size in the same way for that ability type compared with the reference. For example, if the deviation between speed with knowledge and the reference is significant, that would mean that age is a stronger or weaker predictor of correlation size in speed with knowledge compared with the reference.

I used the results in each reference category to determine whether there was evidence of linear or non-linear dedifferentiation for that type. I used the deviations to determine when those linear and non-linear patterns were significantly stronger or weaker across types.

Unfortunately, formal outlier or influence analyses are not available for the specific type of model I needed to control for dependencies. However, I can visually examine whether a small number of correlations seem to be responsible for my patterns of results. To do this, I divided the data into separate data sets for each correlation of two specific tests. I tested whether age predicted correlation size for each of these data sets for every combination of tests and plotted the results for each cognitive ability type. I then examined these plots to determine whether there were any points or clusters of points that had much higher or lower correlations than the rest. If this appeared to be the case, I investigated the correlations to determine if there was any relationship (larger age ranges, same tests) between these outlier points.

Results

Speed with Mental Transformation as Reference Type

The estimates of the meta-regression model where speed with mental transformation served as the reference are presented in Table 1. The average correlation was significantly non-zero, age was a significantly positive predictor (signaling dedifferentiation), and age^2 was not a significant predictor (suggesting dedifferentiation is linear) for the Speed with Mental Transformation type (S-MT). These patterns were also present in additional models that included age and age^2 as predictors of just the Speed with Mental Transformation correlations, see Tables 6 and 7. The finding that age was a significant predictor was consistent with my hypothesis that dedifferentiation would be found for the S-MT type.

Age predicted an increase in correlation size of .0020 with each additional year after 18. These patterns are presented in Figure 2. This suggests a very small impact on correlation size across the adult lifespan; between the ages of 25 and 75 the average correlation would only differ by about .10. The average correlation below age 48.29 (the mean of the sample, M = 27.67) was .37. The average correlation above 48.29 (M = 65.01) was .44. This equals an average difference of .0020 with each year, identical to that found in the model.

The relations between age and predicted correlation size for each combination of tests are presented in Figure 3. The weighted mean of these standardized estimates is around .39 and show that findings of the above model do not appear to be due to a small number of data points.

There were significant deviations in the predictions for the S-MT type compared with the other types for the size of the intercept, the strength of age as a predictor, and the strength of age^2 as a predictor. The intercept deviations were positive for K-M and K-MT, which means that the average correlations were higher for these types than for S-MT. The intercept deviations were negative for the M-S and MT-M, which means that the average correlations were lower for these types than for S-MT. There was no significant deviation between S-MT and S-K.

The age deviations were negative for the K-M, M-S, K-MT, and MT-M types. This means age was a significantly weaker predictor of correlation size in these groups compared with S-MT type. There was no significant deviation between S-MT and S-K types. The findings that age dedifferentiation was weaker for the K-M, K-MT, and MT-M types were consistent with my hypotheses that dedifferentiation would be strongest for correlations between tests of speed and tests of mental transformation or memory. I expected that S-K would have weaker dedifferentiation, but this was not supported by the data. There was also a significant differences between the S-MT and S-M group which was unexpected, but may not be reliable given the small number of data points in this group.

The age^2 deviations were negative for the K-M, M-S, K-MT, and MT-M types. These results suggest that dedifferentiation is less positively non-linear in these groups.

Speed with Knowledge as Reference Type

The estimates of the meta-regression model where speed with knowledge served as the reference are presented in Table 2. The average correlation was significantly non-zero, age was a significantly positive predictor (signaling dedifferentiation), and age^2 was not a significant predictor (suggesting dedifferentiation is linear) for the speed with knowledge type (S-K). In the additional models that included age and age^2 as predictors of just the speed with knowledge correlations, see Tables 6 and 7, age was no longer a significant predictor ($p = .0562$). The difference between the models seems to be due to the simultaneous inclusion of the deviations. The finding that age approached significance/was a significant predictor was consistent with my hypothesis that dedifferentiation would be found for the S-K type.

Age predicted an increase in correlation size of .0019 with each additional year after 18. These patterns are presented in Figure 2. This suggests a very small impact on correlation size across the adult lifespan; between the ages of 25 and 75 the average correlation would only differ by about .10. The average correlation below age 48.29 (the mean of the sample, $M = 27.65$) was .36. The average correlation above 48.29 ($M = 65.17$) was .44. This equals an average difference of .0020 with each year, identical to that found in the model.

The relations between age and predicted correlation size for each combination of tests are presented in Figure 3. The weighted mean of these standardized estimates is around .39. The findings of the above model do not appear to be due to a small number of data points.

There were significant deviations in the predictions for the speed with knowledge correlations compared with the other types for the size of the intercept, the strength of age as a predictor, and the strength of age^2 as a predictor. The intercept deviations were positive for K-M and K-MT, which means that the average correlations were higher for them than for S-K. The intercept deviations were negative for the M-S and MT-M types, which means that the average correlations were for them than for S-K.

The age deviations were negative for the K-M, M-S, K-MT, and MT-M types. This means age was a significantly weaker predictor of correlation size for them compared with S-K type. The findings are somewhat surprising, as I expected dedifferentiation would likely be comparable or stronger for the MT-M type because memory and mental transformation are more dependent on speed than knowledge is. However, it is possible that patterns are stronger here because speed is assessed directly as part of the S-K correlations and only indirectly for in the MT-M correlations.

The age^2 deviations were negative for the K-M, M-S, K-MT, and MT-M types. These results suggest that dedifferentiation are less positively non-linear in these types.

Mental Transformation with Memory as Reference Type

The estimates of the meta-regression model where mental transformation with memory served as the reference are presented in Table 3. The average correlation was significantly non-zero, age was not a significant predictor, and age^2 was a significant predictor for Mental Transformation with Memory (MT-M). In the additional models that included age and age^2 as predictors of just the Mental Transformation with Memory correlations, see Tables 6 and 7, age^2 was no longer a significant predictor ($p = 0.2370$). The difference between the models seems to be due to the simultaneous inclusion of the deviations. These patterns are inconsistent with my hypothesis that dedifferentiation would be present, albeit weaker, because both mental transformation and memory are both impacted by lower speed of processing capacity.

Relations between age and correlation size for each combination of tests are presented in Figure 3. The weighted mean of these standardized estimates is around -.08. The findings of the above model do not appear to be due to a small number of data points.

There were significant deviations in the predictions for the MT-M type compared with the other types for the size of the intercept, the strength of age as a predictor, and the strength of age^2 as a predictor. The intercept deviations were positive for the K-M and K-MT types, which means that the average correlations were higher for them than for MT-M. There was no significant deviation between MT-M and M-S.

There were no significant deviations for age between any of the types. This further supports the above findings of stability for mental transformation. The age^2 deviations was positive for the K-MT type. These results suggest that the non-linear patterns are more positive for the K-MT type.

Mental Transformation with Knowledge as Reference Type

The estimates of the meta-regression model that included Mental Transformation with Knowledge (K-MT) as a reference are presented in Table 4. The mean correlations were significantly non-zero. Age and age^2 were not significant predictors (no dedifferentiation). These patterns are presented in Figure 2. No differences were found in additional models that included age and age^2 as predictors of just the Knowledge with Mental Transformation correlations, see Tables 6 and 7. These results were consistent with my hypotheses that dedifferentiation would be weak or nonexistent for correlations including knowledge (which I predicted would be less dependent on speed) and where speed was not directly assessed.

Relations between age and correlation size for each combination of tests are presented in Figure 3. The weighted mean of these standardized estimates is around .12. The findings of the above model do not appear to be due to a small number of data points.

The intercept deviation was positive for the K-M type and weaker for the M-S, which means that the average correlations were higher and lower for these types compared with K-MT, respectively. There were no significant deviations in age as a predictor between types. The age^2 deviations were negative for the K-M and M-S types. These results suggest that dedifferentiation may be non-linear in these groups.

Speed with Memory as Reference Type

The estimates of the meta-regression model where Speed with Memory served as the reference are presented in Table 5. The average correlation was significantly non-zero. Age and age^2 were not significant predictors suggesting stability in the correlations for the Speed with Memory type. In additional models that included age and age^2 as predictors of just the Speed with Memory correlations, see Tables 6 and 7, age and age^2 are still not significant but the direction of the age patterns has reversed. The difference between the models seems to be due to the simultaneous inclusion of the deviations.

Relations between age and correlation size for each combination of tests are presented in Figure 3. The weighted mean standardized estimate is around .34. The small number of points likely explains why no dedifferentiation was found for this type.

The intercept deviation was positive for the M-K type, which means that the average correlations were higher for M-K than S-M. The age deviations were positive for the M-K type. This means age was a significantly stronger predictor of correlation size for M-K compared with S-M type. There were no significant deviations in age^2 as a predictor.

Memory with Knowledge without other Types

The estimates of the meta-regression model that included age and age^2 as predictors of just the Memory with Knowledge correlations are presented in Tables 6 and 7. The mean correlations were significantly non-zero, age was not a significant predictor (no dedifferentiation) and there were no quadratic age patterns. Linear age patterns are presented in Figure 2. Relations between age and correlation size for each combination of tests is presented in Figure 3. The weighted mean of these relations is around .09. The findings of the above model do not appear to be due to a small number of data points. These results were consistent with my hypotheses that dedifferentiation would be weak or nonexistent for correlations including knowledge (which I predicted would be less dependent on speed) and where speed was not directly assessed.

Summary

In summary, the Speed with Mental Transformation showed evidence of age dedifferentiation. As expected, dedifferentiation was stronger in these correlations compared with the correlations that included a test of knowledge but not speed and in comparison to the mental transformation with memory type. The lack of a significant effect for the speed with memory type is not surprising given the low number of samples for that type.

However, there was no significant difference between the speed and general knowledge and the speed with mental transformation types as expected. This may suggest that lower levels of speed constrain performance on other cognitive tests similarly. It was also unexpected that correlations of mental transformation with memory would show no dedifferentiation. However, given the small size of the effects (less than an increase of .10 from 25 to 75) even when speed is directly included in the correlation, it may not be surprising that dedifferentiation was not identified for this combination of abilities.

I predicted that the amount of dedifferentiation should be stronger for older age groups, and that non-linear patterns would be particularly strong for the speed with general knowledge type. However, although there were differences in the strength of the age^2 predictor between types, age^2 was never a significant predictor in conjunction with age (dedifferentiation) and was only inconsistently significant for the mental transformation with memory type. As age^2 was never reliably significant, I did not explore correlations in three different age groups.

Constrained Analyses

One of the issues with the above analyses is that some of the tests assessed multiple abilities. This means that some of the ability types specified in the analysis above contained tests that measured something in addition to the ability they were primarily categorized as, which could have biased results.

I reran all the analyses after excluding these tests and evaluated whether similar patterns were produced for both analyses. As shown in Supplementary Tables 5-11, the patterns found with the constrained analyses were almost identical to that of the unconstrained analyses, even though the number of correlations for most types were reduced, see Supplementary Table 4. The only major difference was that age was a weaker predictor of speed with knowledge compared to speed with mental transformation. Therefore, I believe that my findings are not generally not biased due to the inclusion of these tests that assessed multiple abilities.

Discussion

The goal of the first study was to evaluate whether dedifferentiation would be found, whether it would be stronger in older age, and whether these patterns would be impacted by the cognitive abilities included in the correlations being compared. I found evidence of age dedifferentiation when a speed test was included in the correlation, although the effect was small. In addition, significant nonlinear patterns were not found for any of the types, even though there were significant deviations in how strong age^2 was a predictor.

Linear Age Effects

Overall, it appears that dedifferentiation is the exception, not the rule across the adult lifespan. Consistent with previous reports where dedifferentiation was found when a speed of processing test was correlated with a test of memory and/or mental transformation (Babcock et al., 1997; Ghisletta et al., 2003; Ghisletta et al., 2005; Hertzog et al., 2001) and with hypotheses that lower levels in speed at older ages compress performance on other cognitive tests (Li et al., 2004), dedifferentiation was present for two of the three categories that that included speed (and given that K = 28 for the Speed with Memory type it is not surprising that there were no significant age trends for that type). However, dedifferentiation was not found in any of the other correlation types and the effects were relatively small even for correlations that included (less than .11 increase between 25-75).

I expected that dedifferentiation would be stronger for the speed with mental transformation correlations than speed with general knowledge correlations because general knowledge tests are less reliant on speed of processing and more reliant on life experiences compared with mental transformation tests. This difference in age patterns between the Speed with general knowledge and the speed with mental transformation types was not significant in the overall sample, but did appear for the constrained analyses. These results are consistent with my hypothesis that speed may not constrain performance on tests that purely assess knowledge as much as tests that assess mental transformation.

I predicted that I would find dedifferentiation for the Mental Transformation with Memory type, as both abilities can be postulated to be reliant on speed. However, age was not a significant predictor for this type. This may not be surprising, however, given that the magnitude of dedifferentiation was small even when a direct test of speed was included; indirect impacts of speed may be too minimal to capture.

Finally, I expected to find little to no dedifferentiation in the general knowledge with mental transformation and general knowledge with memory types because there was no direct test of speed and because general knowledge is less reliant on speed as noted above. These hypotheses were consistent with the data.

It is important to note that there may be power issues in detecting effects. Unfortunately, there is little to no information about evaluating power for meta-regression mixed-effects models that include dependent effect sizes. However, factors that can influence the power include number of independent effect sizes, between-studies heterogeneity, and number of predictors included. There was a moderate amount of heterogeneity in each type and although there were a large number of correlations, many of them were dependent on one another. Therefore, although my findings are likely accurate that correlations that included speed produced the strongest dedifferentiation, a larger sample may show there are additional differences between ability types.

Non-Linear Age Effects

I hypothesized that the increase in correlations would be stronger in older than younger adults based on hypotheses that dedifferentiation occurs after either a certain amount of cognitive/neural decline (de Frias et al., 2007) or because of pathological conditions more common in older adults (de Frias et al., 2009; Batterham et al., 2011; Wilson et al., 2012, Sliwinski et al., 2003). I also expected that because general knowledge is less dependent on speed compared with memory and mental transformation, greater deterioration in speed (i.e. a bigger threshold) may be needed before dedifferentiation occurs for correlations of speed with general knowledge compared to correlations of speed with mental transformation or memory. However, although there were differences between the ability types in how strong a predictor age^2 was, age^2 only predicted correlation size for mental transformation with memory and there was no accompanying significant age effect.

It is possible that the lack of significant quadratic patterns in the individual groups, despite the significant deviations, are due to power issues. As described above, the power of my analyses may be limited by the relatively small number of independent correlations and the between-studies heterogeneity. Therefore, more work is needed to establish whether increased dedifferentiation in older samples would be demonstrated when more independent correlations are included.

However, these patterns may be accurate given that the decline of speed is linear across the adult lifespan (Salthouse, 2010). In addition, the norming samples for many of the test batteries (WAIS, 1955; 1981; 1997; 2008b; WASI, 1999; WMS, 1987; 1997; 2009; although not all, Woodcock & Johnson, 1989) included screenings for health conditions that affect cognitive functioning, like possible dementia. If the non-linear patterns found in previous studies were due to a subsample of individuals with pathological issues (Batterham et al., 2011), then it is reasonable that such patterns would not be replicated here.

Categorization using CHC

As described above, the evidence in favor of dedifferentiation was mainly found for specific combinations of cognitive abilities. This raises an important question: are these true patterns or are they artifacts of the categorization process? The factor analytic work was not ideal, as it was conducted in multiple samples ranging in size with different methodological analytic techniques. However, it did provide some objectivity and the use of cross-battery analyses whenever possible provided additional links between the batteries.

A secondary issue is that some of the tests included in the unconstrained analysis were associated with multiple abilities. This could mean that certain correlations assessed more than just one ability and were therefore biasing the patterns found in my analysis. To resolve this issue, I excluded all correlations including tests that measured multiple abilities according to previous factor analytic literature (Benson et al., 2010; Bowden et al., 2004; Flanagan &

Mcgrew, 1998; Hoelzle et al., 2011; Tulsky & Price 2003) and reran the analyses. Patterns were nearly identical, with the exception that age was a significantly weaker predictor of speed with knowledge compared with speed with mental transformation. This supports the assumption that the patterns found in the unconstrained analyses were generally unbiased.

Overall, although the categorizations may not be perfect, the techniques used for classification were primarily objective and when reviewed by an expert were considered acceptable. In addition, after removing tests that assessed multiple abilities the patterns were nearly identical.

Additional Limitations

There was also some concern that a subset of the correlations influenced the results in the overall dataset. For example, if a single test had extremely strong relations between age and correlation size whereas the rest of the tests showed only weak relations, patterns could appear to be stronger in the overall model than was actually the case.

To explore this possibility, I separated the data into individual sets that included the correlations by age for each combination of tests. I then used age to predict the correlations in each of these subsets datasets; the estimates are presented in order of size in Figure 3. Although the individual estimates are not particularly trustworthy in themselves, I felt I could use them together to examine the relations between age and correlation size in my data and look for certain data points that were behaving differently from the overall sample. After a visual examination of the data, none of the effects appear to be driven by a subset of correlations. Based on these results, I conclude that it is unlikely that a minority of cases is responsible for the patterns found in the data.

Summary

In summary, evidence for age dedifferentiation was found, but was specific to certain types and small in magnitude. Consistent with the general slowing account of dedifferentiation, age effects on correlation size were only found when one of the tests included in the correlation measured speed. Increasingly strong dedifferentiation in older age groups was predicted but not found. This may represent the nature of dedifferentiation in health individuals, but could also be due to the limited number of independent samples.

Study 2: Longitudinal Dedifferentiation in Healthy and Pathological Adults

In the first study, I evaluated cognitive dedifferentiation with a number of large representative samples and a wide range of reliable cognitive abilities. Unfortunately, however, I could not explore within-person change or the role of pathology. Therefore, I conducted a second study where I evaluated the size of correlations between cognitive abilities in healthy and pathological individuals at baseline and over a three-year period.

The dataset used in the preparation of this article was obtained from the Alzheimer's Disease Neuroimaging Initiative[1] (ADNI) database (adni.loni.usc.edu). The ADNI was launched in 2003 as a public-private partnership, led by Principal Investigator Michael W. Weiner, MD. The primary goal of ADNI has been to test whether serial magnetic resonance imaging (MRI), positron emission tomography (PET), other biological markers, and clinical and neuropsychological assessment can be combined to measure the progression of mild cognitive impairment (MCI) and early Alzheimer's disease (AD).

Hypotheses

---

[1] Data used in preparation of this article were obtained from the Alzheimer's Disease Neuroimaging Initiative (ADNI) database (adni.loni.usc.edu). As such, the investigators within the ADNI contributed to the design and implementation of ADNI and/or provided data but did not participate in analysis or writing of this report. A complete listing of ADNI investigators can be found at: http://adni.loni.usc.edu/wp-content/uploads/how_to_apply/ADNI_Acknowledgement_List.pdf

H1. T*he presence of pathology will result in greater cognitive dedifferentiation at Timepoint 1 (T1) for AD and MCI individuals compared to healthy individuals of similar age and educational ability.* As described in greater detail above, there is evidence that time to death (Hulur et al., 2015; Wilson et al., 2012) and pathology (de Frias et al., 2007; Sliwinski et al., 2003) can be better predictors of dedifferentiation than age. Similarly, Siedlecki et al. (2008) found that there was a loss of fit for the model when metric invariance was imposed for healthy and impaired individuals. To test for pathological dedifferentiation, I will examine whether correlations between abilities are significantly higher for the individuals with large impairments than for individuals with less impairments. However, using the ADNI dataset with a slightly different model that that used here, Park et al. (2012) found the covariance between their ability factors could be achieved without loss of fit, suggesting that stability in the structure is also likely here.

H2. *Cognitive dedifferentiation will be shown in a period of 3 years, but only in pathological individuals.* I predicted that there would be no significant changes in correlation magnitude for healthy adults because age-related cognitive change in the healthy adults was unlikely to be found, particularly given the large number of time they took each of the tests during that time (Salthouse, 2014). This is supported by previous research; only two of the six longitudinal studies (Ghisletta & Lindenberger, 2003a; Ghisletta & de Ribaupierre, 2005) with intervals under ten years showed any evidence of dedifferentiation (Anstey et al., 2003; Ghisletta & Lindenberger, 2003a; Ghisletta & de Ribaupierre, 2005; Schaie et al., 1998; Tucker-Drob, 2008; Zelinski et al., 2003) and in both cases the evidence was weak.

However, I did predict that there would be changes in correlation magnitude in the impaired individuals. In three years, individuals with MCI decline twice as fast as healthy individuals and those with Alzheimer's decline four times as fast (Wilson et al., 2010). If dedifferentiation is tied to declining function as expected, it should be stronger in these populations. To test this hypothesis, I evaluated whether correlations are larger at T2 compared with T1 for the more impaired and less impaired group, with the expectation that evidence of dedifferentiation over time was more likely to appear for the more impaired versus less impaired group.

Method

Participants

The ADNI study included 229 cognitive normal individuals, 395 participants with MCI, and 190 participants with Alzheimer's Disease between 55-90 who completed a battery of cognitive tests at T1.

Healthy normal participants needed to have Mini-Mental State Examination scores between 24-30 and clinical dementia ratings of 0, non-depressed, non-MCI, and non-demented. Participants in the MCI condition were required to have MMSE scores of 24-30, a memory complaint, objective memory loss measured by a normed score on the WMS logical memory II, a CDR of .5, and an absence of significant impairments in other domains, preserved ADLs, and an absence of dementia. Finally, the participants with Alzheimer's Disease were required to have MMSE scores of 20-26, a global CDR of .5-1, and to meet the NINCDS/ADRDA criteria for probable AD (adni.loni.usc.edu).

