American Psychological Association

Psychological Bulletin

0033-2909/19/$12.00

http://dx.doi.org/10.1037/bul0000179

Coupled Cognitive Changes in Adulthood: A Meta-Analysis

Elliot M. Tucker-Drob
University of Texas at Austin

Andreas M. Brandmaier and Ulman Lindenberger
Max Planck Institute for Human Development, Berlin, Germany, and Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, and London, United Kingdom

With advancing age, healthy adults typically exhibit decreases in performance across many different cognitive abilities such as memory, processing speed, spatial ability, and abstract reasoning. However, there are marked individual differences in rates of cognitive decline, with some adults declining steeply and others maintaining high levels of functioning. To move toward a comprehensive understanding of cognitive aging, it is critical to know whether individual differences in longitudinal changes interrelate across different cognitive abilities. We identified 89 effect sizes representing shared variance in longitudinal cognitive change from 22 unique datasets composed of more than 30,000 unique individuals, which we meta-analyzed using a series of multilevel metaregression models. An average of 60% of the variation in cognitive changes was shared across cognitive abilities. Shared variation in changes increased with age, from approximately 45% at age 35 years to approximately 70% at age 85 years. There was a moderate-to-strong correspondence (r = .49, congruence coefficient = .98) between the extent to which a variable indicated general intelligence and the extent to which change in that variable indicated a general factor of aging-related change. Shared variation in changes did not differ substantially across cognitive ability domain classifications. In a sensitivity analysis based on studies that carefully controlled for dementia, shared variation in longitudinal cognitive changes remained at upward of 60%, and age-related increases in shared variation in cognitive changes continued to be evident. These results together provide strong evidence for a general factor of cognitive aging that strengthens with advancing adult age.

Public Significance Statement
A longstanding question in cognitive aging has been "Does it all go together when it goes?" This meta-analysis indicates that aging-related declines are interrelated across different domains of thinking. For instance, adults who decline steeply in their memory relative to other adults as they get older are also likely to decline steeply in reasoning and processing speed relative to others over the same period of time. These key insights into how changes in different abilities interrelate suggest that theories and interventions for cognitive aging will benefit from considering mechanisms that cut across several different domains of thinking in addition to mechanisms that are specific to each individual domain.

Keywords: cognitive aging, factor analysis, growth curve modeling, latent difference score modeling

Supplemental materials: http://dx.doi.org/10.1037/bul0000179.supp

How Many Causes Are There of Aging-Related Decrements in Cognitive Functioning? (Salthouse, 1994)

With advancing age, adults typically exhibit decreasing performance across many different domains of cognitive function. Although it is sometimes assumed that cognitive aging is a phenomenon confined to very late adulthood that only affects a small subset of diseased individuals, there is now strong evidence that aging-related cognitive declines begin to emerge at least as early as middle adulthood, occur fairly continuously with the passage of time, affect individuals without diagnosed pathologies, and occur throughout the entire distribution of psychological and physical health (Salthouse, 2004a, 2009). Normative aging-related decrements are large. Cross-sectional studies estimate correlations between adult age and abstract reasoning, visuospatial ability, episodic memory, and processing speed at between approximately r = –.40 and r = –.60 (Salthouse, 2004b). Longitudinal studies

Elliot M. Tucker-Drob, Department of Psychology, University of Texas at Austin; Andreas M. Brandmaier and Ulman Lindenberger, Center for Lifespan Psychology, Max Planck Institute for Human Development, Berlin, Germany, and Max Planck UCL Centre for Computational Psychiatry and Ageing Research, Berlin, Germany, and London, United Kingdom.

This research was supported by National Institutes of Health (NIH) Grant R01AG054628. The Population Research Center at the University of Texas is supported by NIH Grant R24HD042849.

Open Science Framework (meta-analytic data and scripts): https://osf.io/hnpr5/?view_only=979d99cfd15d48f38e396e1b191657c2

Correspondence concerning this article should be addressed to Elliot M. Tucker-Drob, Department of Psychology, University of Texas at Austin, 108 East Dean Keeton Stop A8000, Austin, TX 78712-1043. E-mail: tuckerdrob@utexas.edu

indicate shallower decline in earlier adulthood and are more consistent with cross-sectional estimates for later adulthood (Schaie, 1994). When practice effects associated with repeated assessments of individuals followed over time, selective attrition, and cohort effects are taken into account, the apparent gap between longitudinal and cross-sectional estimates of aging-related declines narrows substantially (Lindenberger, Singer, & Baltes, 2002; Lövdén, Ghisletta, & Lindenberger, 2004; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; Rönnlund, Nyberg, Bäckman, & Nilsson, 2005; Salthouse, 2016).

Crucial to a complete account of cognitive aging is its dimensionality, that is, the structure and magnitude of correlations among changes in cognitive abilities (Lindenberger, von Oertzen, Ghisletta, & Hertzog, 2011; Rabbitt, 1993). Whereas research approaches stemming from cognitive-experimental traditions often seek to uncover specific psychological mechanisms for aging-related declines in individual cognitive tasks, approaches stemming from differential psychology have taken seriously the hypothesis that aging-induced changes in general psychological factors, or even a single psychological factor, may largely account for aging-related declines across many different cognitive abilities (Salthouse, 1991). Such *common factor* hypotheses of cognitive aging have been popular for several decades. For instance, according to Salthouse (1988), the hypothesis that an "age-related reduction in some type of general-purpose processing resource contributes to impaired cognitive performance appears to be the only explanation with sufficient generality to account for the age differences observed across a variety of cognitive tasks" (p. 238). Verhaeghen and Salthouse (1997) concluded that "age-related influences on a wide range of cognitive variables are shared" and that "age-related changes in the cognitive system are associated with a decline in some general and fundamental mechanism" (p. 231). Salthouse (2016) more recently commented that "if the contribution of general influences is at least moderate, explanations of domain-specific age relations will need to be supplemented with explanations of general age relations to fully account for cognitive aging phenomena" (p. 1545). Relatedly, Birren (1964), Craik (1983), and Welford (1965) have argued that aging-sensitive psychological resources may limit performance in a large variety of cognitive domains.

Importantly, the question of "how many causes are there" of cognitive aging (Salthouse, 1994) can be addressed at multiple levels of analysis, including the psychological, social, and biological. The key question that has been the topic of much research, and which is the focus of the current meta-analysis, is the dimensionality of cognitive aging at the psychological level of analysis. As Tucker-Drob (2011a) has clarified, a single general psychological cause could "be the outcome of multiple independent biological mechanisms, each broadly affecting cognition" (p. 341). Relatedly, Deary, Cox, and Ritchie (2016, p. 198) have proposed a model of multiple "formative . . . biological elements giving rise to a reflective, psychometric general" psychological dimension, and Lindenberger, Li, and Bäckman (2006) hypothesized that "changes in behavioral repertoires are accompanied by continuous changes in multiple brain-behavior mappings" (pp. 713–714). In other words, rather than directly seeking to identify the many biological and experiential causes that likely exist for cognitive aging, we seek to reveal the extent to which aging-related changes in different cognitive abilities occur along a common statistical dimension. Even though we cannot presently identify the totality of specific causal processes that underlie aging-related cognitive declines, or directly enumerate the number of such specific causal mechanisms, analyses that characterize the dimensionality of aging-related cognitive changes are an important descriptive step that may prove invaluable for guiding ongoing research into specific mechanisms of cognitive aging, and the cognitive dimensions on which they act.

What's Change Got to Do With It? (Lindenberger et al., 2011)

Historically, approaches to testing for shared aging-related effects across multiple cognitive domains have relied on cross-sectional mediation approaches. For instance, early work tested the extent to which cross-sectional age differences in cognitive abilities such as reasoning, visuospatial ability, and episodic memory were mediated by hypothesized "processing resources," such as information processing speed (Lindenberger, Mayr, & Kliegl, 1993; Salthouse, 1996) and working memory capacity (Salthouse, 1990). More recent work of this sort (e.g., Salthouse, 2004b) has tested the extent to which cross-sectional age differences in a range of cognitive abilities are mediated by a common higher order general intelligence factor. Such cross-sectional *shared influence* approaches (Tucker-Drob & Salthouse, 2011) have generally indicated that substantial proportions of age-related effects on different cognitive abilities are mediated by a general intelligence factor, although some residual age effects on individual abilities typically remain.