Cognitive Tests

Fourteen cognitive tests from the ADNI1 dataset were used in the analyses. Information on these tests can be found at the following paper (Park et al., 2012). Descriptions of the tests are listed in Supplementary Table 12. The MMSE was used as a screening measure and was not included in any of the four ability factors included in this analysis. Speed was measured using digit symbol substitution, the ADAS number cancellation test, and Trails A. Knowledge was measured using category fluency (animals & vegetables), the Boston naming test, and the ADAS naming test. These are not ideal tests for assessing general knowledge because according to recent CHC standards fluency tests are more consistent with measures of long-term memory than knowledge (Schneider & McGrew, 2012). However, they are the closest tests available to the construct in this battery, and therefore I chose to use them and simply interpret my results carefully with regard to this factor. Memory was measured using the auditory verbal learning test (trials 1-5, short delay, long delay, recognition), the ADAS delayed recall test, and the ADAS recognition test. Finally, Mental Transformation was measured using digit span backward, clock drawing, and the ADAS construction test.

Analytic Plan

Confirmatory factor analyses were conducted in AMOS (Arbuckle, 2010) and maximum likelihood was used to deal with missing data. Two models were used to evaluate cross-sectional and longitudinal dedifferentiation.

For the cross-sectional analyses, I tested whether the presence of pathological conditions result in higher correlations at T1. Model 1, presented in Figure 4, is a hierarchical model in which the correlations between the cognitive tests at T1 are explained by first-order latent factors presumed to represent Mental Transformation, Memory, Speed, & Knowledge. These factors are correlated with one another in each impairment group and these correlations are the focus of my first analysis.

In order to make the comparison of these correlations reasonable, I tested the fit of a metric invariance version of Model 1, which forces the loadings between the observed tests and latent factors to be the same to determine if there was a loss of fit. Using this approach ensures the meaning of latent factors is the same across conditions. I evaluated the fit of this model using absolute measures (CFI, RMSEA) and by comparing it to a configural model where the restrictions of the metric model are relaxed I did not constrain the covariances between the factors to be the same. I will use the metric invariance model if there is not a significant loss of fit, or if the absolute fit is still excellent (although this means that that the loadings may not be equivalent).

I first attempted to fit a model where the loadings were constrained across the normal, MCI, and AD diagnostic conditions. However, this model failed to converge and therefore was not usable. Therefore, I abandoned this option and instead divided the sample into groups based on sum-of-boxes CDR scores. The sum-of-boxes CDR scores is a total of the ratings (0-3) on six measures of cognitive and daily functioning: memory, orientation, judgment and problem solving, community affairs, home and hobbies, and personal care. This method was used in a previous analysis of the ADNI neuropsychological battery (Park et al., 2012).

Although the full range for the sum CDR scores is 0-18, the median score in this sample was 1.5. To get roughly equal samples, individuals below 1.5 were categorized as less functionally impaired, and those with a score of 1.5 or above were categorized as more functionally impaired. Mean age and proportion female were similar across conditions, see Table 8, but there were a far greater percentage of individuals who were cognitively normal in the less impaired group and a far greater percentage of individuals who have Alzheimer’s Disease in the more impaired group.

As described in greater detail below, the metric invariance comparing CDR groups also provided a poorer fit than the configural model. Looking at the loadings in the two groups, the differences appear to stem from higher loadings on the speed and general knowledge tests and lower loadings for the memory tests in the more impaired groups. However, the absolute fit measures were still excellent and so I decided to use these groups keeping in mind that this model is not the most accurate representations of the true patterns across groups in my data.[2]

I performed z-tests comparing the correlations between each of the impairment groups. If the correlations were significantly higher in the more impaired (CDR > 1.5) versus less impaired conditions (CDR < 1.5), this would suggest that pathological dedifferentiation was occurring. The primary value of these analyses was to determine if whether my modified model based more closely on CHC constructs would be consistent with the Park et al. (2012) study that found no reduction of fit when constraining the relations between cognitive ability factors.

For the longitudinal analyses, I tested whether the progression of pathological conditions result in higher correlations at T2 versus T1 for the less impaired and more impaired conditions individually. Model 2 (Figure 5) is composed of two hierarchical organizations including first-order latent factors presumed to represent Mental Transformation, Memory, Speed, & Knowledge. These factors are correlated with one another at each time point for each impairment group and these correlations are the focus of my first analysis.

As with the baseline model, I tested for metric invariance over time for each of the impairment conditions. To do this, I fit a version of Model 2 where the loadings from each factor to its latent factor were the same at T1 and T2. I evaluated this model by looking at absolute fit statistics and comparing it to a configural invariance model where these strictures were not imposed. As described in greater detail in the results, the fit of this model was significantly

[2] Similar patterns were found when groups were compared using a median split of MMSE scores, but due to the use of the CDR measure in a previous CFA of the ADNI database, I chose to use the CDR scores.

poorer but still excellent. Therefore, I moved forward with the metric invariance model while keeping in mind that it was not the ideal fit for the data and some patterns may be obscured.

I then compared whether forcing the covariances between the latent factors at T1 to be the same at T2 reduced model fit. This would suggest that relations between cognitive abilities vary across time. If model fit was reduced, I released each of the covariances individually and examined which improved model fit. If a covariance that was allowed to vary improved model fit and the correlation was higher at T2 than T1, this would provide evidence of longitudinal dedifferentiation. I hypothesized that this pattern would be more evident for more impaired versus less impaired condition if pathological dedifferentiation was taking place. This approach is admittedly limited because it only indirectly compares changes between the groups. However, given the difficulty of comparing these groups directly and my expectation that dedifferentiation should be weaker and therefore not found in the less impaired groups, it was still sufficient.

Results

Baseline Model

Correlation matrices for the two impairment conditions at T1 are presented in Supplementary Tables 13 and 14. Tests where higher scores represented poorer cognitive performance were multiplied by -1 so that for all tests higher scores signified higher ability. In all conditions, the tests hypothesized to represent the four cognitive abilities were generally positively correlated to one another, although the strength of these correlations varied between conditions.

I first compared a model with metric invariance, this means that the parameters connecting the observed tests to the latent factors had to be the same across conditions, with one where these parameters were allowed to vary. It is important to use a metric invariance model because it ensures that the latent factors have the same qualitative meaning and are comparable across conditions. Although the metric model had a significantly poorer fit than the configural model where these parameters were free to vary, ($\chi^2 = 146.69$, df = 12, $p < .001$), it still provided an excellent overall fit. Absolute fit statistics, unstandardized loadings, and standardized loadings are in Table 9. Given that metric invariance model still provided an excellent fit I chose to use, while taking into consideration that it was not the best way to model the data.

The standardized loadings from each latent construct and its respective observed variables were moderate to large across groups, providing evidence of convergent validity.

The correlations were z-transformed and compared across impairment conditions. I predicted that if dedifferentiation occurs with pathology, correlations should be significantly higher in the more impaired condition compared with the less impaired condition. For speed with mental transformation, correlations were higher for the more impaired condition than the less impaired ($z = 6.82$, $p < .0001$) condition. For memory with mental transformation, correlations were lower for the more impaired versus less impaired ($z = 4.28$, $p < .0001$) condition. There were no other significant differences. Correlations between abilities are presented in Table 10.

Overall, one of the correlations was larger in the more impaired versus less impaired condition, consistent with pathological dedifferentiation. However, one of the other correlations showed the opposite pattern and the others were stable. Therefore, there is little evidence of pathological dedifferentiation. These results are consistent with those found in Park et al. (2012) who found that the covariances between groups could be constrained without a loss in fit and demonstrates that altering their model to fit more closely with CHC constructs did not significantly impact the overall patterns.

Longitudinal Models

Correlation matrices for the two impairment conditions at T2 are presented in Supplementary Tables 15 and 16. Tests where higher scores represented poorer cognitive performance were multiplied by -1 so that for all tests higher scores signified higher ability. Consistent with T1, the tests hypothesized to represent the four cognitive abilities were generally positively correlated to one another for both conditions. Correlations between tests at T1 and T2 (presented in Table 11) were moderate to high, suggesting both stability and some change over time.

To explore whether dedifferentiation occurred over time, I evaluated whether allowing correlations between latent factors to vary across the three-year period improved model fit. I conducted separate analyses for the more impaired versus less impaired groups.

I first tested for each impairment group whether the loadings between each test and its respective latent construct could be constrained to be equal at T1 and T2. As described in greater detail below, although absolute fit for the 'just loadings constrained' models provided excellent fits in both impairment groups, both models performed significantly worse than models without these constraints. As with the baseline model, I chose to continue my analyses using the constrained model while keeping in mind that it did not provide the best fit for the data.

In the next step, I evaluated whether the covariances between the latent constructs could also be constrained to be equal at T1 and T2. If these loadings and covariances constrained models had significantly poorer fits compared with the just loadings constrained model, it would mean that there were significant differences in at least one of the correlations across conditions.

I then tested whether model fit would be improved by freeing each covariance individually. If freeing a covariance between latent structures improved fit, and the correlation was higher at T2 than T1, that would suggest increased dedifferentiation over time. I expected to find greater evidence of dedifferentiation in the more impaired versus less impaired condition as the progression of pathology should have a greater impact on the more impaired condition.

Less Impaired Group

Although both the loadings constrained model (χ^2 = 726.48, df = 418, CFI = .95, RMSEA = .044) and the unconstrained model (χ^2 = 691.38, df = 406, CFI = .95, RMSEA = .043) provided excellent fits for the data, the loadings constrained model was a poorer fit to the data ($\Delta\chi^2$ = 35.10, $p < .001$). As with the baseline model, I chose to move forward using this model with the awareness that it is not the best possible fit for the data and therefore some part of the true nature of these data may be lost.

Constraining the covariances in addition to the loadings provided an excellent fit for the data in the less impaired condition (χ^2 = 790.82, df = 424, CFI = .94, RMSEA = .047). However, the loadings and covariances constrained model provided a significantly poorer fit for the data than the loadings constrained model ($\Delta\chi^2$ = 63.07, $p < .001$). This means that at least one of the relations between cognitive abilities differs significantly between T1 and T2. The standardized loadings from each latent construct and its respective observed variables were moderate to large, providing evidence of convergent validity. All the cognitive constructs for T1 and T2 were significantly correlated with one another. Absolute fit statistics, unstandardized loadings, and standardized loadings for both these models are in Tables 12 & 13.

I then evaluated whether freeing each additional parameter improved fit. Allowing the covariance between the memory and speed factors to vary with time resulted in a significant improvement in fit over the loadings and covariances constrained model, ($\Delta\chi^2$ = 6.83, p = .009). The correlation between speed and memory was higher at T2 (r = .55) than T1 (r = .50).

Allowing the covariance between the memory and knowledge factors to vary with time resulted in a significant improvement in fit over the loadings and covariances constrained model, ($\Delta\chi^2$ = 4.25, p = .039). The correlation between memory and knowledge was higher at T2 (r = .74) than T1 (r = .68). Allowing the covariance between the mental transformation and speed factors to vary with time resulted in a significant improvement in fit over the loadings and covariances constrained model, ($\Delta\chi^2$ = 5.71, p = .017). The correlation between mental transformation and speed was higher at T2 (r = .70) than T1 (r = .69). Releasing the other covariances did not significantly improve model fit (p's > .10). In sum, it appears that some of the correlations were larger at T2 than at T1, consistent with dedifferentiation in the less impaired group.

More Impaired Group

Although both the loadings constrained model (χ^2 = 715.62, df = 418, CFI = .94, RMSEA = .041) and the unconstrained model (χ^2 = 647.97, df = 406, CFI = .95, RMSEA = .037) provided excellent fits for the data, the loadings constrained model was a poorer fit to the data ($\Delta\chi^2$ = 67.65, p < .001). As with the baseline model, I chose to move forward using this model with the awareness that it is not the best possible fit for the data and therefore some part of the true nature of these data may be lost.

Constraining the covariances in addition to the loadings provided an excellent fit for the data in the less impaired condition (χ^2 = 777.33, df = 424, CFI = .93, RMSEA = .044). However, the loadings and covariances constrained model provided a significantly poorer fit for the data than the loadings constrained model ($\Delta\chi^2$ = 61.71, p < .001). This means that at least one of the relations between cognitive abilities differs significantly between T1 and T2. The standardized loadings from each latent construct and its respective observed variables were moderate to large, providing evidence of convergent validity. All the cognitive constructs for T1 and T2 were significantly correlated with one another. Absolute fit statistics, unstandardized loadings, and standardized loadings for both these models are in Tables 12 and 13.

I then evaluated whether freeing each covariance improved fit. Allowing the covariance between the mental transformation and speed factors to vary with time resulted in a significant improvement in fit over the loadings & covariances constrained model, ($\Delta\chi^2 = 5.61$, $p = .018$). The correlation between speed and mental transformation was higher at T2 ($r = .93$) than T1 ($r = .85$), consistent with dedifferentiation. However, it is important to note that although there was a significant difference between these models, the model fit for the model where the memory-speed covariance was free to vary was extremely similar to that of the loadings and covariances constrained model. Releasing the other covariances individually did not significantly improve model fit (p's > .07).

In summary, and contrary to expectations, the less impaired condition actually showed more evidence of dedifferentiation over time compared with the more impaired condition. In addition, although there were significant differences in model fit, the increases in correlation size over time tended to be very small and did not notably impact model fit.

Discussion

The goal of the second study was to evaluate whether pathological dedifferentiation would be found when comparing individuals with less or more impairment at one time point and over a three-year time period. In contrast to previous research that showed differences between pathological and healthy individuals, I found only weak evidence of pathological dedifferentiation when comparing correlations between less and more severe levels of impairment. Dedifferentiation was shown across the time in both impairment conditions, although there were more significant differences in the less impaired condition. Overall, the present study provides somewhat inconsistent evidence of pathological dedifferentiation.

Cross-sectional Pathological Dedifferentiation

Previous work has found increased dedifferentiation in pathological conditions and individuals in terminal decline (de Frias et al., 2009; Batterham et al., 2011; Wilson et al., 2012, Sliwinski et al., 2003). They attributed the increase in correlation size to the gross changes in cognitive performance (Laukka et al., 2006; Lovden et al., 2005; Sliwinski et al., 2003; Wilson et al., 2010) and in the brain (Schneider et al., 2009; Uylings & de Brabander, 2002) that characterize pathological and terminal decline. Consistent with these hypotheses, pathology and terminal decline have been shown to be stronger predictors of dedifferentiation than age (Batterham et al., 2011; Wilson et al., 2012, Sliwinski et al, 2003). In the present cross-sectional analysis, however, only one of the correlations, speed with mental transformation, was larger in the more impaired versus less impaired condition as would be consistent with dedifferentiation. The other correlations were stable with the exception of memory with mental transformation, which was higher in the less impaired versus more impaired condition.

These cross-sectional results suggest stability rather than dedifferentiation even when pathology is present. However, it is possible that the methodological limitations may have occluded real group differences. In the present study, the failure to achieve metric invariance for any model across the AD, MCI, and normal conditions meant that these groups could not be used for comparisons. This lead to the necessity of comparing conditions on CDR Sum-of-Boxes scores and in the inclusion of MCI individuals in both the less impaired and more impaired conditions. This reduced the differences between the impairment conditions used in this analysis, which in turn may have obscured dedifferentiation between the healthy and pathological individuals.

This explanation is supported by the results of a similar study de Frias et al. (2009) where the comparison groups did not overlap as in the current analysis. They found that a multidimensional model of executive functioning provided the best fit for cognitively elite and normal older adults whereas a simpler, unidimensional model provided the best fit for cognitively impaired older adults. This suggests that the specificity of cognitive abilities decreases in those with cognitive impairments, consistent with pathological dedifferentiation. Both the de Frias et al. (2009) study and the present study had a cross-sectional element, both had multiple tests for their cognitive constructs, and both used structural equation modelling. The major differences were that de Frias et al. (2009)'s groups were defined by performance on tests (CE above the mean in all five tests; CN: scores on all five between -1.5 and 1.5 standardized deviations, CI, at least one test more than 1.5 SDs below the relevant group mean) and the abilities assessed by the cognitive tests.

In summary, the evidence of pathological dedifferentiation when comparing the less versus more impaired groups at T1 was inconsistent and more in favor of stability. This is the similar to the results found in Park et al., 2012, which means that altering their model to more closely align with CHC abilities did not affect the overall patterns. This could mean that dedifferentiation does not occur in the presence of pathology. However, evidence for dedifferentiation may have been obscured by the inclusion of individuals with MCI in both the less and more impaired conditions.

Longitudinal Pathological Dedifferentiation

In the longitudinal analysis, evidence of dedifferentiation was found for both the more and less impaired conditions. Interestingly, dedifferentiation was actually more common in the less impaired than more impaired group. For the less impaired condition, allowing three of the six correlations (memory with knowledge, memory with speed, mental transformation with speed) to vary at T1 and T2 improved model fits. When released, these correlations were uniformly higher at T2 than T1, consistent with dedifferentiation. In contrast, only one of the six correlations (mental transformation with speed) improved model fit when allowed to vary from T1 to T2. For both conditions, the changes in model fit were minimal, suggesting that pathological dedifferentiation is a small effect.

Given previous research establishing model invariance over time (Salthouse, 2012) in healthy individuals and the inclusion of pathological individuals in both impairment groups, it is likely that the patterns found in both conditions are examples of pathological dedifferentiation. However, it was still somewhat surprising that patterns were weaker, not stronger, in the more impaired group. The interpretation of these results appears to be due to the grouping issues discussed above and the differential rates of attrition across the diagnostic groups. Only 11 of the AD patients included at T1 remained at T2 and the more impaired condition had only 157 participants as compared to 430 participants at T1. The lower N in the more impaired group reduced the power to detect change, and the additional loss of distinctiveness between the conditions made it less likely to identify between-group differences in patterns.

In summary, evidence of longitudinal dedifferentiation that is likely due to pathology was found over a three-year time period. Contrary to expectations, these patterns were stronger in the less impaired versus more impaired condition, but these results are likely due to the low distinctiveness between groups and the small N at T2 in the more impaired condition.

Types of Cognitive Abilities

Although the evidence for dedifferentiation was inconsistent, it is interesting to note that several of the patterns identified in Study 1 were mirrored in Study 2. Specifically, speed with mental transformation correlations were significantly higher in the more impaired versus less impaired groups at baseline and higher at T2 than T1 for both impairment groups. In addition, the speed with memory correlations showed evidence of dedifferentiation for the less impaired group. These results add some support to the hypothesis of a role for speed in dedifferentiation, and reasserts the importance of evaluating specific combinations of cognitive abilities. The additional findings of dedifferentiation over time for the less impaired group for memory with knowledge and an absence of dedifferentiation for speed with knowledge suggest a difference in the mechanisms underlying age and pathological dedifferentiation.

Limitations

The most notable limitation in Study 2 was the inability to achieve metric invariance without losing fit in the data. Although the overall fit was still excellent after forcing the loadings to be the same across impairment conditions or over time the loss of fit suggests that there are patterns in the data that were not represented. However, in order to make the assumption that the meaning of my factors was comparable across groups and time, I had to constrain the factor loadings despite a loss in fit.

Additionally, the tests included in the neuropsychological battery for the ADNI were fairly limited and did not always provide the purest representations of the cognitive abilities. Therefore, it is possible that the pattern of results presented above would not be replicated with alternative tests that better represented the CHC constructs of interest and had higher loadings.

Another issue is that in the longitudinal groups I did not directly compare the change over time in correlation magnitude between the impairment groups. However, this does not appear to be a major issue because fewer correlations were greater in T2 than T1 in the more impaired group, which provides acceptable evidence that dedifferentiation was not be stronger for the more impaired compared with the less impaired group as I predicted.

Finally, as noted above, the use of impairment groups as opposed to diagnostic groups and the greater attrition rates over time in the more impaired group meant that there was a fair amount of overlap in pathology. Therefore, true differences between these groups may have been obscured both at baseline and over time.

Summary

In summary, evidence for pathological dedifferentiation was inconsistent cross-sectionally. There was evidence of a small amount of longitudinal dedifferentiation in both the impairment categories, although the patterns were stronger in the less versus more impaired condition. These patterns are likely due to the presence of MCI individuals in both groups and the limited N in the more impaired group at T2. Overall, results suggest that a small amount of dedifferentiation occurs for individuals with progressing pathology.

General Discussion

Previous research on age dedifferentiation has been extremely inconsistent. Various studies have found evidence of dedifferentiation (Cunningham & Birren, 1980; Deary et al., 2004; de Frias et al., 2007; Ghisletta & Lindenberger, 2003; Hertzog et al., 2003; Hulur et al., 2015; Adrover-Roig et al., 2012; Babcock et al., 1997; Balsamo & Romanelli, 2010; Cunningham & Birren 1980; de Frias et al., 2007; Hedden et al., 2006; Hertzog, 1989; Hertzog et al., 2001; Li et al., 2004; Lindenberger & Baltes, 1997; Nyberg et al., 2003; Salthouse &

Saklofske, 2010; Schaie et al., 1989; Schultz et al., 1980) whereas other studies have supported stability across the adult lifespan (Anstey et al., 2003; Batterham et al., 2011; Finkel et al., 2007; Schaie et al., 1998; Tucker-Drob, 2009; Zelinski & Stewart, 1998; Zelinski et al., 2003, Juan-Espinosa et al., 2003; Hale et al., 2011; Hildebrandt et al., 2011; Hull et al., 2011; Johnson et al., 2010; Juan-Espinosa et al., 2000; Park et al., 2002; Sims et al., 2009; Singer et al., 2003; Vaughn & Giovanello, 2010).