As has been pointed out by several scholars, cross-sectional mediation approaches reflect, to a large extent, patterns of mean age differences across domains but are unable to directly test whether individual differences in rates of cognitive change are shared across domains (Hofer & Sliwinski, 2001; Horn, 1970; Kalveram, 1965; Lindenberger & Pötter, 1998). When the goal is to test for mediation of aging-related differences in cognitive abilities, either by a processing resource (e.g., processing speed) or by a general factor, cross-sectional approaches are quite limited. As put by Lindenberger and Pötter (1998), cross-sectional mediation "does not offer a test of the basic mediation assumption. All it does is tell us how the world may look if that assumption were true" (p. 227). Hofer, Flaherty, and Hoffman (2006) similarly wrote that "high levels of association between time-dependent processes can result simply from average population age differences and not necessarily from associations between individual 'rates of aging'" (p. 165), and Maxwell and Cole (2007) referred to cross-sectional mediation approaches as "often highly misleading" (p. 23). More recently, Lindenberger et al. (2011) formally demonstrated that high levels of "explained age-related variance" obtained using cross-sectional mediation approaches may stem from either similar average mean age trends, from within-time (but not longitudinal) correlations between the putative mediator and outcome variable, or some mixture of the two. They characterized the link between cross-sectional mediation approaches and developmental codependencies over time about which researchers seek to make inferences as "brittle and volatile" (p. 40). Indeed, examples of strong cross-sectional overlap but much weaker associations among longitudinal changes have been reported for the associations between cognitive abilities and sensory functions

(e.g., Anstey, Hofer, & Luszcz, 2003; Lindenberger & Ghisletta, 2009), and cognitive abilities and physical functions (Ritchie, Tucker-Drob, Starr, & Deary, 2016). Note that the opposite also appears to occur: Abilities that dissociate cross-sectionally, such as verbal knowledge and perceptual speed, have been found to show correlated change when probed longitudinally (e.g., Ghisletta, Rabbitt, Lunn, & Lindenberger, 2012). Thus, the weaknesses of cross-sectional data for indexing the interrelations among time-based developmental associations are not only a logical possibility, but—in at least some substantively important empirical circumstances—a reality.

Does It All Go Together When It Goes? (Rabbitt, 1993)

Fundamental to accurately assessing the dimensionality of cognitive aging are longitudinal approaches that index within-person changes in multiple cognitive abilities over time. Such approaches can be used to test whether individual differences in rates of intraindividual longitudinal changes across different cognitive abilities are interrelated, as would be predicted by common factor theories of cognitive aging. Given that individual differences in cognitive abilities are moderately stable from middle childhood forward (e.g., Deary, 2014; Humphreys & Davey, 1988; Tucker-Drob & Briley, 2014), correlated individual differences in static levels of different cognitive abilities in adulthood could be a vestige of an interdependence that came into existence earlier in life (e.g., effects of schooling on multiple cognitive functions during childhood) but no longer exists in adulthood. In contrast, correlated (i.e., "coupled") individual differences in longitudinal rates of adult cognitive *change* are more likely to reflect systems of influence that are unfolding during adulthood (Tucker-Drob, 2011b). Examining such associations among individual differences in longitudinal cognitive changes allows one to ask directly whether individuals who are declining particularly rapidly relative to their peers in one ability are more likely to be declining rapidly (or to be improving less) relative to their peers in a different cognitive ability, and whether those remaining relatively intact in one ability are also likely to remain relatively intact in another ability. In other words, longitudinal data allow researchers to ask the question "does it all go together when it goes?" (Rabbitt, 1993) at the level of correlated interindividual differences in intraindividual change (Baltes, Reese, & Nesselroade, 1977). This precise question is foundational for addressing common factor theories of cognitive aging. As Deater-Deckard and Mayr (2005) wrote, "The ultimate answer to the question of whether cognitive aging is a general factor or a multifaceted phenomenon will come from careful longitudinal data . . . that allow uncovering the dimensionality of change across a wide range of cognitive abilities" (p. 25).

Factor Analysis Since Spearman: Where Do We Stand? What Do We Know? (Carroll, 1989)

One reason to suspect that longitudinal aging-related declines might be correlated across cognitive abilities is what Deary (2000, p. 6) has described as "arguably the most replicated result in all psychology," namely, that individual differences in cognitive abilities, measured at a single point in time, are positively correlated with one another. This *positive manifold* of correlations was originally discovered by Spearman (1904), and served as the basis for the hypothesis that a common statistical dimension, or what Spearman termed *general intelligence* (*g*), underlies substantial proportions of variation in different cognitive abilities. Spearman formalized this hypothesis using factor analysis, which tests whether an observed matrix of variable intercorrelations can be closely approximated by a model in which all variables interrelate by way of their mutual relations with an unobserved latent factor. In the time since Spearman (1904), factor analytic methods have established that individual differences in cognitive abilities fit a hierarchical structure in which narrow abilities (often indexed by individual tests) load on broader cognitive ability domains (e.g., abstract reasoning, spatial ability, verbal ability, episodic memory, working memory, and processing speed), which in turn load on a single higher order *g* factor (Carroll, 1993). Typically, *g* accounts for upward of 50% of the variance in the first-order ability domains (Carroll, 1993; Tucker-Drob, 2009). Whether *g* should be treated as a veridical psychological entity or simply as a statistical shorthand for conveniently summarizing an otherwise high dimensional matrix of correlations has been a topic of tremendous theoretical interest (Bartholomew, Deary, & Lawn, 2009; Dickens, 2007; Kievit et al., 2018; Kovacs & Conway, 2016; Thurstone, 1938; van der Maas et al., 2006) and ideological consternation and debate (Gould, 1981) over the past century. One view is that higher-order factors such as *g* are "defining a working reference frame, located in a convenient manner in the 'space' defined by all behaviors of a given type" (Cronbach & Meehl, 1955, pp. 277–278). A different view is that *general intelligence* is a "genuine construct" (Gignac, 2016, p. 69) that causally influences the behaviors through which it is expressed (e.g., Panizzon et al., 2014; Spearman, 1904).

The goal of this article is to provide meta-analytic, descriptive evidence on the dimensionality of cognitive change in adulthood. This goal is compatible with both above-described views of general intelligence, though the two views will diverge in the interpretation of the results. That a moderately strong general factor underlies individual differences in different cognitive abilities at a single point in time suggests the possibility, but does not guarantee, that a general factor may underlie individual differences in rates of change in different abilities over time. In other words, the statistical dimensions along which individual differences in cognitive abilities emerge over the course of development may correspond to the dimensions along which individual differences in cognitive aging occur. Salthouse (1988) proposed that declines in general processing resources may underlie aging-related declines in different cognitive abilities, specifically noting strong "parallels between processing resources and intellectual *g*" (p. 251). Juan-Espinosa et al. (2002) provided an anatomic metaphor for life span growth and decline of cognitive abilities (see also Baltes, Cornelius, Spiro, Nesselroade, & Willis, 1980; Schaie, 1962; Tetens, 1777; Werner, 1948). They proposed that, in the same way that age-related growth and shrinkage of the human bones is organized by the anatomical structure of the human skeleton, individual differences in human cognitive abilities may have an inherent structure along which growth and decline naturally occur (see Baltes, Lindenberger, & Staudinger, 2006, for a summary of this line of thought).