These inconsistencies could be due to a variety of factors. In this project, I specifically examined three major issues: whether the types of cognitive abilities used to test for dedifferentiation impact the results, whether dedifferentiation is linear or non-linear across the adult lifespan, and the role of pathology in age dedifferentiation. Study 1 showed evidence of linear age dedifferentiation in normal individuals when speed was directly assessed as part of the correlation, consistent with the general slowing account of dedifferentiation (Li et al., 2004). This demonstrates that dedifferentiation does not occur after a threshold of decline has been reached or due to pathological dedifferentiation (Batterham et al., 2011; de Frias et al., 2009; Li et al., 2004). However, these patterns were extremely small, which may explain why many previous studies failed to support age dedifferentiation. Inconsistent evidence for dedifferentiation in pathological individuals was found at baseline. Comparisons across time showed small increases in correlation size for certain correlations.

In summary, dedifferentiation does appear to occur with age and when pathology is present under certain condition. However, its influence is weak and are unlikely to confound quantitative comparisons between groups.

Types of Cognitive Abilities

The first goal was to explore whether dedifferentiation was limited to certain cognitive abilities. Specifically, I expected that dedifferentiation would be strongest for speed with memory and mental transformation. This is based on Li et al. (2004)'s argument that dedifferentiation occurs when biological restraints on processing speed limit the expression of other cognitive abilities. This explanation is consistent with the results of previous studies that included a test of processing speed (Babcock et al., 1997; Ghisletta et al., 2003; Ghisletta et al., 2005; Hertzog et al., 2001).

I hypothesized that dedifferentiation would be found for correlations between speed and knowledge as in previous studies (Ghisletta et al., 2003; Ghisletta et al., 2005) but that these patterns would be weaker because knowledge abilities are more associated more with life experiences like education or occupation than other abilities (Ghisletta & Lindenberger, 2003; Ghisletta & de Ribaupierre, 2005).

I predicted that dedifferentiation would be found for memory with mental transformation correlations, given that both these abilities are dependent on speed. However, as with speed and knowledge, I expected that patterns could be weaker since speed was not directly.

Finally, I predicted that dedifferentiation would not be found for knowledge with memory and knowledge with mental transformation because speed was not directly assessed and because knowledge was less likely to be impacted by speed.

The first study explored whether age predicted correlation size using a set of large cross-sectional samples drawn from the technical manuals of various cognitive batteries. Benefits of using these data were that each of the cognitive abilities of interest were assessed with a variety of different tests, these tests were reliable and valid, there were a range of different ages available for testing, and the fairly large samples were healthy and representative of the general population at the time of data collection.

Consistent with my hypotheses, evidence in favor of dedifferentiation was found for correlations that directly assessed speed (speed with mental transformation; speed with knowledge) although the effects were small (Δ r's between 25 & 75 < .11). Age was not a significant predictor for speed with memory, but the sample for this combination was extremely small and therefore any results may not be found for a larger sample. These findings are consistent with the general slowing account of dedifferentiation suggested in Li et al. (2004). As expected, age was a weaker predictor of correlation size for speed with knowledge compared to speed with mental transformation. Finally, age was also not a predictor for knowledge with memory and knowledge with mental transformation.

One surprising finding was that age was not a significant predictor of the mental transformation with memory correlations as predicted. However this is explainable given that the amount of dedifferentiation was extremely small even when speed was directly measured as part of the correlation.

The second study explored whether correlations would be higher in individuals with more versus less cognitive impairments at baseline and over time. The benefit of this study was that it includes multiple time points and individuals with pathological groups. Results at baseline between the impairment groups were inconsistent; one correlation was higher and one correlation was lower in the more impaired groups. Over time, correlations between speed and memory, memory and knowledge, and mental transformation and speed factors were higher in the less impaired groups. Correlations between mental transformation and speed were also higher in the more impaired groups. As with Study 1, these patterns are mainly consistent with the general slowing account of dedifferentiation, although there were some differences that might be specific to pathology.

In summary, evidence of dedifferentiation across both studies was primarily limited to correlations that directly assessed speed. This is consistent with the hypothesis that cognitive abilities become more closely connected to one another lower levels of speed of processing constrains abilities in other types of cognitive performance. In all cases where it was identified, dedifferentiation was extremely small. This may explain why patterns were not found for correlations between tests that are heavily impacted by speed but are not direct assessments of it.

Trajectory of Dedifferentiation

The second and third goals were to explore the trajectory of dedifferentiation across the adult lifespan and the role of pathology. Several previous studies have found that dedifferentiation occurs only after a certain age point and proposed that a certain amount of cognitive decline must have taken place (de Frias et al., 2007; Li et al., 2004; Schaie, 1989; Ward, 2000). An alternative explanation for non-linear dedifferentiation is that 'age' dedifferentiation only occurs because samples of older adults are likely to have a larger number of individuals with pathological conditions. These conditions lead to major deterioration in both one's cognitive abilities and the brain that impact the way cognitive performance is expressed and result in dedifferentiation (Batterham et al., 2011). This hypothesis has been supported by findings of dedifferentiation when comparing healthy and normal samples (de Frias et al., 2009) and studies where terminal decline/pathology is a stronger predictor of dedifferentiation than age (Batterham et al., 2011; Sliwinski et al., 2003; Wilson et al., 2012).

I explored non-linear dedifferentiation in both studies. In Study 1, I evaluated whether the trajectory of dedifferentiation in a healthy adult lifespan sample. If age dedifferentiation is a linear process like cognitive decline (Salthouse, 2010), I expected that age would and age^2 would not significantly predict correlation size. If dedifferentiation is non-linear and a certain threshold of decline is needed, I expected age to be a stronger prediction of correlation size in older adults versus younger adults. Finally, if pathology was a primary cause of age dedifferentiation, I expected no dedifferentiation or weak non-linear dedifferentiation because the participants in many of the samples included (WAIS, 1955; 1981; 1997; 2008b; WASI, 1999; WMS, 1987; 1997; 2009), although not all (Woodcock & Johnson, 1989), were screened for pathological conditions.

In contrast to previous reports, when significant dedifferentiation occurred it was always linear. This was determined based on the lack of a significant age^2 predictor. These results suggest that dedifferentiation does not occur only after a certain amount of decline has occurred. In addition, these findings demonstrate that age dedifferentiation is not just an epiphenomenon of pathological decline as it was found in a primarily healthy sample and was not limited to the older age groups.

In Study 2, I wanted to determine whether pathological impairment was associated with correlations at baseline and over time. If pathology results in dedifferentiation, I expected higher correlations between abilities at baseline for the more impaired group. I also expected more evidence of dedifferentiation in the more impaired group as their more severe pathological conditions progressed.

At baseline, there was more evidence for stability than dedifferentiation or differentiation. Only one correlation was significantly higher for the more impaired versus less impaired groups; four of the other correlations were not significantly different from one another and the final correlation was significantly lower in the more impaired versus less impaired groups. These results do not suggest that increased pathology results in dedifferentiation, although they could be due to the inclusion of pathological individuals in both groups.

Over time, there was evidence for dedifferentiation. However, it was primarily for the less impaired versus more impaired groups. In the less impaired groups, three of six correlations increased in size. In the more impaired groups, only one of the six increased in size. This may not be surprising, given the large attrition, particularly for the participants with Alzheimer's, seen in the more impaired group (only 36% of the more impaired sample remained compared with 76% of the less impaired sample) may have reduced my power to detect significant differences. It seems likely that the differences in both groups are primarily due to pathological change, as cognitively normal individuals tend to show little change in performance over a period as short as three years with a large number of assessments (Salthouse, 2014) and only two out of six studies focusing on healthy adults showed evidence of dedifferentiation under ten years. This supports suggestions that the inclusion of pathological individuals could lead to the appearance of dedifferentiation in otherwise normal individuals (Batterham et al., 2011). In both groups, these increases were small (Δ r's < .08).

In summary, although previous work suggests that age dedifferentiation is non-linear, age^2 was not a significant predictor of correlation size for any of the correlations that exhibited dedifferentiation in the healthy cross-sectional sample. This suggests that at least in normally aging individuals, dedifferentiation does not occur only after a certain threshold of decline has been reached and is not just an epiphenomenon of pathological dedifferentiation. In the second study, evidence of pathological dedifferentiation was inconsistent at baseline. However, small differences in correlation size were found over time and are likely due to the inclusion of pathological individuals, consistent with previous reports (Batterham et al., 2011; de Frias et al., 2007; Li et al., 2004; Schaie, 1989).

Implications for Cognitive Aging Research

One of the major reasons for exploring age dedifferentiation was to determine whether it changed the meaning of the cognitive abilities across the adult lifespan. If cognitive abilities share more variance in older age groups/over time this would suggest that the specific abilities are less or more reflective of general intelligence. If these differences are large, it could mean that quantitative comparisons between young and old adults or over time are confounded.

However, in both the normal and pathological samples, the magnitude of dedifferentiation was small. In addition, these patterns were generally only found when speed was included in the correlation. Based on this evidence, I believe that although dedifferentiation does occur it is too specific and weak to seriously impact cognitive structure and confound quantitative comparisons.

Limitations

In both studies, the tests used to assess the cognitive abilities of interest had their limitations. In Study 1, not all tests were designed with the CHC theory in mind and I had to use previous factor analytic work to determine the abilities they likely represented. In Study 2, several tests had lower than optimal loadings on their factors of interest. In addition, the tests used to represent General Knowledge are more accurately tests of long-term memory in the CHC (Schneider & McGrew, 2012) and therefore the patterns for this group may not be replicated for other tests of General Knowledge.

There were also limitations with the number of participants or samples in each study. For Study 1, there were fewer groups in the older age ranges than in the middle and younger age ranges. In addition, the number of correlations assessing speed and memory were extremely low, meaning that it was impossible to reliably determine whether dedifferentiation occurred in this group. For Study 2, there was a large amount of attrition between T1 and T2, particularly for the participants with Alzheimer's in the more impaired group. This meant that power was limited in this group.

Previous work on dedifferentiation has been impacted by differences in the reliability of the measures, the variability of the scores, and the analytical methods used could produce inconsistencies. Although I did not directly test whether these methodological issues would be more likely to produce differentiation/dedifferentiation, I used standardized tests with documented reliability and variability. Therefore, I reduced the likelihood of these issues occuring.

That being said, I did take this issues into consideration when designing my research approach. For Study 1, the use of normed tests meant that I could ensure that all tests included had acceptable reliabilities and similar variabilities in the different age groups. I used a meta-analytic method that allowed me to control for the dependencies between data and to test moderators. For Study 2, all the tests included were well-established and had been evaluated in previous work (see Park et al. 2012 for details). However, constraining the loadings from factors to tests resulting in a loss fit; this means that there could have been differences between age groups and times in the tests' reliability.

Conclusion

This project was designed to explore the inconsistencies in previous research regarding age dedifferentiation. Study 1 revealed that although age dedifferentiation does occur, patterns are weak and limited to correlations that included direct assessments of speed. In addition, patterns were primarily linear; this suggests that dedifferentiation does not occur after a certain threshold of decline or is primarily due to the presence of higher number of pathological individuals in older age. There was no consistent evidence of cross-sectional pathological dedifferentiation and although correlations increased over time, the increases were small and primarily found for the less impaired individuals. Overall, it appears that although dedifferentiation does occur, it is unlikely to impact quantitative comparisons.

References

Abad, F. J., Colom, R., Juan-Espinosa, M., Garcia, L. F. (2003). Intelligence differentiation in adult samples. *Intelligence, 31,* 157-166.

Adrover-Roig, D., Sese, A., Barcelo, F. & Palmer, A. (2012). A latent variable approach to executive control in healthy aging. *Brain and Cognition, 78,* 284-299.

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). Cross-sectional and longitudinal patterns of dedifferentiation in late-life cognitive and sensory function: The effect of age, ability, attrition, and occasion of measurement. *Journal of Experimental Psychology: General, 132,* 470-487.

Arbuckle, J. L. (2010). Amos (Version 19.0) [Computer Program]. Chicago: SPSS.

Babcock, R. L., Laguna, K. D., Roesch, S. C. (1997). A comparison of the factor structure of processing speed for younger and older adults: Testing the assumption of measurement equivalence across age groups. *Psychology & Aging, 12,* 268-276.

Balsamo, M. & Romanelli, R. (2010). The de-differentiation hypothesis in normal elderly persons. *Perceptual and Motor Skills, 110,* 85-88.

Baltes, P. B., Cornelius, S. W., Spiro, A., Nesselroade, J. R., & Willis, S. L. (1980). Integration versus differentiation of fluid/crystallized intelligence in old age. *Developmental Psychology, 16,* 625-635.

Batterham, P. J., Christensen, H., & Mackinnon, A. J. (2011). Comparison of age and time-to-death in the dedifferentiation of late-life cognitive abilities. *Psychology & Aging, 26,* 844-851.

Benson, N., Hulac, D. M., Kranzler, J. H. (2010). Independent examination of the

Wechsler Adult Inelligence Scale – Fourth Education (WAIS-IV): What does the WAIS-IV measure? *Psychological Assessment, 22,* 121-130.

Bowden, S. C., Cook, M. J., Bardenhagen, F. J., Shores, E. A., Carstairs, J. R. (2004). Measurement invariance of core cognitive abilities in heterogeneous neurological and community samples. *Intelligence, 32,* 363-389.

Bowden, S. C., Weiss, L. G., Holdnack, J. A., & Lloyd, D. (2006). *Psychological Assessment, 18,* 334-339.

Burton, D. B., Ryan, J. J., Paolo, A. M., & Mittenberg, W. (1994). Structural equation analysis of the Wechsler Adult Intelligence Scale- Revised in a normal elderly sample. *Psychological Assessment, 6,* 380-385.

Canivez, G. L., Konold, T. R., Collins, J. M., Wilson, G. (2009). Construct validity of the Wechsler Abbreviated Scale of Intelligence and Wide Range Intelligence Test: Convergent and structural validity. *School Psychology Quarterly, 24,* 252-265.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor analytic studies.* Cambridge, England: Cambridge University Press.

Carroll, J. B. (2003). The higher-stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors In H. Nybord (Ed.), The scientific study of general intelligence: Tribute to Arthur R. Jensen (pp. 5-21). Boston: Pergamon.

Cohen, J. (1957). The factorial structure of the WAIS between early adulthood and old age. *Journal of Consulting Psychology, 21,* 283-290.

Cunningham, W. R. & Birren, J. E. (1980). Age changes in the factor structure of intellectual abilities in adulthood and old age. *Educational and Psychological Measurement, 40,* 271-290.

de Frias, C. M., Dixon, R. A., & Strauss, E. (2006). Structure of four executive functioning tests in healthy older adults. *Neuropsychology, 20,* 206-214.

de Frias, C. M., Dixon, R. A., & Strauss, E. (2009). Characterizing executive functioning in older special populations from cognitively elite to cognitively impaired. *Neuropsychology, 23,* 778-791.

de Frias, C. M., Lovden, M., Lindenberger, U., & Nilsson, L.-G. (2007). Revisiting the dedifferentiation hypothesis with longitudinal multi-cohort data. *Intelligence, 35,* 381-392.

Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86,* 130-147.

Escorial, S., Juan-Espinosa, M., Garcia, L. F., Rebollo, I., & Colom, R. (2003). Does *g* variance change in adulthood? Testing the age de-differentiation hypothesis across sex. *Personality and Individual Differences, 34,* 1525-1532.

Finkel, D., Reynolds, C. A., McArdle, J. J., & Pederson, N. L. (2007). Age changes in processing speed as a leading indicator of cognitive aging. *Psychology & Aging, 22,* 558-568.

Flanagan, D. P. & McGrew, K. S. (1998). Interpreting intelligence tests from contemporary Gf-Gc theory: Joint confirmatory factor analysis of the WJ-R and KAIT in a non-white sample. *Jounral of School Psychology, 36,* 151-182.

Ortiz, S. O., Flanagan, D. P., & Alfonso, V. C. (2013). Essentials of cross-battery assessment (3rd ed.). New York: Wiley.

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 Psychiatry Research. 12,* 189-198.

Fox, M. C., Berry, J. M. & Freeman, S. P. (2014). Are vocabulary tests measurement invariant between age groups? An item response analysis of three popular tests, *Psychology & Aging, 29,* 925-938.

Ghisletta, P. & Lindenberger, U. (2003). Age-based structural dynamics between perceptual speed and knowledge in the Berlin Aging Study: Direct evidence of ability dedifferentiation in old age. *Psychology and Aging, 18,* 696-713.

Ghisletta, P. & de Ribaupierre, A. (2005). A dynamic investigation of cognitive dedifferentiation with control for retest: Evidence from the Swiss Interdisciplinary longitudinal study on the oldest old. *Psychology and Aging, 20,* 671-682.

Goldstein, G., & Shelly, C. H. (1972). Statistical and normative studies of the Halstead Neuropsychological Test Battery relevant to a neuropsychiatric hospital setting. *Journal of Perceptual and Motor Skills, 34*, 603-620.

Goodwin, L. D. & Leech, N. L. (2006). Understanding correlation: Factors that affect the size of *r. Journal of Experimental Education, 74,* 251-266.

Gustafsson, J. (1984). A unifying model of the structure of intellectual abilities. *Intelligence, 8,* 179-203.

Hale, S., Rose, N. S., Myerson, J., Strube, M. J. Sommers, M. Tye-Murray, & Spehar, B. (2011). The structure of working memory abilities across the adult life span. *Psychology & Aging, 26,* 92-110.

Hedden, T. & Yoon, C. (2006). Individual differences in executive processing predict susceptibility to interference in verbal working memory. *Neuropsychology, 20,* 511-528.

Higgins, J. P. T. and Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. Statistics in Medicine, 21: 1539–1558.

Hoelzle, J. B., Nelson, N. W., & Smith, C. A. (2011). Comparison of Wechsler Memory Scale – Fourth Edition (WMS-IV) and Third Edition (WMS-III) dimension structures: Improved ability to evaluate auditory and visual constructs. *Journal of Clinical and Experimental Neuropsychology, 33,* 283-291.

Hertzog, C. (1989). Influences of cognitive slowing on age difference in intelligence. *Developmental Psychology, 25,* 636-651.

Hertzog, C. & Bleckley, M. K. (2001). Age differences in the structure of intelligence: Influences of information processing speed. *Intelligence, 29,* 191-217.

Hertzog, C. Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. S. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology & Aging, 18,* 755-769.

Hildebrandt, A. Wilhelm, O. Schmiedek, F., Herzmann, G., & Sommer, W. (2011). On the specificity of face cognition compared with general cognitive functioning across adult age. *Psychology & Aging, 26,* 701-715.

Hull, R., Martin, R. C., Beier, M. E., Lane, D. & Hamilton, A. C. (2008). Executive function in older adults: A structural equation modeling approach. *Neuropsychology, 22,* 508-522.

Hulur, G. Ram, N., Willis, S. L., Schaie, K. W., & Gerstorf, D. (2015). Cognitive dedifferentiation with increasing age and proximity of death: Within-person evidence from the Seattle Longitudinal Study. *Psychology & Aging, 30,* 311-323.

Johnson, W. Logie, R. H., & Brockmole, J. R. (2010). Working memory tests differ in factor structure across age cohorts: Implications for dedifferentiation. *Intelligence, 38,* 513-528.

Juan-Espinosa, M., Garcia, L. F., Escorial, S. Rebollo, I., Colom, R. & Abad, F. J. (2002). Age dedifferentiation hypothesis: Evidence from the WAIS III. *Intelligence, 30,* 395-408.

Kaufman, A. S., & Kaufman, N. L. (1992). *Kaufman Adolescent and Adult Intelligence Test* (2nd ed.). Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (2004). *Kaufman Brief Intelligence Test* (2nd ed.). Bloomington, MN: Pearson, Inc.

Keith, T. Z. & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools, 47,* 635-650.

Laukka, E. J., Macdonald, S. W. S., Backman, L. (2006). Contrasting cognitive trajectories impending death and preclinical dementia in the very old. *Neurology, 66,* 833-838.

Li, S.-C., Lindenberger, U., Hommel, B. Aschersleben, G., Prinz, W., & Baltes, P. B. (2004). Transformations in the couplings among intellectual abilities and constituent cognitive processes across the life span. *Psychological Science, 15,* 155-163.

Lindenberger, U. & Baltes, P. B. (1997a). Emergence of a powerful connection between sensory and cognitive functions across the adult life span: A new window to the study of cognitive aging? *Psychology & Aging, 12,* 12-21.

Lindenberger, U. & Baltes, P. B. (1997b). Intellectual functioning in old and very old age: Cross-sectional results from the Berlin Aging Study. *Psychology & Aging, 12,* 410-432.

Lovden, M., Bergman, L., Adolfsson, R. Lindenberger, U., Nilsson, L.-G. (2005). Studying individual aging in an interindividual context: Typical paths of age-related, dementia-related, and mortality-related cognitive development in old age. *Psychology & Aging, 20,* 303-316.

Molenaar, D., Dolan, C. V., Wicherts, J. M., van der Maas, H. L. J. (2010). Modeling differentiation of cognitive abilities within the higher-order factor model using moderated factor analysis. *Intelligence, 38,* 611-624.

Niileksela, C. R., Reynolds, M. R., & Kaufman, A. S. (2013). An alternative Cattell-Horn-Carroll (CHC) factor structure of the WAIS-IV: Age invariance of an alternative model for ages 70-90. *Psychological Assessment, 25,* 291-404.

Nyberg, L. Maitland, S. B., Ronnlund, M., Backman, L., Dixon, R. A., & Wahlin, A. (2003). *Psychology & Aging, 18,* 149-160.

Orsini, A. & Laicardi, C. (1997). WAIS-R. *Contribution to the Italian Standardization*. Florence, Itali: Organizzazioni Speciali.

Park, L. Q., Gross, A. L., McLaren, D. G., Pa, J., Johnson, J. K., Mitchell, M., & Manly, J. J. (2012). Confirmatory factor analysis of the ADNI neuropsychological battery. *Brain Imaging & Behavior, 6,* 528-539.