Of course, the factor structure of cognitive aging could barely, if at all, resemble the structure of individual differences in cognitive abilities measured at a static point in time. For instance, the

structure of static individual differences in cognitive ability levels in early adulthood may primarily be reflective of how heterogeneity in environmental experience is structured over childhood (e.g., experiences that foster growth in one ability tend to co-occur with other experiences that foster growth in other abilities), from the broad effects of cognitively enriching experiences on many different cognitive abilities over child development, or from the broad effects of intellectual engagement and achievement motivation on many different cognitive abilities over child development (Dickens, 2007; Tucker-Drob, 2013; Tucker-Drob, Briley, & Harden, 2013; for an early exploration of these ideas, see Baltes, Nesselroade, & Cornelius, 1978). In contrast, heterogeneity in aging-related cognitive declines may stem from specific neurodegenerative processes in different neural structures and functions that each subserves a different ability. Indeed, the correlates of levels of cognitive abilities in both childhood and adulthood, including indices of socioeconomic status and physical health have typically failed to significantly predict individual differences in aging-related cognitive declines (Ritchie et al., 2016; Tucker-Drob, Johnson, & Jones, 2009). Moreover, there is little consistent evidence that individual differences in levels of cognitive abilities are systematically correlated with individual differences in changes in those abilities (Verhaeghen, 2013). Thus, although individual differences in cognitive aging may conform to a similarly low dimensional structure as that underlying static individual differences in cognitive abilities, individual differences in cognitive aging do not appear to simply reconstitute static individual differences present during earlier periods of life.

The factor structure of individual differences in cognitive changes over adulthood may drive transformations in the factor structure of individual differences in cognitive abilities with advancing age. This becomes clear when one considers that individual differences in a trait at a particular adult age represent a mixture of individual differences in the levels in that trait that have existed since early adulthood and individual differences in changes in that trait thereafter. As per Hofer and Sliwinski (2001; cf. Hertzog, 1985), the covariance between abilities x and y at time t is a function of the covariance between the abilities at time baseline, the level-change covariances, and the covariance between changes from baseline to time t. If individual differences in levels of different abilities covary moderately in early adulthood, and individual differences in subsequent changes in those abilities are uncorrelated or correlate very weakly, then we would expect the correlation between ability levels to decline with advancing age. If individual differences in levels of different abilities covary moderately in early adulthood, and individual differences in subsequent changes in those abilities are strongly correlated, then we would expect the correlation between ability levels to increase with advancing age. Finally, if the magnitude of the correlation between individual differences in levels of different abilities in early adulthood is similar to the magnitude of the correlation between individual differences in subsequent changes in those abilities, then we would expect the correlation between ability levels to remain relatively constant with advancing age. As Hofer and Sliwinski (2001) write, "as time elapses, the magnitude of the covariance becomes increasingly due to the covariance associated with rates of change relative to the other sources of covariance. Therefore, in older samples of individuals, more time will have transpired and this will increase the contribution . . . that reflects individual differences in rates of ageing" (p. 346). Indeed, this rationale has been the motivation behind several investigations of the *dedifferentiation hypothesis* that abilities become increasingly correlated with adult age (see Baltes et al., 1980 and Reinert, 1970 for early investigations of the dedifferentiation hypothesis; see Cox et al., 2016, for an investigation of the dedifferentiation hypothesis with respect to neurostructural connectivity). In a factor analytic model, the prediction is that a common g factor should account for increasing variance in abilities with age. Evidence for aging-related dedifferentiation of cognitive abilities has been mixed (see Tucker-Drob, 2009, for a review and negative evidence from a large cross-sectional sample).

Building on earlier work by Baltes, Nesselroade, Reinert, and others, de Frias, Lövdén, Lindenberger, and Nilsson (2007) extended the dedifferentiation hypothesis to predict transformations in the factor structure of *aging-related changes* over time. Theorizing that "an ensemble of common sources *increasingly* dominates development of intellectual abilities" (de Frias et al., 2007, p. 382, italics in original) in adulthood, de Frias et al. (2007) predicted that there are age-related increases in "the degree to which changes in a single cognitive ability are associated with changes in other abilities," (de Frias et al., 2007, p. 382) which they termed the *dynamic dedifferentiation hypothesis.*

How Should We Measure "Change" – Or Should We? (Cronbach & Furby, 1970)

Historically, major impediments to progress in longitudinal research on individual differences in change over time stemmed from the unavailability of suitable methods for analyzing longitudinal data. Perhaps the most intuitive approach to analyzing change is to calculate observed difference scores between test scores at baseline and follow-up measurement occasions. However when such "raw change scores" are calculated from measures that themselves have less than perfect reliability, issues surrounding unreliability and regression to the mean compound to such a degree that the true signal of interest—individual differences in cognitive change—becomes highly obscured. Cronbach and Furby (1970) for example wrote that "'raw change' or 'raw gain' scores formed by subtracting pretest scores from posttest scores lead to fallacious conclusions, primarily because such scores are systematically related to any random error of measurement" (p. 68), and that "investigators who ask questions regarding gain scores would ordinarily be better advised to frame their questions in other ways" (p. 80).

Sophisticated methods now exist for analyzing longitudinal data that avoid the many pitfalls associated with raw change scores. These include growth curve models (which are typically specified as structural equation models, hierarchical linear models, mixed effects models, or random coefficient models; McArdle & Nesselroade, 2003; Raudenbush & Bryk, 2002) and latent difference score models (which are typically specified as structural equation models; Kievit et al., 2017; McArdle & Nesselroade, 1994). Both growth curve and latent difference score models estimate latent factors representing change in systematic variance over time. Growth curve models form latent slope factors (random coefficients) that represent systematic individual differences in longitudinal change that conform to a specified functional form, such that random error (which is, by definition, unsystematic over time) is

captured by time-specific residuals. Latent difference score models form latent change factors from occasion-specific latent factors that use psychometric measurement models to confine random measurement error to test-specific uniquenesses. Thus, by modeling systematic variance in longitudinal change, both growth curve models and latent difference score models are in principle able to limit biases that would otherwise result from random error, such as regression to the mean, variance inflation, and correlation attenuation for which "raw change" score approaches are infamous.[1] We provide a formal treatment of multivariate growth curve and latent difference score models in Appendix A.

Over the past 15 years or so, several studies have capitalized on multivariate growth curve and latent difference score models of interrelations among individual differences in changes over time. For instance, noting that the processing speed theory of cognitive aging had been rarely examined in longitudinal data, Zimprich and Martin (2002) reasoned that "if processing speed constitutes an important limiting factor for cognitive functioning, then a person with a specific longitudinal change in processing speed should show a comparable change in other intellectual abilities" (p. 690). Applying latent difference score models to four-year longitudinal data from older adults, they reported that individual differences in changes in processing speed were correlated with individual differences in changes in fluid intelligence at $r = .53$. Wilson et al. (2002) extended this work from the bivariate to the multivariate context. They used growth curve modeling to estimate correlations among individual differences in seven different cognitive variables, including measures of working memory, visual spatial ability, perceptual speed, fluency, episodic memory, and verbal knowledge. When they submitted this correlation matrix to principal component analysis, they found that a single component accounted for 61.6% of the variance in individual differences in cognitive changes. Several more recent studies have combined factor analytic models and growth curve approaches in the form of "factors of curves" models (McArdle, 1988) to estimate common variance in individual differences in cognitive changes. Lindenberger and Ghisletta (2009) reported that a single common factor accounted for 60% of the variance in 13-year longitudinal declines in multiple cognitive variables from the Berlin Aging Study. Tucker-Drob (2011a) reported that a single common factor accounted for 63% of the variance in longitudinal changes in abstract reasoning, spatial visualization, episodic memory, and processing speed composites in participants from the Virginia Cognitive Aging Project over up to 7 years. Using 20-year longitudinal data from middle-aged to very old adults from the U.K., Ghisletta et al. (2012) reported that a single common factor accounted for two thirds of the variance in longitudinal changes in fluid and crystallized intelligence, perceptual speed, and memory.

Several questions remain. First, whereas the studies highlighted above have indicated that approximately 60% of the variance in aging-related cognitive declines is shared across domains, other reports have reported much lower estimates of shared variance. For instance, in longitudinal data from a subset of participants from the Einstein Aging Studies, Sliwinski and Buschke (2004) reported correlations between individual differences in longitudinal changes in memory, speed, and fluency ranging between $r = .16$ and $r = .33$. Thus, a meta-analytic estimate of the magnitude of shared variance across aging-related changes in cognitive abilities is necessary to distinguish whether the true effect is in the range of 50–60% or more, as suggested by some studies, or in the range of 15–30%, as suggested by others. It is particularly informative to compare the magnitude of shared variance in longitudinal change to shared variance in levels from the same studies, to test whether cognitive aging is to a greater, lesser, or comparable extent domain-general as has been established for static individual differences in cognitive abilities (Carroll, 1993; Spearman, 1904). Meta-analysis also provides the opportunity to test whether shared variance differs according to the type of cognitive ability, and according to other moderators, such as the age range within adulthood under study. We therefore conducted a meta-analysis to answer these questions. We also conduct the first formal test of the congruence of factor loading patterns of levels and slopes, allowing us to determine the extent to which the common factor of cognitive aging represents a similar dimension as the general intelligence factor (Carroll, 1993; Spearman, 1904).

Method

Literature Search

Our goal was to collate a comprehensive meta-analytic database containing estimates of shared variation in normal-range aging-related longitudinal changes in two or more cognitive abilities from the corpus of published research. For a study to be considered for inclusion in our meta-analysis, it needed to report an estimate of shared variation in normative aging-related longitudinal changes in measures of two or more different cognitive abilities. Shared variation in change could take the form of correlations or covariances between longitudinal growth curve slopes or latent difference scores, or loadings of longitudinal growth curve slopes or latent difference scores on a common factor (see Appendix A for an overview of the statistical basis for such multivariate models of longitudinal changes). We compiled an initial set of articles based on Table 1 from Tucker-Drob, Briley, Starr, and Deary (2014), which listed (but did not meta-analyze) past major studies reporting relations among rates of change in two or more cognitive variables using a statistical method (e.g., growth curve modeling or latent difference score modeling) for modeling systematic change over time, as separate from random error. We then sought to expand this set in a number of ways. First, we examined the reference sections of each of the papers that met our inclusion criteria to identify further papers that might warrant inclusion, and performed this process iteratively for every new paper included. Second, we used Google Scholar to search for papers citing the included papers to identify further papers that might warrant inclusion, and performed this process iteratively for every new paper included. Third, we performed searches using Google Scholar with combinations of at least one search term from each of the following categories longitudinal change (longitudinal, change,

[1] Both growth curve models and latent difference score models are advantageous for limiting bias that is associated with random measurement error, though they are less effective in limiting bias that results from systematic measurement confounds, such as the influences of variation in manual dexterity on performance on several different tasks. In many circumstances, statistical modeling approaches have limited effectiveness in controlling for confounds that result from study design and task selection.

slope, growth curve, difference score), cognitive (cognitive, cognition, ability, intelligence), and aging (aging, ageing, adult, adulthood).

We excluded studies that primarily focused on clinical populations, studies of child and adolescent populations (individuals under 18 years of age), studies that solely employed dementia screening instruments (e.g., the Mini Mental State Exam [MMSE]) to index cognitive abilities, studies that only examined shared variation in aging-related longitudinal changes in different markers of the same cognitive ability, studies that did not correct (or provide information that could be used to correct) estimates for unreliability (e.g., not using growth curve modeling, latent difference score modeling, or disattenuated correlations among raw change scores), studies that only reported change in shared variation (e.g., a "curve of factors" model) but not shared variation in change (e.g., a "factor of curves" model, a bivariate or multivariate [parallel process] growth model), publications that examined within-person correlations in abilities over time but did not examine between-person correlations in wave-, time-, or age-based longitudinal change, and publications in languages other than English.

When more than one publication based on the same sample met our inclusion criteria, we retained the study that reported the longest longitudinal timespan or changes in the largest number of abilities measured. In instances in which two different publications based on the same sample each contained unique information (e.g., one publication reported results for more abilities, but the other publication analyzed data from a longer longitudinal timespan), we entered results from both studies, and included appropriate downweights and clustering terms for each, as further described below.

Recording Effect Sizes

The key effect sizes that we recorded for the current meta-analysis were estimates of *communality* from a factor model fit to longitudinal changes in indicators of two or more ability domains. These communality estimates can intuitively be conceptualized as indices of shared variation in cognitive changes across ability domains, or as the proportion of variation in change in an ability that is accounted for by a common factor of changes in multiple abilities (see the end of Appendix A for a formal treatment of communality). When a factor model is fit directly to longitudinal slopes, the communality is computed as the standardized factor loading squared. When a single correlation is reported between longitudinal changes in only two variables, that correlation is a direct estimate of communality. This is because, when a factor model is fit to two variables, the standardized loading of each of the two variables on that factor is calculated as the square root of their correlation. When the loading is squared to compute proportion of variance accounted for, this returns the original correlation. Thus, (a) when a correlation was reported between rates of longitudinal changes in only two cognitive variables, we recorded this correlation as the communality; (b) when standardized loadings on a common factor of longitudinal changes in three or more cognitive variables was reported, we recorded the squared standardized loading as the communalities; (c) when correlations or covariances were reported between rates of longitudinal changes in three or more cognitive variables were reported, we fit a factor model to the matrix so as to derive standardized factor loadings, which we then squared and recorded as the communalities. We also recorded the proportions of shared variance in levels (i.e., communalities for growth curve and latent difference score levels) for each variable, such that we could make direct comparisons of shared variance in change to shared variance in levels.

We recorded supplemental information, including information necessary for calculating meta-analytic precision weights, as described below. For growth curve models, this included the variance in latent levels and slopes of the respective cognitive variables, the time-specific residual variances, and the within-variable level-slope correlation. For latent difference score models, this included the variance in latent levels and changes of the respective cognitive ability factors, the within-factor level-change correlation, and the loadings and unique variances for each for each individual indicator of the latent factors. To facilitate interpretability, all parameters were rescaled to reflect latent level variances of 1.0. We additionally recorded the time intervals between each assessment wave, and the sample size at each assessment wave.

We made a number of additional decisions according to the following guidelines. When results were broken down by age group, we entered the parameters from the age groups, rather than parameters from the pooled analysis. When results were available for both individual tests and composites based on multiple tests representing the same ability, we entered the results for the composites. In cases in which an article only provided sample sizes for the number of completed waves (e.g., 400 individuals completed four waves, 350 individuals completed only three waves, 250 completed only three waves), but sample sizes per assessment wave were not provided, we treated the missingness as if it were entirely attributable to dropout (as opposed to, e.g., enrolling new participants at later waves, or participants skipping waves). Standardized estimates greater than 1.0 were top-coded to 1.0. All parameters were coded to reflect scaling in which higher scores indicated better performance. For instance, if a reaction time (RT) measure had a negative loading on a factor, the sign was changed to a positive loading. In situations in which the entire set of parameters required was not provided in the article, but full longitudinal multivariate covariance matrices were provided, we analyzed the covariance matrices with multivariate growth curve or latent difference score models to derive the full set of parameters. In situations in which the complete set of pairwise correlations between latent slopes was provided for three or more cognitive ability domains, we derived factor loadings on a common slope factor for use in the meta-analysis. Further information about specific coding decisions made for individual studies is provided in the online supplemental materials.

Coding Moderators

We additionally coded a variety of characteristics of the effect sizes, samples, and studies as potential moderators of effect size magnitudes.

Mean Age at Baseline Wave was recorded for all studies. When the mean age at baseline was not provided, but an age range at baseline was provided, we coded the midpoint of that age range.

Mean Age at Level was calculated for all studies based on the best available information. The latent level represents individual differences at the age or point in time at which the growth curve basis coefficient is set to zero. Thus, the choice of how to center

time or age in a growth curve model affects the interpretation of the growth curve level. If, for example, age is centered at 70, then the level in an age-based growth model should be interpreted as representing individual differences at age 70, even if the average age at baseline is different (e.g., 60). For age-based modeling, the age at latent level is the age at which the growth curve slope is equal to 0. For time-based modeling (including latent difference score modeling), the age at latent level is the average age of the participants at Time 0 (i.e., baseline).

Longitudinal Time Lag was recorded as the average amount of time that passed between the first and last occasion of measurement for those individuals who completed the final wave of the study.

Broad versus Narrow Ability was recorded for each outcome measure under study. An outcome measure was coded as a broad ability (1) if it was a latent factor or composite index formed from multiple different measures of the same cognitive domain. An outcome measure was coded as a narrow ability (0) if it was measured with a single test. Latent factors derived from alternate forms of the same test were considered indices of narrow abilities.