Park, D. C., Lautenschlager, G., Hedden, T., Davidson, N. S. & Smith, A. D. (2002). Models of visospatial and verbal memory across the adult life span. *Psychology & Aging, 17,* 299-320.

Parker, K. (1983). Factor analysis of the WAIS-R at nine age levels between 16 and 74 years. *Journal of Consulting and Clinical Psychology, 51,* 302-308.

R Development Core Team (2008). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. http://www.R-project.org.

Salthouse, T. A. (1995). Influence of processing speed on adult age differences in learning. *Swiss Journal of Psychology, 54,* 102-112.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403-428.

Salthouse, T. A. (2000). Aging and measures of processing speed. *Biological Psychology, 54,* 35-54.

Salthouse, T. A. (2010). Major issues in cognitive aging. Oxford: Oxford University Press.

Salthouse, T. A. (2014). Frequent assessments may obscure cognitive decline. *Psychological Assessment, 26,* 1063-1069.

Salthouse, T. A. & Becker, J. T. (1998). Independent effects of Alzheimer's Disease on neuropsychological functioning. *Neuropsychology, 12,* 242-252.

Salthouse, T. A. & Ferrer-Caja, E. (2003). What needs to be explained to account for age-related effects on multiple cognitive variables? *Psychology & Aging, 18,* 91-110.

Salthouse, T. A. & Saklofske, D. H. (2010). Do the WAIS-IV tests measure the same aspects of cognitive functioning in adults under and over 65? In Weiss, L. G., Saklofske, D., H., Coalson, D., & Raiford, S. E., (Ed.s) WAIS-IV Clinical Use and Interpretation. Elsevier Science.

Salthouse, T. A. (2012). Does the level at which cognitive change occurs change with age? *Psychological Science,* 23, 18–23.

Schaie, K. W., Maitland, S. B., Willis, S. L., Interieri, R. C. (1998). Longitudinal invariance of adult psychometric ability factor structures across 7 years. *Psychology & Aging, 13,* 8-20.

Schaie, K. W., Willis, S. L., Jay, G., Chipuer, H. (1989). Structural invariance of cognitive abilities across the adult life span: A cross-sectional study. *Developmental Psychology, 25,* 652-662.

Schneider, J. A., Aggarwal, N. T., Barnes, L., Boyle, & Bennett, D. A. (2009). The neuropathology of older person with and without dementia from community versus clinic cohorts. *Journal of Alzheimer's Disease, 18,* 691-701.

Schneider, W. J., & McGrew, K. (2012). The Cattell-Horn-Carroll model of intelligence. In, D. Flanagan & P. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues (3rd ed.)* (p. 99-144). New York: Guilford.

Schultz, N. R., Kaye, D. B., Hoyer, W. J. (1980). Intelligence and spontaneous flexibility in adulthood and old age. *Intelligence, 4,* 219-231.

Siedlecki, K. L. Honig, L. S., & Stern, Y. (2008). Exploring the structure of a neuropsychological battery across healthy elders and those with questionable dementia and Alzheimer's Disease. *Neuropsychology, 22,* 400-411.

Sliwinski, M. J., Hofer, S. M., & Hall, C. (2003a). Correlated and coupled cognitive change in older adults with and without preclinical dementia. *Psychology & Aging, 18,* 672-683.

Sliwinski, M. J., Hofer, S. M., Hall, C., Buschke, H., & Lipton, R. B. (2003b). Modeling memory decline in older adults: The importance of preclinical dementia, attrition, and chronological age. *Psychology & Aging, 18,* 658-671.

Sims, R. C., Allaire, J. C., Gamaldo, A. A., Edwards, C. L., & Whitfield, K. E. (2009).

An examination of dedifferentiation in cognition among African-American older adults. *Journal of Cross Cultural Gerontology, 24,* 193-208.

Singer, T. Verhaeghen, Ghisletta, P., Lindenberger, U., Baltes, P. B. (2003). The fate of cognition in very old age: Six-year longitudinal findings in the Berlin Aging Study (BASE). *Psychology & Aging, 18,* 318-331.

Taub, G. E., McGrew, K. S., Witta, E. L. (2004). A confirmatory analysis of the factor structure and cross-age invariance of the Wechsler Adult Intelligence Scale – Third Edition. *Psychological Assessment, 16,* 85-89.

Tucker-Drob, E. M. (2008). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology, 47,* 331-343.

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the lifespan. *Developmental Psychology, 45,* 1097-1118.

Tulsky, D. S. & Price, L. R. (2003). The joint WAIS-III and WMS-III factor structure: Development and cross-validation of a six-factor model of cognitive functioning. *Psychological Assessment, 15,* 149-162.

Uylings, H. B. M. & de Brabander, J. M. (2002). Neuronal changes in normal human aging and Alzheimer's Disease. *Brain & Cognition, 49,* 268-276.

Van den Noortgate, W., Lopez-Lopez, J. A., Marin-Martinez, F., Sanchez-Meca, J. (2013).Three-level meta-analysis of dependent effect sizes. *Behavioral Research Methods, 45,* 576-594.

Vaughn, L., Giovanello, K. (2010). Executive function in daily life: Age-related influence of executive processes on instrumental activities of daily living. *Psychology & Aging, 25,* 343-355.

Viechtbaugher, W. (2010). Conducting Meta-analyses in R with the metaphor package. *Journal of Statistical Software, 36,* 1- 48.

Waller, N. G. & Waldman, I. D. (1990). A reexamination of the WAIS-R factor structure. *Psychological Assessment: A Journal of Consulting and Clinical Psychology, 2,* 139-144.

Ward, L. C. Axelrod, B. N., & Ryan, J. J. (2000). Observations on the factor structure of the WAIS-R. *Assessment, 7,* 79-86.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York, NY: Psychological Corporation.

Wechsler, D. (1981). *Wechsler Adult Intelligence Scale-Revised manual.* New York, NY: The Psychological Corporation.

Wechsler, D. (1987). *Wechsler Memory Scale-Revised manual.* New York, NY: The Psychological Corporation.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition, and Wechsler Adult Intelligence Scale-Third edition technical manual.* San Antonio, TX: Pearson.

Wechsler, D. (1999). *Wechsler Abbreviated Scale of Intelligence (WASI)*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2008b). *Wechsler Adult Intelligence Scale-Fourth Edition: Technical and interpretive manual.* San Antonio, TX: Pearson.

Wechsler, D. (2009). *Wechsler Memory Scale-Fourth Edition: Technical and interpretive manual.* San Antonia, TX: Pearson.

Wilson, R. S., Aggarwal, N. T., Barnes, L. L., Mendes de Leon, C. F., Herbert, L. E., &

Evans, D. A. (2010). Cognitive decline in incident Alzheimer disease in a community population. *Neurology, 74,* 951-955.

Wilson, R. S., Segawa, E., Hizel, L. P., Boyle, P. A., Bennett, D. A. (2012). Terminal dedifferentiation of cognitive abilities. *Neurology, 78,* 1116-1122.

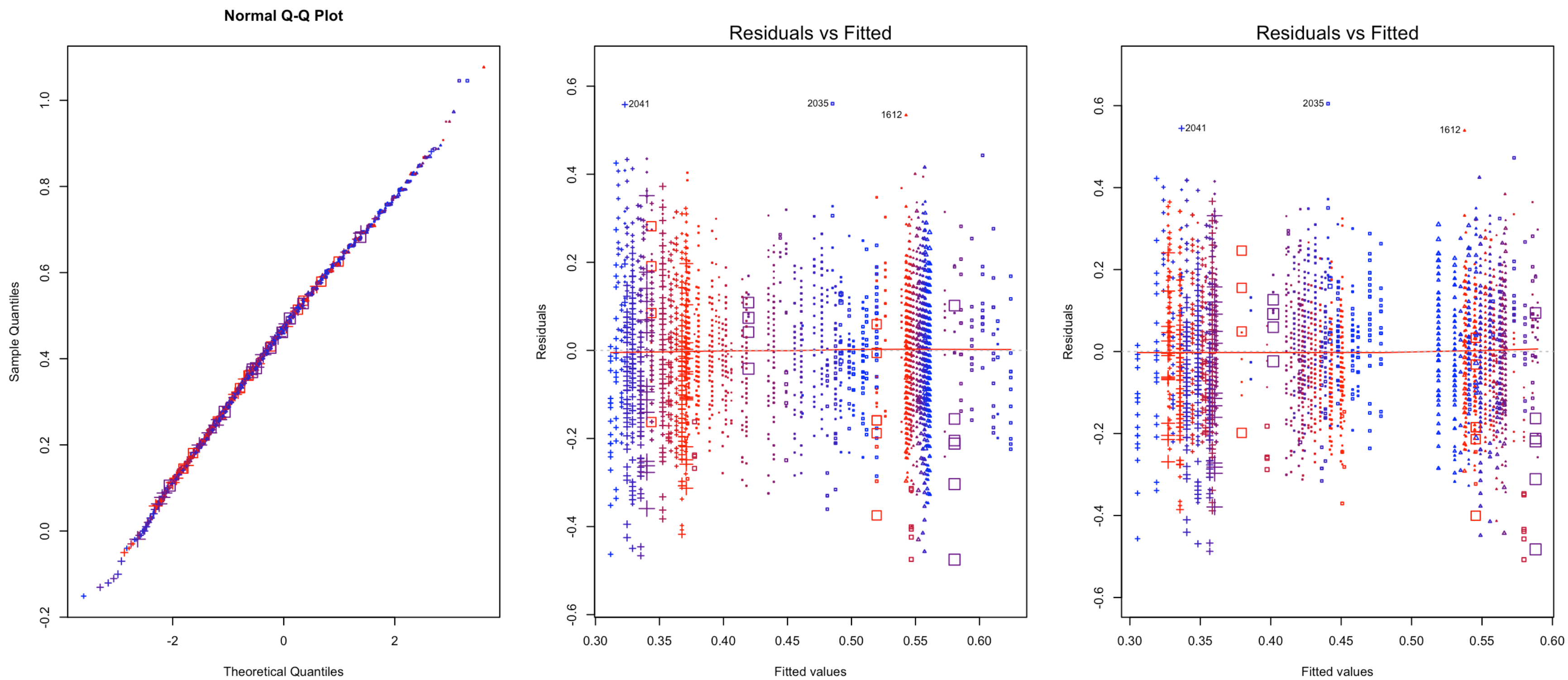
Woodcock, R. W., & Johnson, M. B. (1989). *Woodcock-Johnson Psycho-Educational Battery–Revised*. Itasca, IL: Riverside Publishing.

Zelinski, E. M. & Lewis, K. L. (2003). Adult age differences in multiple cognitive functions: Differentiation, dedifferentiation, or process-specific change? *Psychology & Aging, 18,* 727-745.

Zelinski, E. M. & Stewart, S. T. (1998). Individual differences in 16-year memory changes. *Psychology & Aging, 13,* 622-630.

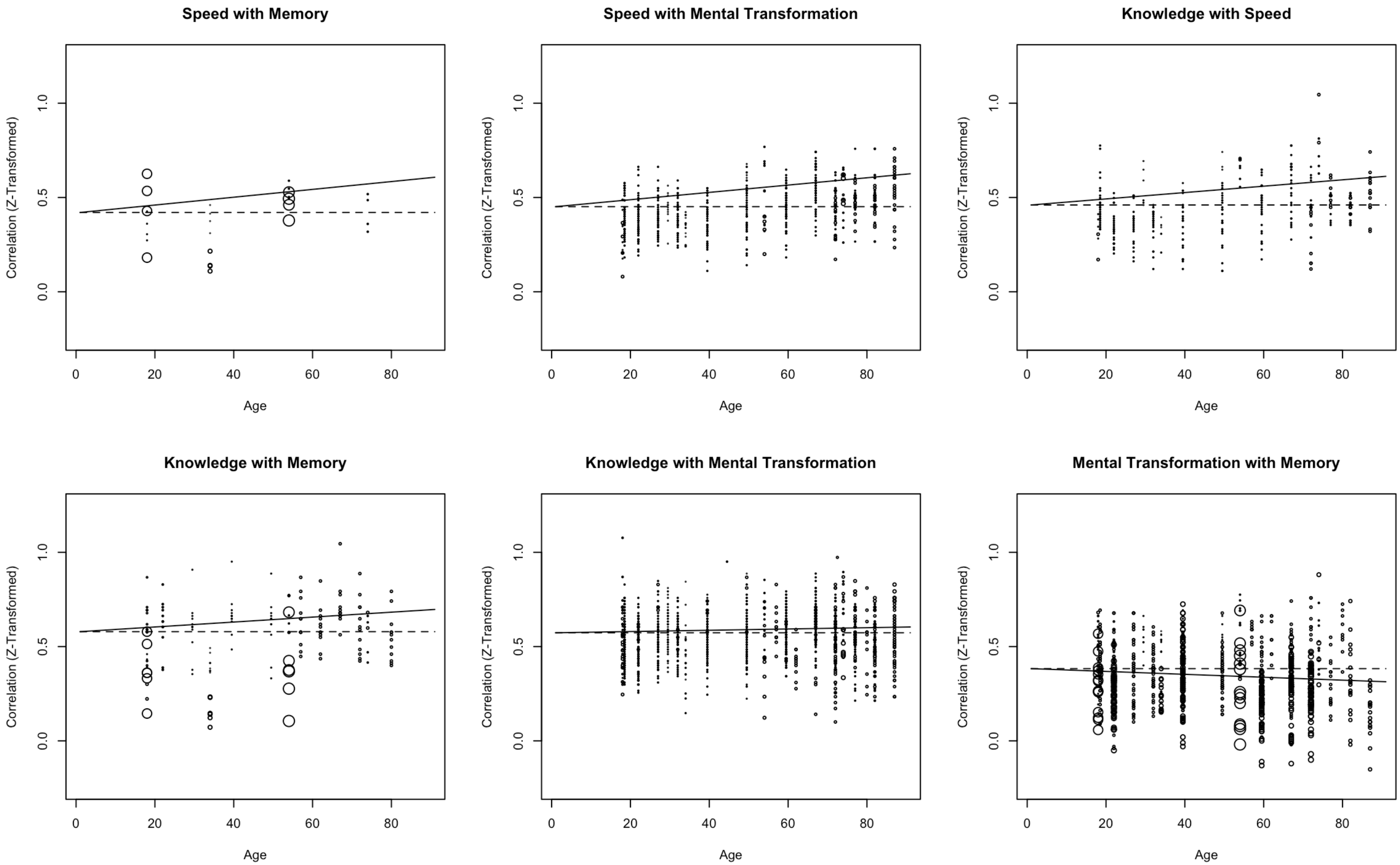
Figures and Tables

Figure 1. QQ-plot and homogeneity plots for age, test type, test type, and the interactions of age with test type and test type.



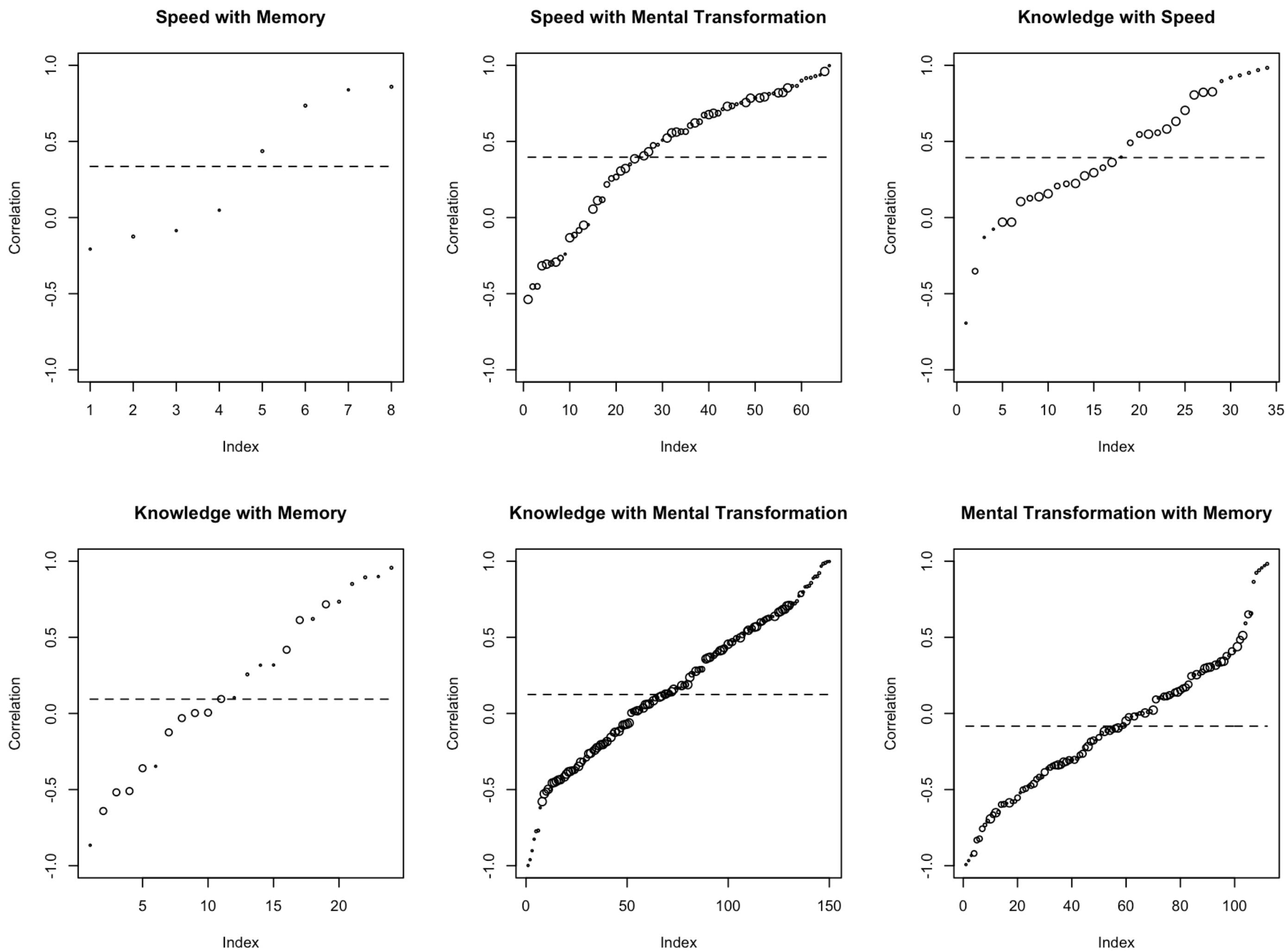
Notes: Within type correlations (S-S, GK-GK, MT-MT, M-M) were not included in this analysis. Age is represented by a gradient going from red to blue (young to old). Symbols range in size based on the weight of the correlation as determined by its sample size. Different symbols represent different types: K-M = circle, M-S = triangle point up, K-MT = plus sign, MT-M = cross, S-K = diamond, MT-S = triangle point down.

Figure 2. Relations between linear age and z-transformed correlations for each of the six types.



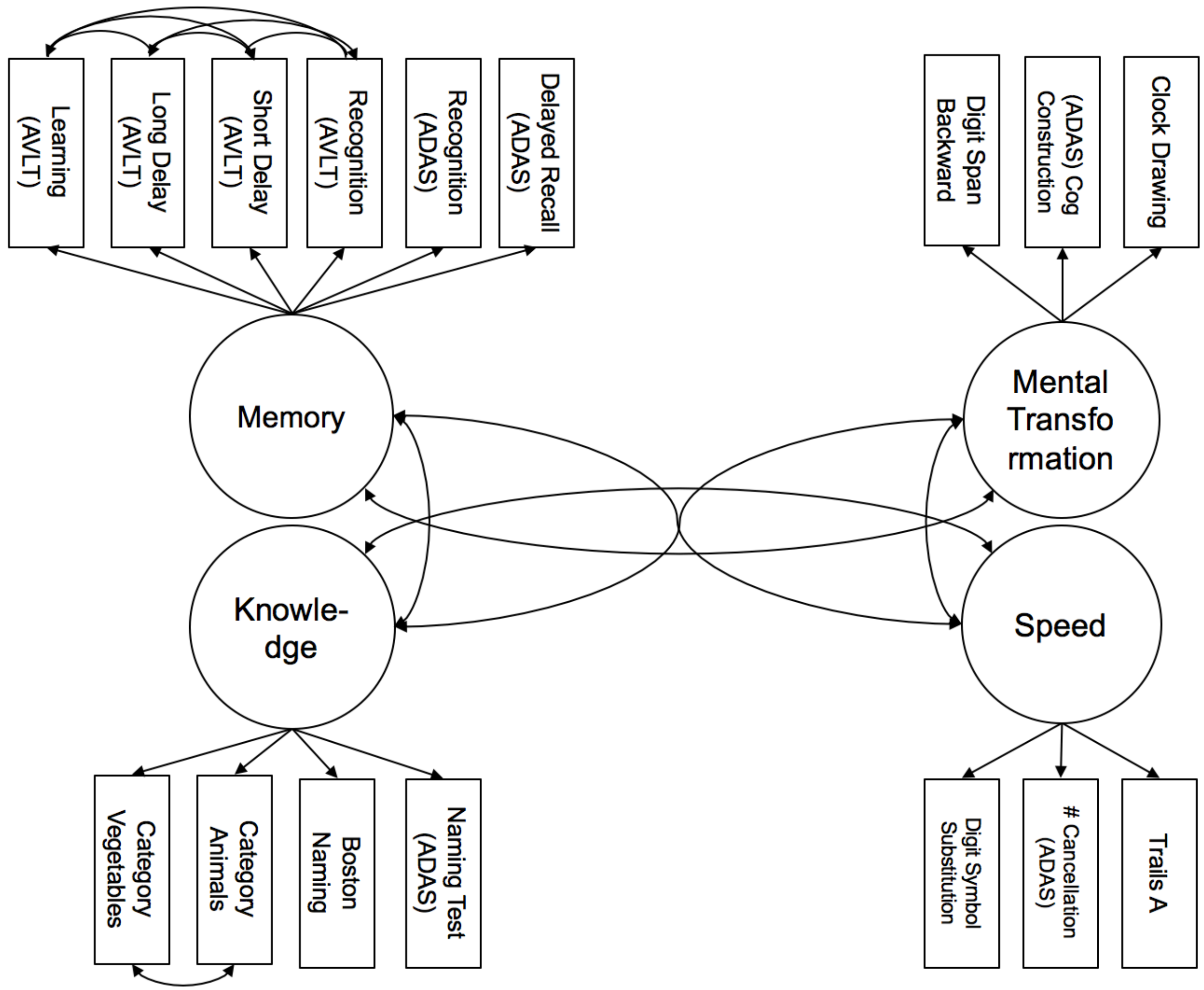
Notes: Increasing circle size represents increasing weights/larger N’s for the individual samples represented.