Cognitive Ability Domains were coded for each outcome under study, regardless of whether the domain was measured with a single test or a broader composite or factor. The following domains were coded: processing speed, episodic memory, working memory, spatial ability, reasoning, verbal knowledge, and prospective memory.

Mean Rate of Longitudinal Change was coded for each outcome under study, when available. This information is reflected in the mean of the growth curve slope (also commonly referred to as a fixed effect) or the mean of the latent difference score. All recorded means were scaled in units of level standard deviations of level per year.

Deriving Meta-Analytic Precision-Weights

Initially, we sought to obtain standard errors for the key effect size estimates of interest (the communalities), or information from which we could derive such standard errors, such as 95% confidence intervals or exact *p* values. This would enable us to weight the contribution of each effect size to the meta-analytic estimate by the precision of that effect size, as is considered best practice in meta-analysis (Cheung, 2015). However, upon reviewing the studies that met inclusion criteria for our meta-analysis, only a minority reported the necessary information for the communalities (standard errors were more consistently available for the mean rates of longitudinal change). Typically, when standard errors are not available, a meta-analysis is conducted using sample size weighting. However, in the context of longitudinal research, weighting by sample size alone is problematic, as several characteristics of the dataset beyond sample size determine the precision of the estimates. For instance, the number of occasions, the amount of time-specific measurement error variance relative to the amount of level and slope variance, the time intervals between occasions, and the degree of attrition across waves all affect the precision of parameter estimates from longitudinal growth curve models (Brandmaier, von Oertzen, Ghisletta, Lindenberger, & Hertzog, 2018; Brandmaier, von Oertzen, Ghisletta, Hertzog, & Lindenberger, 2015; von Oertzen & Brandmaier, 2013). As described in Appendix B, we developed algorithms designed to capitalize on formal mathematical theorems of *effective error* in growth curve modeling and latent difference score modeling (Brandmaier, von Oertzen, et al., 2018; also see Brandmaier, Wenger, et al., 2018) to derive meta-analytic weights proportional to the asymptotic precision of the communality effect size estimates.

Our meta-analytic focus was on the magnitude of shared variance between the individual differences in longitudinal changes in two or more cognitive variables (i.e., the communalities). We therefore developed our weighting algorithm to produce weights proportional to the inverse sampling variances (i.e., inverse of the squared standard error) of the level and slope communalities (equivalent to the correlation, as described earlier) in bivariate growth curve models and latent difference score models. In calculating precision weights, we made a number of simplifying assumptions. First, we assumed that change occurred linearly over time, and that the growth curve model was specified in terms of time since baseline (as opposed to, e.g., occasion number or age). We also assumed that level-slope covariances were negligible (a more complex algorithm that included information about these covariances performed no better at approximating sampling-variance-based weights in a simulation analysis) and that variable-specific autocorrelations in the latent difference score model were negligible. As our algorithm assumed that shared variance was inferred from bivariate "parallel process" growth curve or latent difference score models, we took the following approach to consolidating information from results of multivariate "factor of curves" models, so as to calculate weights for ability-specific loadings when more than one other ability was being modeled. For information pertaining to the first of the two variables, we input the level variance, slope/change variance, time-specific/residual variance, and indicator-loading information for the variable for which we were deriving the weight. For information pertaining to the second of the two variables, we calculated average estimates for level variance, slope/change variance, time-specific/residual variance, and indicator loading taken across the remaining abilities modeled after first standardizing to a common metric.[2] This approach is somewhat conservative, as it does not give extra weight to estimates derived from more complex multivariate models, compared with less complex bivariate models (all effect size estimates are treated as if derived from bivariate models).

Constructing Downweights to Account for Multiple Effect Sizes per Study

In addition to constructing precision weights proportional to the inverse sampling variances, we constructed a downweighting scheme to correct for the fact that many of the studies contributed multiple effect sizes for shared slope variance and shared level

[2] For example, if a study provided loadings for processing speed slope, memory slope, and reasoning slope on a common factor of growth curve slopes, the weight for the squared loading for processing speed would be derived by inputting into the algorithm the level variance, slope variance, and time-specific variance for processing speed as information for the *x* variable. For the *y* variable, the average level variance for memory and reasoning, average slope variance for memory and reasoning, and average time-specific variance for memory and reasoning would be input as information. The time lags and sample sizes are generally constant across variables, such that no specific accommodations need to be made for inputting this information.

variance. For instance a study in which one factor is fit to the levels and a separate factor is fit to the slopes of indices of five different abilities would produce five estimates of shared level variance (i.e., five squared loadings on the first factor) and five estimates of shared slope variance (i.e., five squared loadings on the second factor). All of these estimates contain unique and important information. However, they clearly are not independent. To ensure that studies that measured more variables were not given disproportional leverage on meta-analytic estimates by virtue of contributing a greater number of effect size estimates, we constructed additional "downweights" proportional to the reciprocal number of effect sizes contributed for a given effect size type. Thus, in the above example of a multivariate model of changes in five abilities, the downweights for both shared level variance and shared slope variance would be equal to one fifth (.20). If not for this downweighting, all else being equal, a study with $N = 300$ and 20 variables could have greater leverage on the meta-analytic estimate of shared change variance than a study with $N = 1{,}000$ and only two variables. Downweighting corrects for this potential bias. Moreover, as described in further detail below, we implemented a multilevel modeling approach that accounts for the statistical dependencies that arise when multiple effect sizes are derived from the same sample.

Constructing a More Restrictive Meta-Analytic Dataset for Dementia-Controlled Sensitivity Analyses

We also conducted sensitivity analyses to examine the extent to which results were driven by the presence of substantial subsamples of individuals with dementia in the primary studies (cf. Sliwinski, Hofer, & Hall, 2003). To this end, we constructed an additional meta-analytic dataset based on the above-described procedures using only those studies that met at least one of the following criteria: (a) data from an individual were excluded from any wave at which that individual met criteria for dementia diagnosis or scored in the cognitive impairment range on a dementia screening instrument (e.g., MMSE scores of 23 and below), (b) the reported rate of dementia in the sample was less than 1% over the entire study period, or (c) dementia status was included as a time-varying covariate. In practice, this often involved entering different results from those entered for the main meta-analytic dataset. For instance, Sliwinski et al. (2003) reported separate results for their complete sample ("All Participants") and a dementia-free ("Noncases") subsample. In this case, we entered the results from the dementia-free subsample into this more restrictive dataset in place of those from the complete sample that were entered into the primary meta-analytic dataset. Similarly, Lindenberger and Ghisletta (2009) reported results from models with and without controls for age, time to death, and a dichotomous marker of likely dementia. In this case, we entered the results from the model with controls into this more restrictive dataset in place of those from the model without controls that were entered into the primary meta-analytic dataset. Tucker-Drob (2011a) and Ritchie et al. (2016) did not report full results excluding participants with likely dementia (both reported that sensitivity analyses that excluded likely dementia cases did not substantively change results), but because we had access to the raw data, which contained MMSE scores, we reanalyzed those two datasets excluding data from any wave at which an individual scored in the cognitive impairment range on the MMSE (i.e., scores of 23 and below). Tucker-Drob (2011b) and Ghisletta et al. (2012) did not remove data from participants with dementia or control for dementia status. Unverzagt et al. (2012; their Table 3) reported a dementia event rate of 7.2% (and an incidence rate of 19.2/1,000 person-years) in the sample analyzed by Tucker-Drob (2011b) during the longitudinal period observed. We therefore excluded Tucker-Drob (2011b) from this more restrictive dataset. In contrast, Ghisletta et al. (2012) estimated a dementia prevalence of 20 of 6,203 in their sample. Because this estimated dementia prevalence was less than 1%, we retained results from Ghisletta et al. (2012) in this more restrictive dataset. Finally, Tucker-Drob et al. (2014) analyzed data from a sample of 857 individuals, 48 of whom (5.6%) were diagnosed with dementia and provided cognitive scores during the study period. Results were reported for a model that controlled for dementia as a time-varying covariate, and were therefore included in both the primary meta-analytic dataset and the more restricted meta-analytic dataset. Studies included in the primary meta-analytic dataset that did not indicate the dementia rate or report a method for dealing with participants with dementia (e.g., all results reported by Rast & Hofer, 2014), along with those reporting nontrivial rates of dementia (e.g., Lemke & Zimprich, 2005, as reported in Sattler et al., 2015), were excluded from this more restrictive dataset.

Analytic Approach: Multilevel Metaregression Models

We meta-analyzed effects sizes using a metaregression framework in which the effect sizes of interest were regressed onto hypothesized moderators of the effect sizes, using a weighted fit function that incorporates precision weights and downweights of the outcome variables. Because many of the individual studies included in the meta-analytic dataset contributed effect sizes for multiple variables, we specified metaregression models as two-level models, in which total effects were decomposed into within- and between-study components.

An unconditional multilevel metaregression model of effect size $ES_{i,j}$ for outcome i in study j can be written as:

$$ES_{i,j} = ES_j + u_{i,j}, \tag{1}$$

where ES_j represents an inferred study-specific effect size that is allowed to have a mean and a (between-study) variance $\sigma^2_{ES_j}$ (a so-called random effect), and $u_{i,j}$ is a within-study deviation from the study-specific mean that is assumed to have a mean of zero and a (within-study) variance $\sigma^2_{u_{i,j}}$. In this unconditional model the total variance of $ES_{i,j}$ is therefore specified as the sum of between-study and within-study variation, i.e.:

$$\sigma^2_{ES_{i,j}} = \sigma^2_{ES_j} + \sigma^2_{u_{i,j}}. \tag{2}$$

The unconditional metaregression model can be expanded to allow for predictors (so-called moderators) at the within-study level, the between-study level, or a combination of the two. Such a conditional multilevel metaregression model can be written as:

$$ES_{i,j} = b_0 + \sum_1^k b_{k_j} \cdot x_{k_{i,j}} + u_{i,j} + \sum_1^l b_l \cdot x_{l_j} + u_j, \tag{3}$$

where b_0 is a regression intercept, b_{k_j} is a regression coefficient of the effect size on within-study variable $x_{k_{i,j}}$, b_l is a regression

coefficient of the effect size on between-study variable x_l, $u_{i,j}$ is a within-study residual, and u_j is a between-study residual. Within-study and between-study residuals are specified to have means of zero and freely estimated variances. Variances may also be estimated for selected within-study regression coefficients (b_k), such that they constitute *random slopes* representing between-study variation in the magnitude of the within-study regression effect. Random slopes may be allowed to covary with between-study residuals.

When relevant, we included precision weights by specifying them at the level of the individual effect sizes (i.e., as *within*-cluster weights). Downweights were specified at the level of the contributing samples (i.e., as *between*-cluster weights). Weights were rescaled such that the products of the within-cluster and between-cluster weights sum to the total number of effect sizes in the metaregression model (Asparouhov, 2008). Models were estimated in M*plus* (Muthén & Muthén, 1998–2017). Data cleaning, derivation of weights, and plotting were conducted in *R* (R Core Team, 2016).

Results

Description of Dataset

Descriptive information for each of the studies that contributed effect size estimates to the meta-analysis is provided in Table 1. In total, we identified 89 effect sizes representing shared variance in cognitive change in 98 cognitive outcomes from 22 unique datasets composed of more than 30,000 unique individuals in total. Note that multivariate analyses in which a factor model is fit to the slopes provide individual estimates for the slope variance of each of the variables analyzed and individual estimates for shared variance of each of the variables analyzed. However, bivariate analyses provide individual estimates for the slope variance of each of the two variables but only provide a single estimate for shared variance between the two variables. Because some studies employed bivariate approaches, the number of total shared variance estimates is slightly lower than the number of total variables.

Of the 89 estimates of shared variance, 74 were derived from growth curve models and 15 were derived from latent difference score models. Across the 89 shared variance estimates, number of waves ranged from 2 to 12, with a median of 5, a mean of 5.45, and a standard deviation of 2.40. Across the 89 shared variance estimates, the total time elapsed from beginning to end of the study ranged from 2.81 to 21 years, with a median of 8.41, a mean of 10.33 years, and a standard deviation of 4.59 years. Across the 89 estimates of shared variance, the average age at baseline wave ranged from 35.42 years to 84.92 years, with a median of 64.90, a mean of 66.27 years, and a standard deviation of 11.32 years. The average age at latent level ranged from 35.42 years to 85.00 years, with a median of 69.53, a mean of 66.72 years, and a *SD* of 11.64 years. The correlation between the average age at baseline and the average age at latent level was .91.

Of the 98 outcomes analyzed, 26 indexed processing speed, 35 indexed episodic memory, 3 indexed working memory, 9 indexed spatial ability, 12 indexed reasoning, 12 indexed verbal knowledge, and 1 indexed prospective memory. Forty-three outcomes were classified as broad ability composites or factors and 55 were classified as specific (narrow) measures. Rast and Hofer (2014) did not report mean rate of longitudinal change for any outcome in their paper, and Lemke and Zimprich (2005) did not report mean rate of longitudinal change for memory. Of the remaining 75 outcomes, 20 indexed processing speed, 27 indexed episodic memory, 1 indexed working memory, 8 indexed spatial ability, 10 indexed reasoning, and 9 indexed verbal knowledge. Of these, 40 outcomes were classified as broad ability composites or factors and 35 were classified as specific measures.

Mean Change

To produce a meta-analytic estimate of mean cognitive change, we fit a two-level unconditional metaregression model, with the individual effect size estimates of mean change weighted by their inverse sampling variance and the inverse number of effect sizes contributed by the associated dataset to the complete meta-analytic dataset. The mean change was –.051 ($SE = .007$, $p < .0005$).[3] This indicates that, on average, cognitive performance decreased by approximately one-twentieth of a standard deviation per year, that is, half a standard deviation per decade. In this two-level model that decomposed effect size variation into within-sample and between-sample components, the within-sample standard deviation of mean change estimates was .022 ($SE = .002$, $p < .0005$; variance = $.022^2 = .00048$) and the between-sample standard deviation of mean change estimates was .027 ($SE = .004$, $p < .0005$, variance = $.027^2 = .00073$). This indicates that mean rates of longitudinal change varied substantially across outcomes and across samples.

Level-Slope Correlations

The variable-specific level-slope correlations ranged from –.67 to .84. In a two-level unconditional metaregression model, with the individual effect sizes weighted by the inverse number of effect sizes contributed by the associated dataset, the mean level-slope correlation was –.042 ($SE = .047$, $p = .362$). Five level-slope correlations were exactly zero. Two of these came from Zimprich and Martin (2002), and two came from Lemke and Zimprich (2005), who appear to have fixed these parameters to 0 rather than estimating them. One came from Sliwinski and Buschke (2004), who appeared to have freely estimated the association to be exactly zero. When these five effect sizes were excluded, the weighted mean remained at very close to zero ($-.047$; $SE = .049$, $p = .347$). Finally, to facilitate comparisons with results pertaining to slope communalities, which were the primary focus of the current meta-analysis, we calculated the within-variable level-slope correlations, weighting by the corresponding slope communality precisions and the inverse number of effect sizes contributed by the associated dataset. This estimate was .001 ($SE = .045$, $p = .989$) in the complete dataset, and .000 ($SE = .047$, $p = .996$) when excluding the five effects that were exactly zero. A histogram of within-variable level-slope correlations is depicted in the left panel of Figure 1. It is important to note that the level-slope correlation is dependent on how time is coded in the growth curve model (Biesanz, Deeb-Sossa, Papadakis, Bollen, & Curran, 2004; Rovine & Molenaar, 1998). Because the tendency was for time to be coded

[3] To maintain consistency with results of other analyses, we used the downweights constructed for the complete dataset, even though there were some missing estimates for mean rate of change.