Figure 3. Correlations for age with predicted correlation size for each combination of tests.



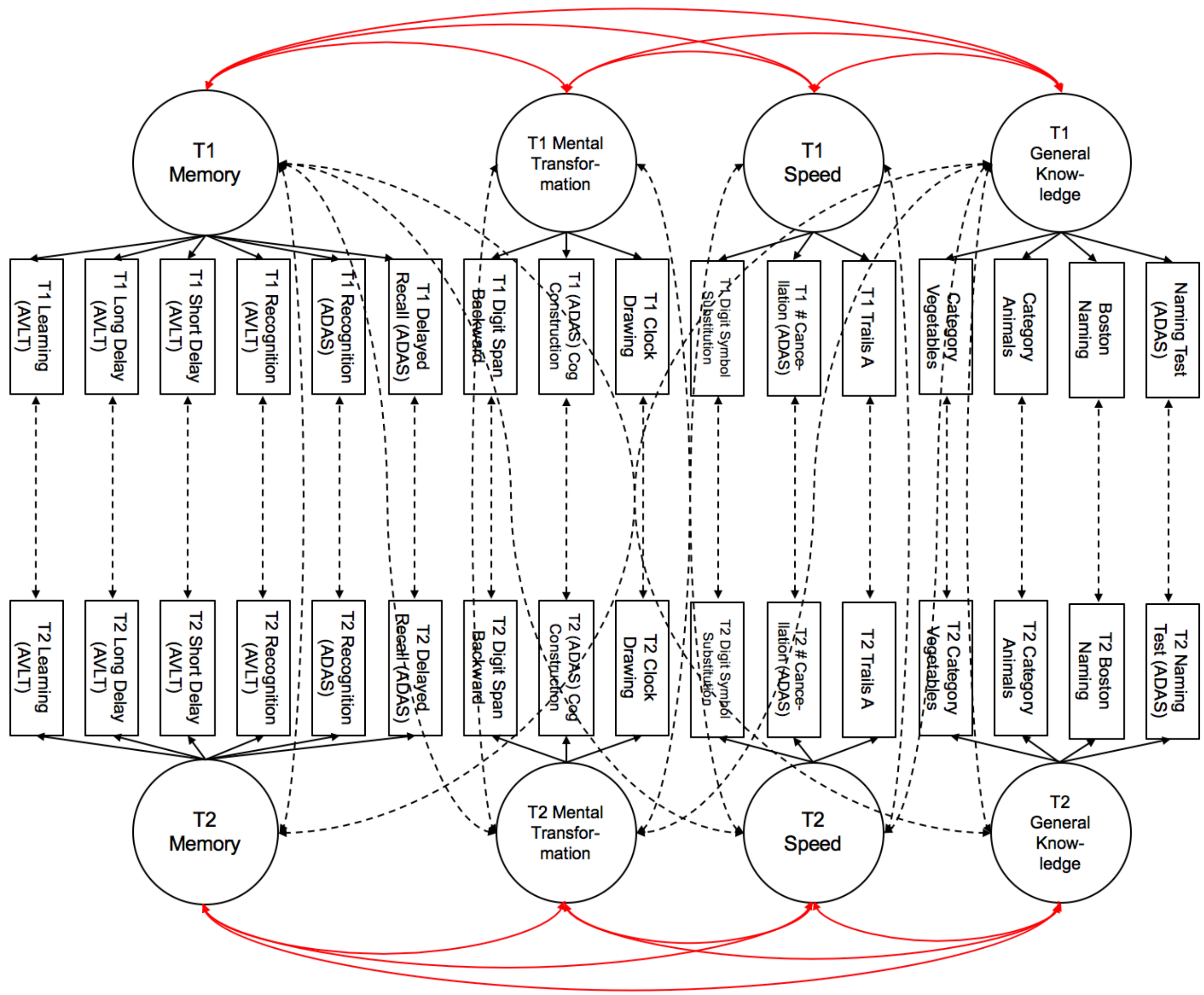
Notes: Increasing circle size represents increasing weights/larger N's for the individual samples represented. The correlations were ordered in terms of correlation size and assigned order numbers, which are on the x-axis (Index). The dashed lines are the weighted (by the N's of each correlation) mean of the correlations.

Figure 4. Four-factor hierarchical model used for the baseline and alternative baseline model.



Notes: ADAS = Alzheimer's Disease Assessment Scale. AVLT: Adult Verbal Learning Test.

Figure 5. Proposed four-factor hierarchical temporal invariance model.



Notes: In the loadings constrained model, the solid black lines were held constant for T1 and T2. For the loadings and covariances model, the solid black and red lines were held constant for T1 and T2. Method correlations excluded for simplicity. ADAS = Alzheimer's Disease Assessment Scale. AVLT: Adult Verbal Learning Test.

Table 1. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the S-MT type and then compared to the five other groups.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-MT)	0.4727	0.0123	0.4485	0.4969	38.34	**< .0001**
Age (S-MT)	0.0020	.0005	0.0010	0.0031	3.76	**0.0002**
Age^2 (S-MT)	0.0096	0.0145	-0.0188	0.038	0.66	0.5083
Intercept (K-M)	0.0916	0.0091	0.0737	0.1095	10.01	**< .0001**
Intercept (S-M)	-0.0593	0.0184	-0.0953	-0.0233	-3.23	**0.0012**
Intercept (K-MT)	0.0974	0.0039	0.0898	0.1050	25.09	**< .0001**
Intercept (MT-M)	-0.0494	0.0053	-0.0598	-0.0391	-9.36	**< .0001**
Intercept (S-K)	-0.0052	0.0054	-0.0157	0.0053	-0.96	0.3349
Age (K-M)	-0.0020	.0004	-0.0029	-0.0011	-4.44	**< .0001**
Age (S-M)	-0.0037	.0009	-0.0055	-0.0019	-4.11	**< .0001**
Age (K-MT)	-0.0016	.0002	-0.0019	-0.0012	-8.51	**< .0001**
Age (MT-M)	-0.0030	.0003	-0.0035	-0.0025	-11.96	**< .0001**
Age (S-K)	-.0002	.0003	-.0007	.0003	-0.68	0.4995
Age^2 (K-M)	-0.0550	0.012	-0.0785	-0.0315	-4.58	**< .0001**
Age^2 (S-M)	-0.0596	0.0236	-0.1059	-0.0134	-2.53	**0.0115**
Age^2 (K-MT)	-0.0168	0.005	-0.0265	-0.0071	-3.38	**.0007**
Age^2 (MT-M)	-0.0399	0.0068	-0.0533	-0.0266	-5.86	**< .0001**
Age^2 (S-K)	0.0078	0.0069	-0.0057	0.0212	1.13	0.2567

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table 2. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the S-K type and then compared to the four other groups.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-K)	0.4677	0.0127	0.4427	0.4926	36.78	<.0001
Age (S-K)	0.0019	0.0006	0.0008	0.0030	3.37	0.0007
Age^2 (S-K)	0.0178	0.015	-0.0117	0.0473	1.18	0.2364
Deviations from						
Intercept (K-M)	0.0912	0.0097	0.0722	0.1102	9.41	<.0001
Intercept (S-M)	-0.0722	0.0188	-0.1091	-0.0353	-3.83	0.0001
Intercept (K-MT)	0.1034	0.0049	0.0938	0.113	21.16	<.0001
Intercept (MT-M)	-0.0461	0.0061	-0.0581	-0.0341	-7.53	<.0001
Age (K-M)	-0.0019	0.0005	-0.0028	-0.001	-4.03	0.0001
Age (S-M)	-0.004	0.0009	-0.0058	-0.0022	-4.30	<.0001
Age (K-MT)	-0.0014	0.0002	-0.0018	-0.0009	-6.06	<.0001
Age (MT-M)	-0.0028	0.0003	-0.0034	-0.0023	-9.78	<.0001
Age^2 (K-M)	-0.0688	0.0127	-0.0936	-0.044	-5.43	<.0001
Age^2 (S-M)	-0.0852	0.0242	-0.1326	-0.0377	-3.51	0.0004
Age^2 (K-MT)	-0.0244	0.0062	-0.0366	-0.0121	-3.91	0.0001
Age^2 (MT-M)	-0.0496	0.0079	-0.0651	-0.0342	-6.29	<.0001

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table 3. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the MT-M type and then compared to the three other groups.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (MT-M)	0.4179	0.0119	0.3945	0.4413	35	<.0001
Age (MT-M)	-0.0008	0.0005	-0.0019	0.0002	-1.57	0.1153
Age^2 (MT-M)	-0.0302	0.0142	-0.0581	-0.0024	-2.13	0.0334
Deviations from						
Intercept (K-M)	0.1414	0.0089	0.1238	0.1589	15.81	<.0001
Intercept (S-M)	-0.0151	0.0187	-0.0518	0.0215	-0.81	0.4184
Intercept (K-MT)	0.1526	0.0048	0.1431	0.1621	31.61	<.0001
Age (K-M)	0.0010	0.0004	0.0001	0.0019	2.27	0.0235
Age (S-M)	-0.0007	0.0009	-0.0025	0.0011	-0.76	0.4446
Age (K-MT)	0.0013	0.0002	0.0008	0.0018	5.63	<.0001
Age^2 (K-M)	-0.0210	0.0118	-0.0441	0.0021	-1.78	0.0755
Age^2 (S-M)	-0.0380	0.0241	-0.0851	0.0092	-1.58	0.1144
Age^2 (K-MT)	0.0236	0.0063	0.0113	0.0359	3.77	0.0002

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table 4. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the K-MT type and then compared to the two other groups.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (K-MT)	0.5747	0.0101	0.5549	0.5944	56.96	<.0001
Age (K-MT)	.0005	.0004	-.0004	0.0014	1.09	0.2757
Age^2 (K-MT)	-0.0111	0.0119	-0.0345	0.0123	-0.93	0.3525
Deviations from						
Intercept (K-M)	0.0444	0.0093	0.0261	0.0627	4.76	<.0001
Intercept (S-M)	-0.1358	0.0192	-0.1734	-0.0982	-7.07	<.0001
Age (K-M)	.0006	.0005	-.0003	0.0015	1.29	0.1975
Age (S-M)	-0.0016	.0009	-0.0034	.0003	-1.69	0.0918
Age^2 (K-M)	-0.0533	0.0123	-0.0773	-0.0292	-4.34	<.0001
Age^2 (S-M)	-0.0796	0.0248	-0.1281	-0.0311	-3.22	0.0013

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table 5. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in S-M type and then compared to K-M.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-M)	0.4890	0.0344	0.4217	0.5564	14.23	<.0001
Age (S-M)	-0.0009	0.0015	-0.0039	0.0021	-0.58	0.5643
Age^2 (S-M)	-0.0357	0.0419	-0.1177	0.0464	-0.85	0.3945
Deviations from						
Intercept (K-M)	0.0891	0.0228	0.0443	0.1338	3.90	0.0001
Age (K-M)	0.0023	0.0011	0.0001	0.0045	2.02	0.0433
Age^2 (K-M)	0.0275	0.0294	-0.0301	0.0851	0.94	0.3487

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table 6. Estimates of the correlation meta-analysis where the average age predicts correlation size for each of types.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
S-M (28)						
intercept	0.4183	0.0396	0.3406	0.4960	10.55	**<.0001**
age	0.0021	0.0019	-0.0016	0.0057	1.11	0.2679
S-MT (554)						
intercept	0.4494	0.0119	0.4261	0.4726	37.85	**<.0001**
age	0.0019	0.0005	0.0009	0.0030	3.67	**0.0002**
S-K (292)						
intercept	0.4585	0.0199	0.4195	0.4975	23.04	**<.0001**
age	0.0017	0.0009	-0.0000	0.0034	1.91	**0.0562**
K-M (162)						
intercept	0.5778	0.0280	0.5229	0.6327	20.64	**<.0001**
age	0.0013	0.0011	-0.0009	0.0035	1.15	0.2485
K-MT (1256)						
intercept	0.5729	0.0104	0.5526	0.5933	55.20	**<.0001**
age	0.0003	0.0005	-0.0005	0.0012	0.75	0.4504
MT-M (882)						
intercept	0.3838	0.0203	0.3859	0.4236	18.89	**<.0001**
age	-0.0008	0.0009	-0.0025	0.0010	-0.87	0.3859

Notes: Items in bold are significant at $p < .05$. K's are in parentheses. Used the unconstrained data. Random effects for sample were included.

Table 7. Estimates of the correlation meta-analysis where the average age predicts correlation size for each of types.

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
S-M (28)						
intercept	0.4591	0.0763	0.3095	0.6087	6.02	**<.0001**
age	0.0017	0.0023	-0.0028	0.0062	0.75	0.4513
age^2	-0.0437	0.0644	-0.1699	0.0825	-0.68	0.4972
S-MT (554)						
intercept	0.4538	0.0195	0.4157	0.4920	23.32	**<.0001**
age	0.0020	0.0006	0.0009	0.0031	3.57	**0.0004**
age^2	-0.0042	0.0143	-0.0322	0.0239	-0.29	0.7720
S-K (292)						
intercept	0.4643	0.0327	0.4003	0.5283	14.22	**<.0001**
age	0.0017	0.0009	-0.0001	0.0036	1.88	**0.0605**
age^2	-0.0054	0.0240	-0.0523	0.0416	-0.22	0.8233
K-M (162)						
intercept	0.5859	0.0435	0.5005	0.6712	13.46	**<.0001**
age	0.0014	0.0012	-0.0010	0.0037	1.16	0.2442
age^2	-0.0080	0.0323	-0.0714	0.0554	-0.25	0.8050
K-MT (1256)						
intercept	0.5905	0.0168	0.5576	0.6235	35.14	**<.0001**
age	0.0005	0.0005	-0.0004	0.0014	1.10	0.2702
age^2	-0.0163	0.0122	-0.0403	0.0077	-1.33	0.1824
MT-M (882)						
intercept	0.4126	0.0317	0.3504	0.4748	13.00	**<.0001**
age	-0.0005	0.0009	-0.0023	0.0013	-0.56	0.5787
age^2	-0.0291	0.0246	-0.0775	0.0192	-1.18	0.2370

Notes: Items in bold are significant at $p < .05$. K's are in parentheses. Used the unconstrained data. Random effects for sample were included.

Table 8. Means and standard deviations for each variable in the total sample and each of the conditions at T1 and T2.

	T1			T2		
	Total	CDR < 1.5	CDR >= 1.5	Total	CDR < 1.5	CDR >= 1.5
N	819	389	430	451	294	157
Age	75.19 (6.84)	75.61 (5.95)	74.80 (7.54)	75.19 (6.84)	75.61	74.80 (7.54)
Proportion Female	.42	.42	.42	.42	.42	.42
Proportion CN	.28	.59	.00	.41	.64	.00
Proportion MCI	.48	.41	.55	.56	.36	.92
Proportion AD	.24	.00	.45	.03	.00	.08
MMSE	26.74 (2.67)	28.38 (1.64)	25.27 (2.57)	27.89 (1.93)	28.56 (1.50)	26.63 (2.03)
AVLT Learning (1-5)	3.66 (2.66)	4.93 (2.61)	2.52 (2.15)	3.81 (2.74)	4.54 (2.75)	2.38 (2.09)
AVLT Short Delay	4.49 (3.82)	6.63 (3.81)	2.57 (2.62)	4.82 (4.16)	6.07 (4.20)	2.41 (2.79)
Delayed Recall (ADAS)	*5.82 (2.86)*	*4.02 (2.38)*	*7.45 (2.22)*	*5.48 (3.25)*	*4.35 (2.92)*	*7.65 (2.70)*
AVLT Long Delay	3.62 (4.00)	5.85 (4.10)	1.60 (2.58)	4.10 (4.19)	5.40 (4.22)	1.58 (2.74)
AVLT Recognition	10 (4.01)	11.80 (3.21)	8.37 (3.97)	10.20 (4.55)	11.57 (3.75)	7.53 (4.79)
Word Recognition (ADAS)	*4.51 (3.00)*	*3.32 (2.60)*	*5.58 (2.94)*	*4.05 (3.31)*	*3.07 (2.86)*	*5.91 (3.32)*
Digit Span Backward	6.19 (2.17)	6.90 (2.16)	5.53 (1.97)	6.52 (2.34)	6.96 (2.35)	5.62 (2.04)
Construction (ADAS)	*.58 (.62)*	*.44 (.54)*	*.70 (.66)*	*.55 (.69)*	*.47 (.59)*	*.72 (.82)*
Clock Drawing	4.12 (1.11)	4.49 (.82)	3.79 (1.23)	4.22 (1.14)	4.48 (.91)	3.73 (1.36)
Category Vegetables	11.15 (4.35)	13.21 (4.17)	9.28 (3.59)	11.21 (5.27)	12.81 (4.87)	8.04 (4.58)
Category Animals	16.18 (5.79)	18.33 (5.62)	14.23 (5.22)	16.39 (6.56)	18.09 (6.14)	13.09 (6.09)
Boston Naming	25.44 (4.74)	26.97 (3.30)	24.06 (5.39)	25.74 (5.29)	27.16 (3.98)	23.00 (6.32)
Naming (ADAS)	*.28 (.58)*	*.14 (.36)*	*.41 (.70)*	*.27 (.70)*	*.13 (.40)*	*.54 (1.00)*
Trails A	*47.80 (27.22)*	*38.05 (13.94)*	*56.66 (32.78)*	*45.24 (29.55)*	*37.90 (20.21)*	*59.82 (38.52)*
Digit Span Substitution	37.03 (13.15)	42.93 (11.03)	31.64 (12.61)	39.70 (14.75)	43.68 (13.38)	31.75 (38.52)
Cancellation (ADAS)	*1.01 (1.09)*	*.60 (.80)*	*1.39 (1.19)*	*.94 (1.20)*	*.64 (.91)*	*1.55 (1.45)*

Note: Standard deviations reported in parentheses. For tests in italics, lower scores represent poorer performance.

Table 9. Measurement invariance across functional impairment group for the baseline analyses.

Indicator	Configural Invariance		Metric Invariance	
	Less Impaired	More Impaired	Less Impaired	More Impaired
N	389	430	389	430
Memory				
AVLT Learning (1-5)	1.2 (.70)*	0.8 (.56)*	1.1 (.63)*	1.1 (.63)*
AVLT Short Delay	1.5 (.83)*	0.8 (.64)*	1.2 (.74)*	1.2 (.74)*
Delayed Word Recall	*1.2 (.82)**	*1.2 (.87)**	*1.3 (.79)**	*1.3 (.79)**
AVLT Long Delay	1.5 (.80)*	0.8 (.66)*	1.1 (.76)*	1.1 (.76)*
AVLT Recognition	1 (.69)	1 (.56)	1 (.49)	1 (.49)
Word Recognition	*0.8 (.53)**	*1.0 (.55)**	*0.9 (.47)**	*0.9 (.47)**
Mental Transformation				
Digit Span Backward	.9 (.38)*	0.6 (.50)*	0.7 (.53)*	0.7 (.53)*
Construction	*1.1 (.52)**	*0.8 (.60)**	*0.9 (.61)**	*0.9 (.61)**
Clock Drawing	1 (.55)	1 (.70)*	1 (.68)	1 (.68)
Speed				
Trails A	0.4 (.59)*	1.1 (.78)*	0.9 (.72)*	0.9 (.72)*
Digit Symbol Substitution	1 (.86)	1 (.88)	1 (.91)	1 (.91)
Cancellation	*0.5 (.48)**	*0.9 (.70)**	*0.8 (.68)**	*0.8 (.68)**
General Knowledge				
Category Vegetables	1.2 (.59)*	0.7 (.71)*	0.9 (.78)*	0.9 (.78)*
Category Animals	1.3 (.65)*	0.8 (.76)*	1.0 (.82)*	1.0 (.82)*
Boston Naming	1 (.69)	1 (.74)	1 (.67)	1 (.67)
Naming	*0.6 (.49)**	*0.8 (.59)**	*0.7 (.45)**	*0.7 (.45)**
Model Fit Statistics				
RMSEA	.036		.045	
CFI	.962		.934	
χ^2, df	370.60, 182		517.28, 194	

Notes: Raw (standardized) factor loadings are shown. Standardized loadings are correlations between the observed indicator and latent factor. The difference in degrees of freedom between the configural and metric invariance models is 12 because 12 factor loadings were constrained to be the same. RMSEA = root mean squared error of approximation, CFI = comparative fit index. Tests in italics are reverse coded so that all scores represented higher performance.

Table 10. Correlations between latent constructs for less impaired versus more impaired in the metric invariance model.