Table 1
Characteristics of Studies Contributing to the Meta-Analytic Dataset

Unique sample ID	Dataset	Reference	Method	*N*	Mean age at baseline wave	Age at level	Total study time span	Waves	Number of shared variance estimates	Number of cognitive outcomes	Variables[a]
1	Advanced cognitive training for independent and vital elderly (nonintervention group)	Tucker-Drob (2011b)	Growth curve	698	74.05	74.05	5.23	6	3	3	Executive reasoning, perceptual speed, episodic memory
2	Age, lead exposure, and neurobehavioral decline	Ferrer, Salthouse, McArdle, Stewart, and Schwartz (2005)	Growth curve	834	56.50	40.00	2.81	4	1	2	Memory, processing speed
3	Berlin Aging Study	Lindenberger & Ghisletta (2009)	Growth curve	516	84.92	84.85	13.00	6	5	5	Digit letter, identical pictures, paired associates, memory for text, categories
4	Bronx Aging Study	Sliwinski, Hofer, and Hall (2003)	Growth curve	466	80.10	85.00	11.00	12	3	3	Fluency, memory, speed
5	Einstein Aging Studies	Sliwinski and Buschke (2004)	Growth curve	650	78.10	78.10	5.00	6	3	3	Memory, speed, fluency
6	English Longitudinal Study of Ageing	Rast and Hofer (2014)	Growth curve	6197	49.00	49.00	6.19	4	3	3	Delayed word recall, animal fluency, prospective memory
7	Health and Retirement Study	Rast and Hofer (2014)	Growth curve	10,060	55.00	55.00	8.04	5	2	4	Immediate word recall, subtract 7s (×2 estimates), delayed word recall
8	Interdisciplinary Longitudinal Study on Adult Development	Lemke and Zimprich (2005)	Latent difference score	474	62.94	62.94	4.00	2	1	2	Processing speed, memory
8	Interdisciplinary Longitudinal Study on Adult Development	Zimprich and Martin (2002)	Latent difference score	417	62.96	62.96	4.00	2	1	2	Fluid intelligence, processing speed
9	Long Beach Longitudinal Study (Cohort 1)	Zelinski and Stewart (1998)	Latent difference score	82	64.00	64.00	16.00	2	5	5	Text recall, list recall, recognition, reasoning, recognition vocabulary
10	Long Beach Longitudinal Study (Cohort 2)	Rast and Hofer (2014)	Growth curve	284	71.00	71.00	8.41	4	3	3	Reasoning, speed, vocabulary
11	Longitudinal Aging Study Amsterdam	Rast and Hofer (2014)	Growth curve	1,367	55.00	55.00	13.15	5	1	2	Ravens Colored Progressive Matrices, alphabet coding task
12	Lothian Birth Cohort of 1936	Ritchie et al. (2016)	Growth curve	1,091	69.53	69.53	6.75	3	4	4	Visuospatial ability, processing speed, episodic memory, crystallized intelligence
13	National Growth and Change Study	Ferrer et al. (2005)	Growth curve	382	52.90	40.00	7.99	3	1	2	Memory, processing speed

(*table continues*)

Table 1 (*continued*)

Unique sample ID	Dataset	Reference	Method	*N*	Mean age at baseline wave	Age at level	Total study time span	Waves	Number of shared variance estimates	Number of cognitive outcomes	Variables[a]
14	Origins of Variance in the Old-Old: Octogenarian Twins	Rast and Hofer (2014)	Growth curve	241	80.00	80.00	8.03	5	3	3	Digit symbol substitution, memory-in-reality free recall, block design
15	Seattle Longitudinal Study	Rast and Hofer (2014)	Growth curve	430	55.00	55.00	21.00	4	3	3	Delayed word recall, word fluency, number comparison
16	Swedish Adoption/ Twin Study of Aging (ages 65 and younger)	Tucker-Drob, Reynolds, et al. (2014)	Growth curve	515	57.50	65.00	16.00	5	4	4	Verbal, spatial, memory, speed
17	Swedish Adoption/ Twin Study of Aging (ages 66 and older)	Tucker-Drob, Reynolds, et al. (2014)	Growth curve	671	80.50	65.00	16.00	5	4	4	Verbal, spatial, memory, speed
18	University of Manchester Longitudinal Study of Cognition and Normal Health	Ghisletta, Rabbitt, Lunn, and Lindenberger (2012)	Growth curve	6,203	64.90	70.00	14.00	8	17	17	Heim Intelligence Tests 1 & 2, Raven Mill Hill Vocabulary version A, verbal free recall, cumulative verbal recall, picture recognition, culture-fair test, visual search, alphabet coding task, semantic reasoning, immediate verbal free recall, delayed verbal recall, shape + spatial locations, propositions about people, memory objects, memory objects + position
19	Victoria Longitudinal Study	Hertzog, Dixon, Hultsch, & MacDonald (2003)	Latent difference score	303	66.65	66.65	6.00	2	8	8	Working memory, reaction time, processing speed, induction, fact recall, word recall, story recall, vocabulary
19	Victoria Longitudinal Study	Rast & Hofer (2014)	Growth curve	293	70.00	70.00	9.50	4	2	4	Simple reaction time, word recall (×2 estimates), identical pictures
20	Virginia Cognitive Aging Project (young adulthood group)	Tucker-Drob (2011a)	Growth curve	455	35.42	35.42	6.00	7	4	4	Abstract reasoning, spatial visualization, episodic memory, processing speed
21	Virginia Cognitive Aging Project (middle adulthood group)	Tucker-Drob (2011a)	Growth curve	562	58.55	58.55	6.00	7	4	4	Abstract reasoning, spatial visualization, episodic memory, processing speed
22	Virginia Cognitive Aging Project (older adulthood group)	Tucker-Drob (2011a)	Growth curve	264	77.30	77.30	6.00	7	4	4	Abstract reasoning, spatial visualization, episodic memory, processing speed

[a] When individual tests were used, the test names are listed. When composite measures or latent factors representing a more general cognitive domain were used, the cognitive domains are listed.