	T1	
	CDR < 1.5	CDR >= 1.5
N	389	430
Memory w/ Mental Transformation	.59*	.36*
Knowledge w/ Memory	.60*	.55*
Memory w/ Speed	.44*	.33*
Knowledge w/ Mental Transformation	.70*	.64*
Mental Transformation w/ Speed	.61*	.83*
Knowledge w/ Speed	.65*	.58*

Note: *p < .05. Items in bold were significantly different across the conditions.

Table 11. Correlations between T1 and T2 for each of the individual tests.

	CDR < 1.5	CDR >= 1.5
AVLT Learning (1-5)	.44*	.38*
AVLT Short Delay	.69*	.54*
Delayed Word Recall	.76*	.67*
AVLT Long Delay	.67*	.63*
AVLT Recognition	.46*	.60*
Word Recognition	.38*	.54*
Digit Span Backward	.61*	.66*
Construction	.36*	.35*
Clock Drawing	.34*	.45*
Category Vegetables	.71*	.61*
Category Animals	.68*	.62*
Boston Naming	.76*	.72*
Naming (ADAS)	.35*	.52*
Trails A	.51*	.65*
Digit Span Substitution	.80*	.77*
Cancellation	.42*	.44*

Note: *p < .05.

Table 12. Measurement invariance across time points for the less impaired condition for the temporal invariance model.

Indicator	Configural		Metric (loadings + correlations constrained)	
	T1	T2	T1	T2
N	389	294	389	294
Memory				
AVLT Learning (1-5)	1.2 (.66)*	1.0 (.68)*	1.1 (.66)*	1.1 (.66)*
AVLT Short Delay	1.5 (.83)*	1.3 (.85)*	1.4 (.82)*	1.4 (.84)*
Delayed Word Recall	1.3 (.84)*	1.2 (.89)*	1.2 (.85)*	1.2 (.87)*
AVLT Long Delay	1.5 (.81)*	1.3 (.83)*	1.4 (.80)*	1.4 (.83)*
AVLT Recognition	1 (.69)*	1 (.80)*	1 (.72)*	1 (.76)*
Word Recognition	0.8 (.53)*	0.9 (.72)*	0.9 (.58)*	0.9 (.64)*
Mental Transformation				
Digit Span Backward	1.0 (.39)*	1.0 (.59)*	1.0 (.44)*	1.0 (.50)*
Construction	1.0 (.48)*	0.7 (.46)*	0.8 (.43)*	0.8 (.47)*
Clock Drawing	1 (.55)*	1 (.73)*	1 (.60)*	1 (.67)*
Speed				
Trails A	0.4 (.58)*	0.6 (.72)*	0.5 (.66)*	0.5 (.62)*
Digit Symbol Substitution	1 (.87)*	1 (.89)*	1 (.86)*	1 (.91)*
Cancellation	0.5 (.49)*	0.7 (.64)*	0.6 (.57)*	0.6 (.55)*
General Knowledge				
Category Vegetables	1.2 (.62)*	1.2 (.75)*	1.2 (.65)*	1.2 (.73)*
Category Animals	1.3 (.67)*	1.2 (.79)*	1.3 (.69)*	1.3 (.76)*
Boston Naming	1 (.69)*	1 (.77)*	1 (.72)*	1 (.72)*
Naming	0.6 (.48)*	0.6 (64)*	0.6 (.52)*	0.6 (.59)*
Model Fit Statistics				
RMSEA	.043		.047	
CFI	.95		.94	
χ^2, df	691.38, 406		790.82, 424	

Notes: *p < .05. Raw (standardized) factor loadings are shown. Standardized loadings are correlations between the observed indicator and latent factor. The difference in degrees of freedom between the configural and metric invariance models is 18 because 12 factor loadings and 6 covariances were constrained to be the same across time. RMSEA = root mean squared error of approximation, CFI = comparative fit index. Tests in italics are reverse coded so that all scores represented higher performance.

Table 13. Measurement invariance across time points for the more impaired condition for the temporal invariance model.

Indicator	Configural		Metric (loadings + correlations constrained)	
	T1	T2	T1	T2
N	430	157	430	157
Memory				
AVLT Learning (1-5)	.84 (.62)*	.51 (.56)*	.67 (.56)*	.67 (.57)*
AVLT Short Delay	.86 (.75)*	.62 (.78)*	.74 (.72)*	.74 (.77)*
Delayed Word Recall	1.1 (.86)*	.84 (.85)*	.98 (.83)*	.98 (.84)*
AVLT Long Delay	.85 (.79)*	.56 (.74)*	.70 (.72)*	.70 (.74)*
AVLT Recognition	1 (.60)*	1 (.77)*	1 (.65)*	1 (.68)*
Word Recognition	.87 (.53)*	.92 (.78)*	.91 (.60)*	.91 (.68)*
Mental Transformation				
Digit Span Backward	.57 (.50)*	.62 (.70)*	.57 (.53)*	.57 (.55)*
Construction	.81 (.60)*	.87 (.73)*	.82 (.64)*	.82 (.58)*
Clock Drawing	1 (.71)*	1 (.82)*	1 (.74)*	1 (.72)*
Speed				
Trails A	1.12 (.79)*	1.4 (.90)*	1.19 (.81)*	1.19 (.81)*
Digit Symbol Substitution	1 (.87)*	1 (.91)*	1 (.87)*	1 (.90)*
Cancellation	.91 (.71)*	1.5 (.90)*	1.03 (.75)*	1.03 (.72)*
General Knowledge				
Category Vegetables	.84 (.78)*	.66 (.81)*	.74 (.76)*	.74 (.77)*
Category Animals	.93 (.79)*	.80 (.89)*	.84 (.79)*	.86 (.86)*
Boston Naming	1 (.68)*	1 (.84)*	1 (.74)*	1 (.76)*
Naming	.86 (.54)*	1.12 (.80)*	.95 (.64)*	.95 (.65)*
Model Fit Statistics				
RMSEA	.037		.044	
CFI	.95		.93	
χ^2, df	647.97, 406		777.33, 424	

Notes: *p < .05. Raw (standardized) factor loadings are shown. Standardized loadings are correlations between the observed indicator and latent factor. The difference in degrees of freedom between the configural and metric invariance models is 18 because 12 factor loadings and 6 covariances were constrained to be the same across time. RMSEA = root mean squared error of approximation, CFI = comparative fit index. Tests in italics are reverse coded so that all scores represented higher performance.

Supplementary Tables.

Table S1. A summary of studies testing for dedifferentiation as evidence for structural differences, increased g variance, and higher correlations magnitudes.

Source	**Constructs**	**Type**	**N**	**Age Range**	**Comparison**	**Results**	**Strength of Evidence**
Adrover-Roig (2012)	Mental Transformation	C	122	48 - 91	Configural invariance	A two-factor solution was found to provide the best fit an older sample, whereas an earlier sample found the best fit for younger adults was a three factor fit (Miyake et al., 2000)	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age
Anstey (2003)	Speed Memory Knowledge	C/L	1823	70 – 85 + (8 yrs)	Correlation magnitude	No consistent evidence of increased correlations, cross-sectionally or longitudinally	Strong; multiple measures of cognitive abilities; multiple age groups; 8 years over time; statistically tested
Babcock (1997)	Speed	C	144, 105	18 - 24, 55 - 80	Configural invariance, Metric invariance, Variance explained by g	Same number of factors found across age groups. Factor loadings from tests were invariant, but interfactor loadings were higher in older adults. A higher order factor explained more variance in g.	Strong; model fit was statistically significantly worse by forcing interfactor loadings to be the same across age groups.
Balsamo (2010)	Knowledge Mental Transformation	C	267, 256	65 – 74, 74 +	Configural invariance	A 1 factor solution provided the best fit for WAIS-R Italian older sample, whereas studies looking at the younger groups had better fits for 3 and 2 factor solutions, Orsini & Laicardi	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age

						(1997, 2003)	
Baltes (1980)	Speed Memory Mental Transformati on Knowledge	C	109	60 - 89	Configural invariance	The model was more integrated than previous work in done with college students, however that research was done with different variables	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age
Baltes (1997)	Speed Memory Mental Transformati on Knowledge	C	171, 516	25 - 69, 70 - 103	Correlation magnitude	Relations between sensory and intellectual functioning increased from 11% in adulthood to 31% in older adulthood	Strong; except for reasoning, significantly more variance predicted by hearing/vision in old age group
Batterham (2011)	Speed Knowledge Memory	L	687	70 - 97 + (12 yrs)	Variance explained by g	Only 2/7 cognitive tests showed significant dedifferentiation, which was attenuated by controlling for possible cognitive impairment	Strong; multiple measures of cognitive abilities; multiple age groups; linear and quadratic effects; 17 years over time
Benson (2010)	Speed Knowledge Mental Transformati on Memory	C	2200	16-90	Configural invariance, metric invariance	Constraining the factor loadings from subtests to latent factors in the WAIS IV standardization sample across age resulted in a significant loss in fit	Strong; metric invariance resulting in significant loss of fit, but not in the direction of dedifferentiation (higher loadings in younger ages)
Bowden	Speed	C	1299	16-89	Configural	With adjacent groups, strict	Strong; Overall, evidence is

(2006)	Knowledge Mental Transformation Memory				invariance, metric invariance	invariance was found. However, when comparing the youngest and oldest groups, there some evidence of differences in the factor loadings and theta, due to largest factor loading for the picture completion subtest in the oldest age group	in favor with metric invariance. And even though there is a difference, it appears to be due to only one measure and only found between the extreme groups
Burton (1994)	Speed Knowledge Mental Transformation	C	225	75-79, 80-96	Configural invariance	Out of 7 models tested, the same model was found to provide the best fit for both groups	Weak; correlations could still be significantly larger (consistent with dedifferentiation) even though configural invariance is found; this was not tested
Cohen (1957)	Memory Mental Transformation Speed Knowledge	C	1152	18-over 75	Configural invariance, metric invariance	The same model fit in all groups except the 60+ group. In this case, *greater differentiation* was evidence by tests having weaker loadings, as well as a change in what tests load onto which factors for the WAIS-R	Weak; failure of configural invariance due to tests loading onto different factors confounds the meaning of the changes in factor loadings
Cunningham (1980)	Speed Knowledge	C/L	96 L/123 C	College Students (40 yrs, 3 T’s), additional college sample at	Time-lag comparisons; Configural invariance, metric invariance	Comparisons were invariant across college students in 1919 and at 1960. Some increase in factor loadings between 1919 and 1950, no model that fit in these samples could be fitted in	Weak; large differences are found in the factor loadings (direction consistent with dedifferentiation) but the strength of that difference is not statistically tested.

				T3		1960. Correlations above .7 at T3 versus basically 0. (1960)	
De Frias (2006)	Mental Transformation (Executive Functioning)	C/L	427	55 – 85	Configural Invariance	A two-factor solution was found to provide the best fit an older sample, whereas an earlier sample found the best fit for younger adults was a three factor fit (Miyake et al., 2000)	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age
De Frias (2007)	Memory Mental Transformation Knowledge	L	1000 (649 by T3)	35 – 80, (10 yrs, 3 T's)	Correlation magnitude between age groups and across time	Correlations at T1 were stronger in oldest age group versus other; Variances in slope increase in that group as well	Strong; trying to constrain the model covariances to be the same resulted in a statistically significant loss of fit.
Deary (2004)	Mental Transformation Memory	C/L	353, 74	65 VS 78	Correlation Magnitude/ g Variance	Correlations were universally higher in the older vs. younger sample, g explained 60.6 of the variance in the older vs 47.3 in the younger	Strong; all correlations were higher and 6/10 were significantly higher
Escorial (2003)	Knowledge Mental Transformation Speed	C	719	16-54	Variance explained by g	Variance explained by g was not significantly greater in older age groups of the WAIS-III Spanish norming samples	Strong; specific statistical tests of g variance reveal no differences, but age groups relatively young (see age range)
Finkel (2007)	Memory Speed	L	806	50-88 (16 yrs)	Correlation magnitude	Adding coupling parameters between measures of verbal	Weak; testing whether parameters add additional fit,

	Knowledge Mental Transformation					and spatial factors (gc and gf) did not improve model fit. However, including a speed -> memory and speed -> space did improve fit, whereas the reverse did not.	but not whether correlations are increasing. However, directionality is interesting
Ghisletta (2003a)	Speed Knowledge	L		80-85 (3 yrs)	Variance explained by "mechanics" test in "pragmatics" test	With increasing age, changes in a processing speed test explains increasing more variance in a knowledge test, but not the other way around	Weak; testing whether parameters add additional fit, but not whether correlations are increasing. However, directionality is interesting
Ghisletta (2003b)	Speed Knowledge	L	516	70 – 103 (34 yrs, every 2 yrs)	Variance in explained by speed test in knowledge test	With increasing age, changes in a processing speed test explains increasing more variance in a knowledge test, but not the other way around	Weak; testing whether parameters add additional fit, but not whether correlations are increasing. However, directionality is interesting
Ghisletta (2005)	Speed Knowledge	L	377	79.5 – 84.5 (5 yrs, 5 T's)	Variance in explained by speed in knowledge test	With increasing age, changes in a processing speed test explains increasing more variance in a knowledge test, but not the other way around	Weak; testing whether parameters add additional fit, but not whether correlations are increasing. However, directionality is interesting
Hale (2011)	Memory	C	388	20 – 89	Configural invariance/metric invariance	The same model was found to fit in old and young groups; comparisons made with CFI found that factor loadings and intercorrelations could be constrained without loss of fit.	Weak; relations between variables can be constrained to be the same without a statistically significant loss in fit. However, no information about direction

Hedden (2006)	Memory Mental Transformation Speed	C	121	63-82	Configural invariance	A two-factor solution was found to provide the best fit an older sample, whereas an earlier sample found the best fit for younger adults was a three factor fit (Miyake et al., 2000)	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age
Hertzog (1989)	Speed Knowledge Mental Transformation	C	833	College Students, 43 – 89	Correlation magnitude across groups	Trend of increasing correlations across age groups between speed factor and verbal meaning factor. Also, no interaction between speed and age predicting verbal meaning. Finally, patterns not found for other cognitive factors	Weak; patterns only found between one pair of factors, statistical tests of the strength of the magnitude increase not run; also speed*age predicting verbal meaning not significant.
Hertzog (2001)	Speed Mental Transformation Knowledge	L	833	43 – 78	Correlation magnitude across groups	Higher correlations found with increasing age which were eliminated by controlling for speed factor (answer sheet speed)	Weak; testing whether parameters add additional fit, but not whether correlations are increasing. However, directionality is interesting
Hildebrandt (2011)	Memory Speed Mental Transformation	C	448	18 – 82	Configural and Metric Invariance	The same model fit across both samples, and the loadings could be constrained without a significant loss of fit. (The intercepts could not be restricted to be the same without loss of fit.)	Strong; relations between variables can be constrained to be the same without a statistically significant loss in fit.

Hull (2008)	Mental Transformation Knowledge	C	100	51 – 74	Configural Invariance	The model used in younger adults did not fit, but the pattern was not consistent with that of dedifferentiation (fewer factors, but also weaker correlations)	Weak; change in configural invariance consistent with dedifferentiation, but weakened correlations were not
Hulur (2015)	Knowledge Mental Transformation	L	419	22 – 84 (up to 49 yrs)	Correlations magnitude over time*	After removing mean trends in decline, cognitive abilities became more coupled over time	Weak; correspondence to dedifferentiation as measured by increased correlations unclear
Johnson (2010)	Memory	C	95000	18 – 90	Configural Invariance; Metric Invariance	The same single factor model fit across age groups, but the tests had a lot of unexplained variance. In addition, although there were differences in the loadings to the model, they did not consistently reflect dedifferentiation	Strong; metric invariance failure but not consistent with dedifferentiation
Juan-Espinosa (2002)	Memory Mental Transformation Knowledge Speed	C	1369	16 – 94	Variance accounted for by g and by four group factors	There were no significant changes in the variance accounted for by and by the four group factors in the tests in the WAIS-III Spanish norming sample	Strong; specific statistical tests of g/factors variance reveal no differences, but age groups relatively young (see age range)
Li (2004)	Memory Knowledge Speed	C	356	Adult sample 18 – 89 (four	Correlation magnitudes; configural	Larger correlations/more g variance in late adulthood and old age vs young and	Strong; both higher correlations found and change in configural

	Mental Transformation			groups)	invariance; g variance	middle adulthood; PCA revealed 5 components for 2 younger groups VS 2 in the 2 older groups	structure consistent with dedifferentiation.
Lindenberger (1997)	Perceptual Speed Mental Transformation Memory Knowledge	C	516	70 – 103	Correlation magnitude; configural invariance	Correlations higher in older adulthood; older adult data can be well described by a single 2nd higher order factor (but are still differentiated at 1st level)	Strong; both higher correlations found and change in configural structure consistent with dedifferentiation.
Nyberg (2003)	Memory Knowledge	C	925	35 – 50, 55 – 65, 70 – 80	Correlation magnitude	Correlations higher with age, but differences were not significant	Strong; no significant differences despite higher magnitudes
Park (2002)	Memory Speed Sensory Knowledge	C	345	30 – 92	Configural invariance; model comparisons	A differentiated model fit better for both young and older adults; but the model was not exactly the same	Weak; correlations could still be higher even if configural invariance is acheived
Schaie (1998)	Knowledge Speed Mental Transformation Memory	C/L	1998	32,46,53,60,67,76 (7 yrs)	Configural invariance; metric invariance	Configural invariance was found across all groups. Weak factorial invariance was found across time for all cohorts and could be accepted across all groups except the oldest and youngest	Weak; Overall, evidence is in favor with metric invariance. And even though there is a difference, it appears to be due to only one measure and only found between the extreme groups
Schaie	Speed	C	1621	22 – 95	Configural	Configural invariance is	Weak; Failure of fit and

(1989)	Mental Transformation Knowledge				invariance; metric invariance	achieved, but neither factor correlations nor loadings could be constrained to be invariant	reference to increases in covariance across age groups; but also in variance up to 60's. Also large differences in sample sizes between groups
Schultz (1980)	Mental Transformation Knowledge	C	100, 100	19.54, 63.99	Configural invariance; correlation magnitude	The same model did not fit across groups; all but one correlation was larger with age.	Weak; failure of model fit attributed to higher correlations; but no test of significant increase in magnitude
Sims (2009)	Memory Speed Mental Transformation Knowledge	C	512	50 – 79 (3 groups, 10 yrs)	Configural invariance; metric invariance	Constraining the loadings and the factor covariances to be equal across the 3 groups did not reduce the fit compared with the unconstrained model.	Strong; constraining loadings resulted in the no statistically significant loss of fit.
Singer (2003)	Speed Memory Knowledge Mental Transformation	C/L	132	70-103 (7 yrs)	Directionality dedifferentiation	Longitudinally, whereas fluid abilities declined, crystallized abilities were stable	Weak; testing for changes in direction, but no increases in correlation magnitude
Tucker-Drob (2008)	Speed Knowledge Memory Mental Transformation	C/L	1281	18-95 + (7 yrs)	Variance in change explained by g across groups	g was not responsible for a greater amount of change seen in older groups versus younger adults	Strong; large age range, multiple measures of cognitive abilities; multiple age groups; 7 years over time; specifically testing for changes in correlation

Tucker-Drob (2009)	Knowledge Mental Transformation Speed Memory	C	6273	4 – 101	Variance explained by g across groups	*g* was not responsible for a greater amount of change seen in older groups versus younger adults; in fact, in older adulthood patterns suggested greater differentiation	Strong; large age range, multiple measures of cognitive abilities; lifespan sample; specifically testing for changes in correlation
Vaughan (2010)	Mental Transformation	C	95	60 – 90	Configural Invariance, correlation magnitude	A three factor model of executive function fit in older adults as it does in younger adults (Miyake et al., 2001), and correlations between factors were actually tended to be weaker with age	Weak; comparison age groups were not included, and therefore differences could be due to differences in measures/analytic techniques etc., not age
Waller (1990)	Knowledge Mental Transformation	C	1880	16-74	Configural invariance	A three-factor model fit 8/9 age groups best. In the oldest group however, a two-factor model provided a better fit	Strong; statistical loss of fit when old group fit with 3 factor model versus 2 factor model
Ward (2000)	Knowledge Mental Transformation	C	1880	16-74	Configural invariance	One of the tests loaded onto a different factor in younger versus older adults. Modeling a fourth improved fit for all age groups except the oldest.	Weak; Differences in configural invariance could be consistent with dedifferentiation; however other changes (like what tests load onto what factors) may confound this.
Zelinski (1998)	Memory Knowledge	L	82	55 – 81, + (16 yrs)	Comparing predictors of change	Although two measures shared some predictors of change, they were not	Weak; not directly testing for dedifferentiation, but whether certain factors

						consistent across measures- suggesting there is not a unitary source of change nor are the mechanisms of change completely differentiated.	explained age differences comparable across measures.
Zelinski (2003)	Memory Speed Mental Transformati on Knowledge	C/L	613 (289 by T3)	30 - 97	Configural invariance, metric invariance	Cross-sectionally the age groups had configural invariance, but only partial invariance. However, longitudinally, there was evidence of configural and metric invariance	Strong; large age range, multiple measures of cognitive abilities; lifespan sample; specifically testing for changes in correlation

Table S2. Data sources included in the meta-analysis, with their age ranges, cognitive variables, etc.

Name	Collection Date	# Constructs	Tests Included	Total N	Age Groups	Source
WAIS III	1997	14	Vocabulary Arithmetic Block Design Comprehension Digit Span Digit Symbol Information Letter-Number Sequencing Matrix Reasoning Object Assembly Picture Arrangement Picture Completion Similarities Symbol Search	2450 12 groups: all 200 except 80-84 = 150 85-89 = 100	18-89 *12 groups:* 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-89	WAIS III & WMS III Technical Manual pgs. 218-230
WAIS IV	2008	15	Block Design Arithmatic Cancellation Coding Comprehension Digit Span Figure Weight Information Letter-Number Sequencing Matric Reasoning Picture Completion Similarities Symbol Search Visual Puzzles	2200 12 groups: youngest 9 N= 200 oldest 4 groups N= 100	18-90 *12 groups:* 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-90	WAIS IV Technical Manual: pgs. 138-150

			Vocabulary			
WMS III	1997	11	Faces 1,2 Family Pictures 1,2 LM Recall 1,2 LN 1 Sequence Spatial Span VPA 1,2 Recall Auditory Recognition	1250 13 groups: all 100 except 80-84 = 75 85-89 = 75	18-89 *12 groups:* 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-89	WAIS III & WMS III Technical Manual pgs. 231-243
WMS IV	2009	14	Logical Memory 1,2 Verbal Paired Associates 1,2 Designs 1,2 Visual Reproduction 1,2 Spatial Addition Symbol Span Logos 1,2 Names 1,2	1400 100 in each are band (double collection of 65-69)	18-90 *12 groups:* 18-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-90	WMS – IV Technical Manual pgs. 184-197.
KAIT	1993	10	Auditory Comprehension Auditory Delayed Recall Definitions Double Meanings Famous Faces Logical Steps Memory for Block Designs Mystery Codes	2000 (500 children/ 1500 adults)	17-85+ 10 groups: 17-19, 20-24, 25-34, 35-44, 45-54, 55-59, 60-64, 65-69, 70-74, 75-85+	Kaufman Adolescent & Adult Intelligence Test pgs 130-136

			Rebus Delayed Recall Rebus Learning			
WJ-R	1990	21	Memory for Names Memory for Sentences Visual Matching Incomplete Words Visual Closure Picture Vocabulary Analysis-Synthesis Visual-Auditory Learning Memory for Words Cross Out Sound Blending Picture Recognition Oral Vocabulary Concept Formation Delayed Recall-Memory for Names Delayed Recall-Visual-Auditory Learning Numbers Reversed Sound Patterns Spatial Relations Listening Comprehension Verbal Analogies	2669 144, 208, 316, 312, 254, 1074, 184, 177	18-79 4 groups: 18, 30-39, 50-59, 70-79	Woodcock Johnson II: pgs. 307-328
WMS-R	1987	11	Mental Control Figural Memory Logical Memory 1	316 *5 groups:* 53, 50, 54, 54, 55,	20-74 *5 groups:* 20-24, 35-44,	Wechsler Memory Scale-Revised pgs.

			Visual Paired Associates 1 Verbal Paired Associates 1 Visual Reproduction 1 Digit Span Visual Memory Span Logical Memory 2 Visual Paired Associates 2 Verbal Paired Associates 2 Visual Reproduction 2	50	55-64, 65-69, 70-74	144-150
WASI	1999	4	Vocabulary (gc) Block Design (gf) Similarities (gf) Matrices (gf)	1,245: *each groups:* N = 100 except 75-79; 80-84 = 85 85-89 = 75.	17-89: *12 groups:* 17-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, 85-89	WASI manual: pgs. 203-210
WAIS	1955	11	Information Comprehension Arithmetic Similarities Digit Span Vocabulary Digit Symbol Picture Completion Block Design Picture Arrangement	1700: *3 groups:* 200, 300, 300	18-54: *3 groups*: 18-19, 25-34, 45-54	WAIS manual: pgs. 15-17

			Object Assembly			
WAIS R	1981	11	Information Comprehension Arithmetic Similarities Digit Span Vocabulary Digit Symbol Picture Completion Block Design Picture Arrangement Object Assembly	1880: 9 groups: 200, 200, 200, 300, 250, 250, 160, 160, 160	18-74 *8 groups:* 18-19, 20-24, 25-34, 35-44, 45-54, 55-64, 65-69, 70-74	WAIS-R manual: pgs. 37-45
KBIT	1990	2	Vocabulary Matrices	2022: *4 groups:* 113, 109, 105, 119, 123, 122, 116, 207, 181, 148, 179, 213, 172, 115	17-90: *4 groups:* 17-19, 20-34, 35-54, 55-90	Kaufman Brief Intelligence Test: pg. 58

Table S3. List of tests by battery, name, description, and the type.

Test Battery	Test	Description	Type/ies
KAIT	**Auditory Comprehension**	**Listening to a recording of (or examiner reading aloud) a news story, then answering literal and inferential questions about the story.**	**GK, MT**
KAIT	Auditory Delayed Recall	Answering literal and inferential questions about news stories that were heard during administration of Auditory Comprehension.	M
KAIT	Definitions	Integrating two types of clues-a word with some of its letters missing and an oral clue — about the word's meaning-to identify the word.	GK
KAIT	Double Meanings	Studying two sets of word clues, then thinking of a word with two different meanings that fits both sets of clues.	GK
KAIT	Famous Faces	Naming people of current or historical fame, based on their photographs and a verbal clue. (Also serves as an alternate subtest for the Core Battery Crystallized Scale.)	GK
KAIT	Logical Steps	Attending to logical premises presented both visually and orally, using these to solve a problem.	MT
KAIT	Memory for Block Designs	Studying a printed design that is briefly exposed, then constructing the design using six cubes and a form board. (Also serves as an alternate subtest for the Core Battery Fluid Scale.)	MT
KAIT	Mystery Codes	Cracking a code that is used to identify a set of pictures, and then applying this code to a new set of pictures.	MT
KAIT	Rebus Learning	Learning the word or concept that is represented by a rebus (that is, a picture that stands for a word), and then 'reading' phrases and sentences composed of these rebuses.	M
KAIT	Rebus Delayed Recall	'Reading' phrases and sentences composed of	M

		rebuses that were learned earlier during the administration of Rebus Learning. (see above)	
KBIT	Vocabulary	The examiner says a vocabulary word, and the examinee points to the picture that illustrates the word.	GK
KBIT	Matrices	Finding a relationship or rule in a set of pictures or patterns, and pointing to the picture or pattern that best fits the relationship or rule.	MT
WASI	Vocabulary	For picture items, the examinee names the object presented visually. For verbal items, examinee defines words presented visually and orally.	GK
WASI	Block Design	This subtest consists of two-dimensional designs which the client tries to copy using three dimensional blocks.	MT
WASI	**Similarities**	**The subtest consists of 18 pairs of words. The client is asked to identify the qualitative relationship between the two words.**	**GK, MT**
WASI	Matrices	This is a nonverbal reasoning task in which individuals are asked to identify patterns in designs.	MT
WAIS	Information	Examinee answers questions that address a broad range of general knowledge topics.	GK
WAIS	Comprehension	Examinee answers questions based on his/her understanding of general principles and social situations.	GK
WAIS	**Arithmetic**	**Working within a specified time limit, the examinee mentally solves a series of arithmetic problems.**	**MT, GK**
WAIS	**Similarities**	**The subtest consists of 18 pairs of words. The client is asked to identify the qualitative relationship between the two words.**	**GK, MT**
WAIS	Digit Span	Digit Span Forward (individual tries to repeat digits forward). Digit Span Backward (individual tries to	MT

		repeat digits backward). Digit Span Sequencing (individual tries to repeat digits in ascending order)	
WAIS	Vocabulary	For picture items, the examinee names the object presented visually. For verbal items, examinee defines words presented visually and orally.	GK
WAIS	Digit Symbol	Examinee goes through a grid of numbers and places the correct symbol above each number.	S
WAIS	**Picture Completion**	**Working within a specified time limit, the examinee views a picture with an important part missing and identifies the missing part.**	**MT, GK**
WAIS	Block Design	This subtest consists of two-dimensional designs which the client tries to copy using three dimensional blocks.	MT
WAIS	Picture Arrangement	Eleven items. Each item consists of 3 to 6 cards containing pictures. The examinee must arrange the pictures from left to right to tell the intended story.	MT
WAIS	Object Assembly	Four items, each item being a "cut up" object, like a puzzle. Examinee must correctly assemble the parts of the puzzle.	MT
WAIS-R	Information	Examinee answers questions that address a broad range of general knowledge topics.	GK
WAIS-R	Comprehension	Examinee answers questions based on his/her understanding of general principles and social situations.	GK
WAIS-R	**Arithmetic**	**Working within a specified time limit, the examinee mentally solves a series of arithmetic problems.**	**MT, GK**
WAIS-R	**Similarities**	**The subtest consists of 18 pairs of words. The client is asked to identify the qualitative relationship between the two words.**	**GK, MT**
WAIS-R	Digit Span	Digit Span Forward (individual tries to repeat digits forward). Digit Span Backward (individual tries to	MT

		repeat digits backward). Digit Span Sequencing (individual tries to repeat digits in ascending order)	
WAIS-R	Vocabulary	For picture items, the examinee names the object presented visually. For verbal items, examinee defines words presented visually and orally.	GK
WAIS-R	Digit Symbol	Examinee goes through a grid of numbers and places the correct symbol above each number.	S
WAIS-R	**Picture Completion**	**Working within a specified time limit, the examinee views a picture with an important part missing and identifies the missing part.**	**MT, GK**
WAIS-R	Block Design	This subtest consists of two-dimensional designs which the client tries to copy using three dimensional blocks.	MT
WAIS-R	Picture Arrangement	Eleven items. Each item consists of 3 to 6 cards containing pictures. The examinee must arrange the pictures from left to right to tell the intended story.	MT
WAIS-R	Object Assembly	Four items, each item being a "cut up" object, like a puzzle. Examinee must correctly assemble the parts of the puzzle.	MT
WAIS III	Block Design	This subtest consists of two-dimensional designs which the client tries to copy using three dimensional blocks.	MT
WAIS III	Comprehension	Examinee answers questions based on his/her understanding of general principles and social situations.	GK
WAIS III	Digit Span	Digit Span Forward (individual tries to repeat digits forward). Digit Span Backward (individual tries to repeat digits backward). Digit Span Sequencing (individual tries to repeat digits in ascending order)	MT
WAIS III	Digit Symbol	Examinee goes through a grid of numbers and places the correct symbol above each number.	S
WAIS III	Vocabulary	For picture items, the examinee names the object	GC

		presented visually. For verbal items, examinee defines words presented visually and orally.	
WAIS III	**Arithmetic**	**Working within a specified time limit, the examinee mentally solves a series of arithmetic problems.**	**MT, GK**
WAIS III	Information	Examinee answers questions that address a broad range of general knowledge topics.	GK
WAIS III	Letter-Number Sequencing	The examinee is read a sequence of numbers and letters and recalls the numbers in ascending order and the letters in alphabetical order.	MT
WAIS III	Matrix Reasoning	This is a nonverbal reasoning task in which individuals are asked to identify patterns in designs.	MT
WAIS III	Object Assembly	Four items, each item being a "cut up" object, like a puzzle. Examinee must correctly assemble the parts of the puzzle.	MT
WAIS III	Picture Arrangement	Eleven items. Each item consists of 3 to 6 cards containing pictures. The examinee must arrange the pictures from left to right to tell the intended story.	MT
WAIS III	**Picture Completion**	**Working within a specified time limit, the examinee views a picture with an important part missing and identifies the missing part.**	**MT, GK**
WAIS III	**Similarities**	**The subtest consists of 18 pairs of words. The client is asked to identify the qualitative relationship between the two words.**	**GK, MT**
WAIS III	Symbol Search	The client, under time pressure, scans a search group and indicates whether one of the symbols in the target group matches.	S
WAIS IV	Block Design	This subtest consists of two-dimensional designs which the client tries to copy using three dimensional blocks.	MT
WAIS IV	**Arithmetic**	**Working within a specified time limit, the examinee mentally solves a series of arithmetic problems.**	**MT, GK**
WAIS IV	Cancellation	Working within a specified time limit, the examinee scans a structured arrangement of shapes and marks target shapes.	S

WAIS IV	Coding	In this subtest individuals are asked to record associations between different symbols and numbers within time limits.	S
WAIS IV	Comprehension	Examinee answers questions based on his/her understanding of general principles and social situations.	GK
WAIS IV	Digit Span	Digit Span Forward (individual tries to repeat digits forward). Digit Span Backward (individual tries to repeat digits backward). Digit Span Sequencing (individual tries to repeat digits in ascending order)	MT
WAIS IV	Figure Weights	Working within a specified time limit, the examinee views a scale with missing weight(s) and selects the response option that keeps the scale balanced.	MT
WAIS IV	Information	Examinee answers questions that address a broad range of general knowledge topics.	GK
WAIS IV	Letter-Number Sequencing	The examinee is read a sequence of numbers and letters and recalls the numbers in ascending order and the letters in alphabetical order.	MT
WAIS IV	Matrix Reasoning	This is a nonverbal reasoning task in which individuals are asked to identify patterns in designs.	MT
WAIS IV	**Picture Completion**	**Working within a specified time limit, the examinee views a picture with an important part missing and identifies the missing part.**	**MT, GK**
WAIS IV	**Similarities**	**The subtest consists of 18 pairs of words. The client is asked to identify the qualitative relationship between the two words.**	**GK, MT**
WAIS IV	Symbol Search	The client, under time pressure, scans a search group and indicates whether one of the symbols in the target group matches.	S
WAIS IV	Visual Puzzles	In this subtest individuals view a completed puzzle and then select three response options that can be combined to reconstruct the puzzle.	MT
WAIS IV	Vocabulary	For picture items, the examinee names the object presented visually. For verbal items, examinee defines words presented visually and orally.	GK
WMS-R	Figural Memory	The examinee looks briefly at abstract designs, and is then asked to identify them from a larger group of designs.	M
WMS-R	Logical Memory 1	The examinee is asked to retell a story from memory immediately after hearing it.	M

WMS-R	Verbal Paired Associates 1	This subtest assesses verbal memory for associated word pairs after seeing them.	M
WMS-R	Visual Reproduction 1	The examinee looks at a geometric design and is then asked to draw it from memory.	M
WMS-R	Digit Span	Digit Span Forward (individual tries to repeat digits forward). Digit Span Backward (individual tries to repeat digits backward). Digit Span Sequencing (individual tries to repeat digits in ascending order)	MT
WMS-R	Visual Memory Span	The two parts of the visual memory span subtest, tapping forward and tapping backward. Tapping forward the examinee watches the examiner touch the red squares on a card and are then asked to repeat the sequences. For tapping backward, the examinee watches the examiner touch the green squares on a card and are then asked to repeat the performance in reverse.	MT
WMS-R	Logical Memory 2	The delayed condition assesses long-term narrative memory with free recall and recognition tasks	M
WMS-R	Visual Paired Associates 2	The delayed condition assesses long-term recall for visually paired information with cued recall and recognition tasks, and includes a free recall task	M
WMS-R	Verbal Paired Associates 2	The delayed condition assesses long-term recall for verbally paired information with cued recall and recognition tasks, and includes a free recall task	M
WMS-R	Visual Reproduction 2	The examinee looks at a geometric design and is then asked to draw it from memory after a delay.	M
WMS III	Faces 1,2	The examinee is presented photographs of 24 target faces. The examinee is then presented photographs of 48 faces, included 24 faces and 24 new faces. The examinee must identify each face as either a target face of a new one. For faces 2, the examinee must again identify each face as a target or new face.	M

WMS III	Family Pictures 1,2	A family portrait and four subsequent scenes involving the family characters and family dog are shown in to the examinee. The examinee is asked to identify who was in each scene and each character's activity and location. Family pictures 2 requires the examinee to recall the same information after a delay.	M
WMS III	LM Recall 1,2	The examinee is asked to retell a story from memory immediately after hearing it. The delayed condition assesses long-term narrative memory with free recall and recognition tasks	M
WMS III	LN 1 Sequence	The examinee is read a sequence of numbers and letters and recalls the numbers in ascending order and the letters in alphabetical order.	MT
WMS III	Spatial Span	Spatial span is a visual analogue of the Digit Span subtest.	MT
WMS III	VPA 1,2 Recall	This subtest assesses verbal memory for associated word pairs The delayed condition assesses long-term recall for verbally paired information with cued recall and recognition tasks, and includes a free recall task.	M
WMS IV	Logical Memory 1,2	The examinee is asked to retell a story from memory immediately after hearing it. The delayed condition assesses long-term narrative memory with free recall and recognition tasks	M
WMS IV	Verbal Paired Associates 1,2	This subtest assesses verbal memory for associated word pairs The delayed condition assesses long-term recall for verbally paired information with cued recall and recognition tasks, and includes a free recall task	M
WMS IV	Designs 1,2	This subtest assesses spatial memory for unfamiliar visual material. The delayed condition assesses long- term spatial and visual	M

		memory with free recall and recognition tasks.	
WMS IV	Visual Reproduction 1,2	This subtest assesses memory for nonverbal visual stimuli The delayed condition assesses long-term visual- spatial memory with free recall and recognition tasks, and includes a direct copy task	M
WMS IV	Spatial Addition	Assesses visual-spatial working memory using a visual addition task.	MT
WMS IV	Symbol Span	The examinee is briefly shown a series of abstract symbols on a page and then asked to select the symbols from an array of symbols, in the same order they were presented on the previous page	MT
WJ-R	Memory for Names	Participants must remember an increasingly large number of names of novel cartoon characters	M
WJ-R	**Memory for Sentences**	**Participants must repeat complete sentences, using sentence meaning to aid recall.**	**MT, GK**
WJ-R	Visual Matching	Participants must quickly find and circle two identical numbers in a row of six numbers in 3 minutes.	S
WJ-R	Picture Vocabulary	Participants must name familiar and unfamiliar pictured objects	GK
WJ-R	Analysis-Synthesis	Participants must analyze the components of an incomplete logic puzzle and to determine and name the missing components.	MT
WJ-R	Visual-Auditory Learning	Participants must associate new visual symbols with familiar words and to translate a series of sentences into a controlled learning situation.	M
WJ-R	Memory for Words	Participants must repeat lists of unrelated words in the correct order	MT
WJ-R	Cross Out	Participants must mark drawings that are identical to the first drawing in the row in 3 mins.	S
WJ-R	Oral Vocabulary	Participants hear words and then must state either a synonym or antonym to those words.	GK
WJ-R	Concept Formation	Participants must identify rules that make up	MT

		geometric figures after being exposed to concepts.	
WJ-R	Delayed Recall- Memory for Names	Participants must recall and relearn (after a 30-minute to 8-day delay) names of novel cartoon	M
WJ-R	Delayed Recall- Visual-Auditory Learning	Participants are asked to name the words associated with the symbols learned during visual auditory learning 1-8 after original test.	M
WJ-R	Numbers Reversed	Participants must repeat a series of random numbers backward	MT
WJ-R	Spatial Relations	Participants must select the component parts of whole shape	MT
WJ-R	**Listening Comprehension**	**Participants must supply a single word missing at the end of pre-recorded passage.**	**GK, MT**
WJ-R	**Verbal Analogies**	**Participants must complete phrases with words that indicate appropriate analogies**	**MT, GK**

Notes: Tests that load on multiple constructs based on previous reports are in bold.

Table S4. Number of dependent and independent samples (K) in each of the individual correlation types.

	Total K	**Independent K**
Total Sample	**3174 (2355)**	**68 (68)**
Speed with Memory	**28 (28)**	**4 (4)**
Speed with Mental Transformation	**554 (416)**	**39 (39)**
Speed with Knowledge	**292 (217)**	**39 (39)**
Knowledge with Memory	**162 (118)**	**12 (12)**
Knowledge with Mental Transformation	**1256 (722)**	**63 (63)**
Mental Transformation with Memory	**882 (854)**	**37 (37)**

Notes: K's for unconstrained (constrained) overall/independent K's for the total dataset and the individual types. There are overlaps between the independent K's in each of the groups. Total K's represent the number of correlations included in each type. Independent K's represent the number of separate samples from which these correlations were drawn from.

Table S5. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the S-MT type and then compared to the five other groups (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-MT)	0.4758	0.0134	0.4496	0.5019	35.61	**<.0001**
Age (S-MT)	0.0024	0.0006	0.0012	0.0035	4.00	**0.0001**
Age^2 (S-MT)	0.0089	0.0155	-0.0215	0.0392	0.57	0.5664
Deviations from						
Intercept (K-M)	0.0631	0.0106	0.0424	0.0839	5.95	**<.0001**
Intercept (S-M)	-0.0612	0.0188	-0.098	-0.0244	-3.26	**0.0011**
Intercept (K-MT)	0.0752	0.0047	0.0659	0.0845	15.91	**<.0001**
Intercept (MT-M)	-0.0505	0.0057	-0.0617	-0.0394	-8.88	**<.0001**
Intercept (S-K)	-0.008	0.0062	-0.0202	0.0042	-1.28	0.2008
Age (K-M)	-0.003	0.0005	-0.004	-0.0019	-5.66	**<.0001**
Age (S-M)	-0.0047	0.0009	-0.0065	-0.0029	-5.11	**<.0001**
Age (K-MT)	-0.0021	0.0002	-0.0025	-0.0017	-9.43	**<.0001**
Age (MT-M)	-0.0032	0.0003	-0.0038	-0.0027	-11.94	**<.0001**
Age (S-K)	-0.0006	0.0003	-0.0012	-0.0001	-2.19	**0.0287**
Age^2 (K-M)	-0.0507	0.0138	-0.0777	-0.0237	-3.68	**0.0002**
Age^2 (S-M)	-0.0574	0.0239	-0.1042	-0.0106	-2.4	**0.0162**
Age^2 (K-MT)	-0.0205	0.006	-0.0321	-0.0088	-3.43	**0.0006**
Age^2 (MT-M)	-0.0362	0.0072	-0.0503	-0.0220	-5.00	**<.0001**
Age^2 (S-K)	0.0026	0.0078	-0.0127	0.0179	0.33	0.7402

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table S6. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the S-K type and then compared to the four other groups (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-K)	0.4679	0.0139	0.4408	0.4951	33.75	**<.0001**
Age (S-K)	0.0018	0.0006	0.0006	0.0030	2.87	**0.0041**
Age^2 (S-K)	0.0119	0.0161	-0.0198	0.0435	0.74	0.4620
Deviations from						
Intercept (K-M)	0.0664	0.0113	0.0443	0.0885	5.90	**<.0001**
Intercept (S-M)	-0.0721	0.0196	-0.1105	-0.0337	-3.68	**0.0002**
Intercept (K-MT)	0.0841	0.0059	0.0726	0.0956	14.34	**<.0001**
Intercept (MT-M)	-0.0439	0.0068	-0.0572	-0.0306	-6.48	**<.0001**
Age (K-M)	-0.0024	0.0006	-0.0035	-0.0014	-4.41	**<.0001**
Age (S-M)	-0.0047	0.0010	-0.0066	-0.0028	-4.88	**<.0001**
Age (K-MT)	-0.0015	0.0003	-0.0020	-0.0009	-5.34	**<.0001**
Age (MT-M)	-0.0026	0.0003	-0.0033	-0.0020	-8.12	**<.0001**
Age^2 (K-M)	-0.059	0.0146	-0.0876	-0.0304	-4.04	**0.0001**
Age^2 (S-M)	-0.0799	0.0249	-0.1287	-0.0311	-3.21	**0.0013**
Age^2 (K-MT)	-0.0227	0.0074	-0.0371	-0.0083	-3.09	**0.0020**
Age^2 (MT-M)	-0.0410	0.0086	-0.0578	-0.0241	-4.76	**<.0001**

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table S7. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the MT-M type and then compared to the three other groups (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (MT-M)	0.4179	0.0119	0.3945	0.4413	35.00	**<.0001**
Age (MT-M)	-0.0008	0.0005	-0.0019	0.0002	-1.57	0.1153
Age^2 (MT-M)	-0.0302	0.0142	-0.0581	-0.0024	-2.13	**0.0334**
Deviations from						
Intercept (K-M)	0.1414	0.0089	0.1238	0.1589	15.81	**<.0001**
Intercept (S-M)	-0.0151	0.0187	-0.0518	0.0215	-0.81	0.4184
Intercept (K-MT)	0.1526	0.0048	0.1431	0.1621	31.61	**<.0001**
Age (K-M)	0.0010	0.0004	0.0001	0.0019	2.27	**0.0235**
Age (S-M)	-0.0007	0.0009	-0.0025	0.0011	-0.76	0.4446
Age (K-MT)	0.0013	0.0002	0.0008	0.0018	5.63	**<.0001**
Age^2 (K-M)	-0.021	0.0118	-0.0441	0.0021	-1.78	0.0755
Age^2 (S-M)	-0.038	0.0241	-0.0851	0.0092	-1.58	0.1144
Age^2 (K-MT)	0.0236	0.0063	0.0113	0.0359	3.77	0**.0002**

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table S8. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in the K-MT type and then compared to the two other groups (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (K-MT)	0.5536	0.0112	0.5316	0.5755	49.43	**<.0001**
Age (K-MT)	0.0003	0.0005	-0.0007	0.0013	0.58	0.5622
Age^2 (K-MT)	-0.016	0.013	-0.0416	0.0096	-1.22	0.2206
Deviations from						
Intercept (K-M)	0.0598	0.0111	0.0381	0.0816	5.39	**<.0001**
Intercept (S-M)	-0.0898	0.0208	-0.1306	-0.0489	-4.31	**<.0001**
Age (K-M)	0.0002	0.0005	-0.0009	0.0012	0.28	0.7782
Age (S-M)	-0.0023	0.001	-0.0044	-0.0003	-2.28	**0.0225**
Age^2 (K-M)	-0.0403	0.0145	-0.0687	-0.0119	-2.79	**0.0053**
Age^2 (S-M)	-0.0700	0.0266	-0.1222	-0.0178	-2.63	**0.0086**

Notes: Items in bold are significant at p < .05. Random effects for sample were included.

Table S9. Estimates of the correlation meta-analysis where the average age, type of correlation, age^2 predict correlation size in S-M type and then compared to K-M (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
Intercept (S-M)	0.4598	0.0321	0.3968	0.5228	14.31	**<.0001**
Age (S-M)	-0.0008	0.0015	-0.0038	0.0021	-0.55	0.5845
Age^2 (S-M)	-0.0334	0.0403	-0.1123	0.0456	-0.83	0.4075
Deviations from						
Intercept (K-M)	0.0991	0.025	0.0501	0.1481	3.96	**0.0001**
Age (K-M)	0.0018	0.0012	-0.0006	0.0042	1.47	0.1405
Age^2 (K-M)	0.0107	0.0319	-0.0519	0.0733	0.33	0.7383

Notes: Items in bold are significant at $p < .05$. Random effects for sample were included.

Table S10. Estimates of the correlation meta-analysis where the average age predicts correlation size for each of types (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
S-M (28)						
intercept	0.4179	0.0396	0.3403	0.4954	10.56	**<.0001**
age	0.0021	0.0019	-0.0016	0.0057	1.11	0.2679
S-MT (416)						
intercept	0.4458	0.0121	0.4222	0.4695	36.90	**<.0001**
age	0.0022	0.0005	0.0012	0.0033	4.10	**<.0001**
S-K (217)						
intercept	0.4558	0.0210	0.4146	0.4971	21.66	**<.0001**
age	0.0015	0.0009	-0.0004	0.0033	1.57	0.1175
K-M (118)						
intercept	0.5584	0.0257	0.5080	0.6088	21.72	**<.0001**
age	0.0009	0.0011	-0.0012	0.0030	0.82	0.4106
K-MT (722)						
intercept	0.5530	0.0114	0.5306	0.5754	48.41	**<.0001**
age	0.0001	0.0005	-0.0009	0.0011	0.16	0.8714
MT-M (854)						
intercept	0.3841	0.0204	0.3441	0.4242	18.81	**<.0001**
age	-0.0008	0.0009	-0.0025	0.0010	-0.84	0.3962

Notes: Items in bold are significant at $p < .05$. K's are in parentheses. Used the unconstrained data. Random effects for sample were included.

Table S11. Estimates of the correlation meta-analysis where the average age predicts correlation size for each of types (constrained analyses).

	Unstandardized Estimate	Standard Error	Lower Bound	Upper Bound	Z	p
S-M (28)						
intercept	0.4587	0.0763	0.3091	0.6084	6.01	<.0001
age	0.0018	0.0023	-0.0027	0.0062	0.78	0.4377
age^2	-0.0428	0.0630	-0.1663	0.0807	-0.68	0.4972
S-MT (416)						
intercept	0.4485	0.0198	0.4097	0.4872	22.67	<.0001
age	0.0022	0.0006	0.0011	0.0034	3.93	<.0001
age^2	-0.0024	0.0143	-0.0304	0.0256	-0.17	0.8667
S-K (217)						
intercept	0.4662	0.0344	0.3987	0.5337	13.54	<.0001
age	0.0016	0.0010	-0.0004	0.0035	1.59	0.1116
age^2	-0.0094	0.0248	-0.0580	0.0391	-0.38	0.7032
K-M (118)						
intercept	0.5811	0.0396	0.5035	0.6588	14.67	<.0001
age	0.0011	0.0011	-0.0011	0.0032	0.95	0.3400
age^2	-0.0223	0.0298	-0.0806	0.0361	-0.75	0.4545
K-MT (722)						
intercept	0.5746	0.0183	0.5388	0.6103	31.47	<.0001
age	0.0003	0.0005	-0.0007	0.0013	0.60	0.5483
age^2	-0.0199	0.0131	-0.0456	0.0058	-1.52	0.1284
MT-M (854)						
intercept	0.4125	0.0319	0.3500	0.4750	12.94	<.0001
age	-0.0005	0.0009	-0.0023	0.0014	-0.51	0.6124
age^2	-0.0281	0.0242	-0.0756	0.0194	-1.16	0.2458

Notes: Items in bold are significant at p < .05. K's are in parentheses. Used the unconstrained data. Random effects for sample were included.

Table S12. List of tests by name, description, and the category.

Test	Description	Type
MMSE	30-point questionnaire that is used extensively in clinical and research settings to measure pathological cognitive impairment.	-
AVLT	This is a test of episodic memory that assesses the ability to acquire 15 words across five immediate learning trials, to recall the words immediately after an intervening interference list, and to recall and recognize the words after a 30-minute delay interval.	M
Delayed Recall (ADAS)	Ask the subject to recall as many words as possible from the words presented in the Immediate Word Recall task. ADAS tests were scored so that higher scores represent poorer performance.	M
Recognition (ADAS)	In the learning portion of this test, the subject is given one trial to learn a list of 12 words. They are then shown a list of words that were on the list and were not and asked to identify the words that were on the list. ADAS tests were scored so that higher scores represent poorer performance.	M
Digit Span Backward	This is a widely used measure of working memory (or attention) in which the subject is read number sequences of increasing length and then asked to repeat each sequence backward. The primary measure of performance is the number of digit sequences correctly reversed.	MT
Cog Construction (ADAS)	This test assesses the subject's ability to copy 4 geometric forms. The forms should be presented one at a time. ADAS tests were scored so that higher scores represent poorer performance.	MT
Clock Drawing	Participants are asked to draw the face of a clock showing the numbers and two hands set to ten after eleven.	MT
Trails A	Part A consists of 25 circles numbered 1 through 25 distributed over a white sheet of 81/2" x 11" paper. The subject is instructed to connect the circles with a drawn line as quickly as possible in ascending numerical order. The subject's performance is judged in terms of the time, in seconds, required to complete each Trail. The time to complete Part A (150-second maximum) is used as measure, with longer scores representing poorer performance.	S

Digit Symbol Substitution	Participant fills in as many squares with corresponding symbols as possible, one after the other, without skipping any in 90 seconds.	S
Number Cancellation (ADAS)	Participants are shown two numbers. Participants are then given 45 seconds to go through a list of random numbers and cross out those two numbers whenever they appear. ADAS tests were scored so that higher scores represent poorer performance.	S
Category Fluency	This is a widely used measure of semantic memory (verbal fluency, language). The subject is asked to name different exemplars from a given semantic category. The number of correct unique exemplars named is scored.	GK
Boston Naming	Participants are shown 30 images and are asked to tell the administrator the name of the object that you see.	GK
Naming Test (ADAS)	The subject is asked to name 12 randomly presented real objects. ADAS tests were scored so that higher scores represent poorer performance.	GK

Notes: ADAS = Alzheimer's Disease Assessment Scale. AVLT: Adult Verbal Learning Test.

Table S13. Correlations between MMSE and cognitive tests at T1 for participants with CDR scores less than 1.5.

	1	2	3	*4*	5	6	*7*	8	*9*	10	11	12	13	*14*	*15*	16
1 MMSE	1															
2. AVLT Learning	.36*	1														
3. AVLT Short D.	.37*	.61*	1													
4. Delayed Recall	.47*	.58*	.68*	1												
5. AVLT Long D.	.34*	.56*	.83*	.66*	1											
6. AVLT Recog.	.33*	.45*	.59*	.55*	.58*	1										
7. Word Recogn.	.25*	.37*	.40*	.46*	.40*	.42*	1									
8. DS backward	.25*	.16*	.24*	.17*	.22*	.13*	.15*	1								
9. Construction	.31*	.24*	.29*	.24*	.27*	.25*	.14*	.22*	1							
10. Clock Drawing	.32*	.25*	.24*	.24*	.22*	.28*	.20*	.15*	.31*	1						
11. Category Veg.	.31*	.33*	.44*	.43*	.43*	.35*	.23*	.19*	.17*	.15*	1					
12. Category Anm.	.32*	.30*	.38*	.39*	.41*	.28*	.22*	.22*	.19*	.21*	.55*	1				
13. Boston Naming	.34*	.28*	.32*	.32*	.32*	.29*	.22*	.19*	.26*	.33*	.34*	.46*	1			
14. Naming	.23*	.21*	.26*	.17*	.24*	.14*	.16*	.18*	.19*	.20*	.31*	.23*	.45*	1		
15. Trails A	.17*	.15*	.23*	.21*	.15*	.22*	.10	-.04	.24*	.22*	.22*	.26*	.33*	.15*	1	
16. DS Substitution	.35*	.27*	.38*	.31*	.36*	.26*	.16*	.23*	.26*	.32*	.34*	.44*	.39*	.23*	.50*	1
17. Cancellation	.19*	.14*	.19*	.19*	.18*	.09	.10	.12*	.11*	.14*	.29*	.17*	.21*	.15*	.29*	.41*

Notes: *p < .05. Tests in italics are reverse coded so that all scores represented higher performance. Correlations in the rectangles are from tests that hypothesized to measure the same construct. AVLT learning = AVLT learning, AVLT Short D. = AVLT short delay, Delayed recall = delayed recall (ADAS), AVLT Long D. – AVLT long delay, Word Recog. = word recognition (ADAS), DS backward = digit span backward, Construction = construction (ADAS), Category Veg. = category vegetables, Category Anm. = category animals, Naming = naming (ADAS), DS Substitution = digit symbol substitution, Cancellation = cancellation (ADAS), ADAS = Alzheimer's Disease Assessment Scale, AVLT: Adult Verbal Learning Test.

Table S14. Correlations between MMSE and cognitive tests at T1 for participants with CDR scores greater than or equal to 1.5.

	1	2	3	*4*	5	6	*7*	8	*9*	10	11	12	13	*14*	*15*	16
1. MMSE	1															
2. AVLT Learning	.30*	1														
3. AVLT Short Delay	.23*	.59*	1													
4. Delayed Recall	.46*	.49*	.55*	1												
5. AVLT Long Delay	.29*	.61*	.77*	.57*	1											
6. AVLT Recognition	.29*	.40*	.48*	.49*	.52*	1										
7. Word Recogn.	.32*	.32*	.36*	.48*	.37*	.29*	1									
8. DS backward	.29*	.09	.05	.08	.02	.05	.05	1								
9. Construction	.28*	.13*	.09	.29*	.13*	.18*	.12*	.27*	1							
10. Clock Drawing	.38*	.21*	.15*	.24*	.17*	.16*	.20*	.31*	.39*	1						
11. Category Veg.	.39*	.29*	.37*	.38*	.38*	.22*	.28*	.26*	.24*	.37*	1					
12. Category Anm.	.39*	.24*	.27*	.36*	.27*	.19*	.21*	.26*	.31*	.34*	.60*	1				
13. Boston Naming	.38*	.18*	.19*	.31*	.19*	.19*	.20*	.16*	.24*	.33*	.48*	.57*	1			
14. Naming	.27*	.12*	.12*	.20*	.13*	.10*	.04	.18*	.13*	.19*	.38*	.38*	.59*	1		
15. Trails A	.29*	.10*	.13*	.16*	.05	.12*	.11*	.30*	.38*	.43*	.37*	.41*	.27*	.16*	1	
16. DS Substitution	-.37*	.18*	.21*	.28*	.16*	.17*	.26*	.42*	.38*	.46*	.43*	.45*	.31*	.14*	.68*	1
17. Cancellation	.29*	.11*	.16*	.17*	.08	.06	.18*	.30*	.26*	.39*	.35*	.36*	.25*	.16*	.53*	.60*

Notes: *p < .05. Tests in italics are reverse coded so that all scores represented higher performance. Correlations in the rectangles are from tests that hypothesized to measure the same construct. AVLT learning = AVLT learning, AVLT Short D. = AVLT short delay, Delayed recall = delayed recall (ADAS), AVLT Long D. – AVLT long delay, Word Recog. = word recognition (ADAS), DS backward = digit span backward, Construction = construction (ADAS), Category Veg. = category vegetables, Category Anm. = category animals, Naming = naming (ADAS), DS Substitution = digit symbol substitution, Cancellation = cancellation (ADAS), ADAS = Alzheimer's Disease Assessment Scale, AVLT: Adult Verbal Learning Test.

Table S15. Correlations between MMSE and cognitive tests at T2 for participants with CDR scores less than 1.5.

	1	2	3	*4*	5	6	*7*	8	*9*	10	11	12	13	*14*	*15*	16
1 MMSE	1															
2. AVLT Learning	.10	1														
3. AVLT Short Delay	.11	.58*	1													
4. Delayed Recall	*.01*	*.41**	*.64**	*1*												
5. AVLT Long Delay	-.03	.50*	.84*	*.68**	1											
6. AVLT Recognition	-.11	.34*	.56*	*.56**	.57*	1										
7. Word Recogn.	*.06*	*.20**	*.39**	*.38**	*.36**	*.50**	*1*									
8. DS backward	.11	.03	.14	*.02*	.08	.01	*-.03*	1								
9. Construction	*.23**	*-.04*	*.11*	*.11*	*.13*	*.01*	*.04*	*.09*	*1*							
10. Clock Drawing	.18*	.03	.14	*.15**	.13	.01	*.02*	.21*	*.34**	1						
11. Category Veg.	.05	.27*	.47*	*.39**	.42*	.20*	*.29**	.02	*.22**	.21*	1					
12. Category Anm.	.12	.17*	.39*	*.28**	.40*	.21*	*.26**	.17*	*.16**	.18*	.47*	1				
13. Boston Naming	.14	.08	.19*	*.03*	.16*	.12	*.11*	.07	*.10*	.11	.14	.33*	1			
14. Naming	*-.05*	*.16**	*.16**	*.07*	*.12*	*.15*	*.07*	*.02*	*-.06*	*.03*	*.12*	*.06*	*.40**	*1*		
15. Trails A	*.19**	*.08*	*.19**	*.10*	*.14*	*.05*	*.18**	*.14*	*.01*	*.02*	*.24**	*.35**	*.17**	*-.06*	*1*	
16. DS Substitution	.23*	.18*	.36*	*.31**	.28*	.17*	*.29**	.17*	*.19**	-.08	.35*	.34*	.15*	*.01*	*.53**	1
17. Cancellation	*.01*	*.10*	*.19**	*.21**	*.21**	*.16**	*.23**	*.07*	*.07*	*-.00*	*.19**	*.14*	*-.09*	*.03*	*.14*	*.37**

Notes: *p < .05. Tests in italics are reverse coded so that all scores represented higher performance. Correlations in the rectangles are from tests that hypothesized to measure the same construct. AVLT learning = AVLT learning, AVLT Short D. = AVLT short delay, Delayed recall = delayed recall (ADAS), AVLT Long D. – AVLT long delay, Word Recog. = word recognition (ADAS), DS backward = digit span backward, Construction = construction (ADAS), Category Veg. = category vegetables, Category Anm. = category animals, Naming = naming (ADAS), DS Substitution = digit symbol substitution, Cancellation = cancellation (ADAS), ADAS = Alzheimer's Disease Assessment Scale, AVLT: Adult Verbal Learning Test.

Table S16. Correlations between MMSE and cognitive tests at T2 for participants with CDR scores greater than or equal to 1.5.

	1	**2**	**3**	***4***	**5**	**6**	***7***	**8**	***9***	**10**	**11**	**12**	**13**	***14***	***15***	**16**
1 MMSE	1															
2. AVLT Learning	.13	1														
3. AVLT Short Delay	.29*	.59*	1													
4. Delayed Recall	*.27**	*.46**	*.78**	*1*												
5. AVLT Long Delay	.25*	.52*	.87*	*.75**	1											
6. AVLT Recognition	.39*	.35*	.57*	*.64**	.60*	1										
7. Word Recogn.	*.27**	*.36**	*.57**	*.65**	*.56**	*.62**	*1*									
8. DS backward	.23*	.18*	.18*	*.23**	.15	.21*	*.13*	1								
9. Construction	*.25**	*.11*	*.29**	*.29**	*.22**	*.27**	*.26**	*.21**	*1*							
10. Clock Drawing	.25*	.26*	.34*	*.31**	.29*	.33*	*.39**	.35*	*.28**	1						
11. Category Veg.	.17	.26*	.41*	*.47**	.41*	.32*	*.44**	.31*	*.25**	.42*	1					
12. Category Anm.	.19*	.36*	.56*	*.54**	.52*	.37*	*.49**	.25*	*.20**	.39*	.68*	1				
13. Boston Naming	.28*	.23*	.37*	*.35**	.27*	.25*	*.37**	.29*	*.31**	.47*	.52*	.65*	1			
14. Naming	*.20**	*.13*	*.23**	*.24**	*.17*	*.13*	*.20**	*.28**	*.28**	*.41**	*.47**	*.47**	*.65**	*1*		
15. Trails A	*.09*	*.20**	*.23**	*.28**	*.22**	*.31**	*.22**	*.39**	*.32**	*.42**	*.41**	*.41**	*.36**	*.33**	*1*	
16. DS Substitution	.12	.32*	.38*	*.38**	.38*	.28*	*.31**	.49*	*.39**	.53*	.40*	.41*	.42*	*.32**	*.66**	1
17. Cancellation	*.11*	*.19**	*.26**	*.34**	*.27**	*.23**	*.31**	*.40**	*.30**	*.42**	*.47**	*.43**	*.34**	*.31**	*.69**	*.72**

Notes: *p < .05. Tests in italics are reverse coded so that all scores represented higher performance. Correlations in the rectangles are from tests that hypothesized to measure the same construct. AVLT learning = AVLT learning, AVLT Short D. = AVLT short delay, Delayed recall = delayed recall (ADAS), AVLT Long D. – AVLT long delay, Word Recog. = word recognition (ADAS), DS backward = digit span backward, Construction = construction (ADAS), Category Veg. = category vegetables, Category Anm. = category animals, Naming = naming (ADAS), DS Substitution = digit symbol substitution, Cancellation = cancellation (ADAS), ADAS = Alzheimer's Disease Assessment Scale, AVLT: Adult Verbal Learning Test.