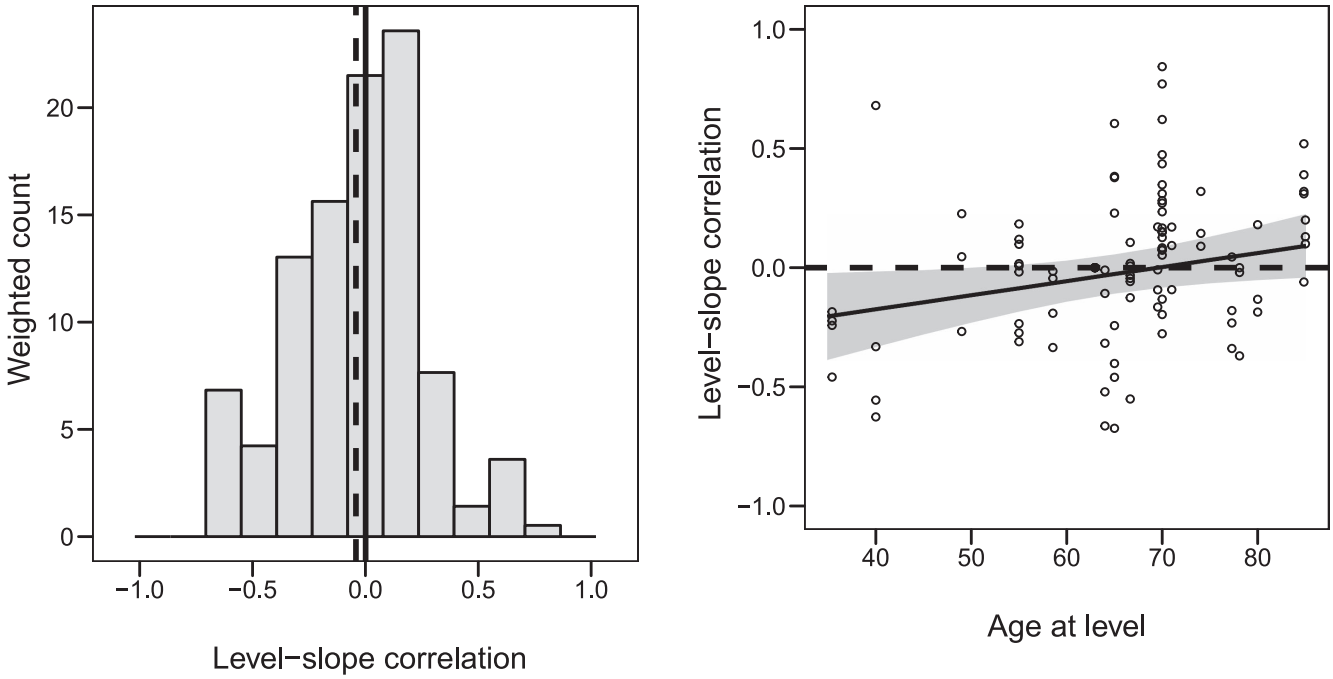

Figure 1. Left: Histogram of the within-variable level-slope correlations, weighted by the inverse number of effect sizes contributed by the associated dataset. Weights are scaled to sum to the total number of effect sizes (98). The dashed vertical line represents the weighted meta-analytic estimate (–.042) for the level-slope correlation. To facilitate comparisons with results pertaining to slope communalities, the solid vertical line depicts the weighted meta-analytic estimate (.001) for the level-slope correlations using the slope communality weights. Right: Level-slope correlation plotted as a function of age at level. The overlaid regression line represents the model-implied trend and its 95% confidence interval from a metaregression model that is weighted by the inverse number of effect sizes contributed by the associated dataset. The positive association between age at level and the level-slope correlation is significant at $p = .031$, but is not significant when restricting analyses to effect sizes for which age at level is greater than 50 (87 observations, $p = .243$).

such that time = 0 corresponded to either the baseline measurement wave or the earlier end of the age range of the sample, it may be most appropriate to interpret these level-slope correlations as representing the relation between early ability levels and subsequent change. To test whether this association was dependent on the age at which the growth curve level was centered (as indicated in the Method section, for age-based modeling, this is the age at which the growth curve slope is equal to 0; for time-based modeling, this is the average age of the participants at Time 0), we fit a two-level metaregression model with age at (latent) level as the independent variable and the level-slope correlation as the dependent variable, weighted by the downweights, and allowing for a random regression intercept. Results of this analysis, superimposed on a scatterplot, are presented in the right panel of Figure 1. Results indicated a significant positive association between age at level and the level-slope correlation ($b = .006$, $SE = .003$, $p = .031$). However, the 95% confidence interval for this age effect only excluded 0 prior to approximately age 50 years, where data were sparse. Moreover, the association was not significant when restricting analyses to effect sizes for which age at level is greater than 50 (87 observations, $p = .243$), or when the metaregression is additionally weighted by the slope communality weights ($p = .050$).

Distributions of Communality Precision Weights

As expected from previous treatments of power to detect correlated change (Hertzog, Lindenberger, Ghisletta, & von Oertzen, 2006; Rast & Hofer, 2014), slope communality precision values were substantially lower than level communality precision values. To index the relative precision of the estimates, we calculated ratios of the precisions of the slope communality estimates to the level communality estimates for each variable. The distribution of these ratios is depicted as a histogram in Figure 2. It can be seen that this distribution was right-skewed, with the majority of ratios falling within the 0 to .20 range. No ratio achieved a value of 1.0 or higher. The median ratio was .072, indicating that the slope communalities tend to be approximately 7% as precise as the level communalities.

Distributions of Level and Slope Communalities

The distributions of level and slope communalities are depicted as histograms in Figure 3. Level communality estimates ranged from .086 to 1.0. In an unconditional two-level metaregression model, weighted by the respective precision of the individual estimates and by the inverse number of effect sizes contributed by the associated dataset, the mean level communality was .558

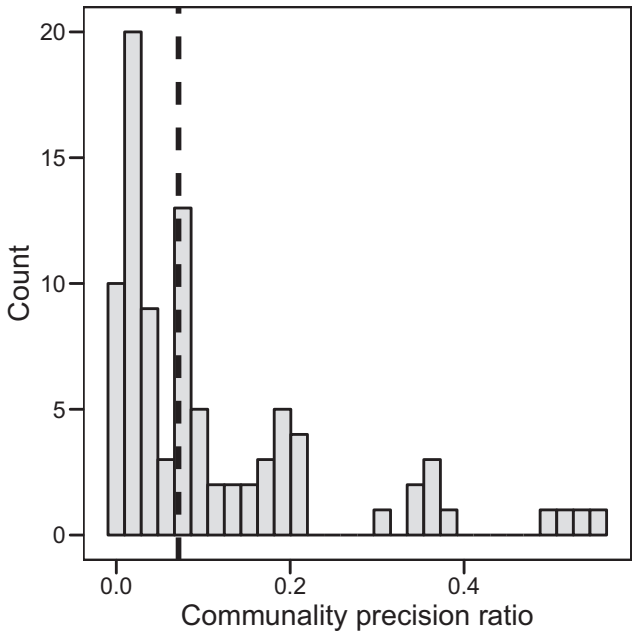

Figure 2. Histogram of the ratios of slope communality precision to level communality precision. The vertical dashed line depicts the median value (.072). All ratios are substantially less than 1.0, indicating that the slope communalities tend to be estimated with substantially less precision than the level communalities.

($SE = .029$, $p < .0005$). This indicates that 56% of the variance in static individual differences is shared across abilities. In this two-level model that decomposed effect size variation into within-sample and between-sample components, the within-sample standard deviation of level communality estimates was .123 ($SE = .026$, $p < .0005$, variance = $.123^2 = .015$) and the between-sample standard deviation of level communality estimates was .047 ($SE = .012$, $p < .0005$, variance = $.047^2 = .002$).

Slope communality estimates ranged from .004 to 1.0. In an unconditional two-level metaregression model, weighted by the respective precision of the individual estimates and by the inverse number of effect sizes contributed by the associated dataset, the mean slope communality was .600 ($SE = .029$, $p < .0005$). This indicates that, on average, 60% of individual differences in aging-related cognitive change is shared across abilities. In this two-level model that decomposed effect size variation into within-sample and between-sample components, the within-sample standard deviation of slope communality estimates was .209 ($SE = .021$, $p < .0005$, variance = $.209^2 = .044$) and the between-sample standard deviation of slope communality estimates was .076 ($SE = .028$, $p = .006$, variance = $.076^2 = .006$). This indicates that the majority of variation in effect size estimates occurs within samples, potentially as a function of characteristics of the individual cognitive abilities or cognitive ability variables. Cross-sample variation, alternatively, could reflect characteristics of the participants (e.g., age) and study design (e.g., time lag, or growth curve vs. latent difference score modeling).[4]

The meta-analytic estimates for the mean slope and level communalities were very similar (.558 for level communalities and .600 for slope communalities). We were interested in whether slight differences in estimates stemmed from differences in the distributions of individual estimates or from differences in the relative contributions of these estimates to the meta-analytic mean resulting from employing differing sets of precision weights for the levels and the slopes. In other words, it is possible that unobserved heterogeneity in effect sizes is correlated with aspects of the study design that are differently correlated with the two sets of precision weights, such that weighting slope communalities and level communalities results in estimates that are representative of different theoretical populations of studies. We therefore ran a sensitivity analysis to determine whether differences in the meta-analytic effect size estimates for levels and slope communalities converged or diverged when using the same set of precision weights. We reran the unconditional metaregression model for the level communalities using the precision weights for the slopes. In this model, the estimate for the mean level communality was .585 ($SE = .024$, $p < .0005$), even closer to the mean slope communality estimate of .600. This can be taken as further evidence that the mean level and slope communalities are extremely similar.

Probing for Publication Bias

One important consideration in meta-analysis is whether the corpus of effect sizes included in the meta-analytic dataset constitute an unbiased representation of the true distribution of effects within the population at large. We would expect the meta-analytic dataset to be biased and unrepresentative if, for example, there is a tendency for authors to be more likely to submit articles, or journal editors to accept articles, reporting results in which effect size estimates are large, or p values are small. This phenomenon is known as *publication bias.* We would also expect the meta-analytic dataset to be biased and unrepresentative in cases in which authors run multiple models, analyze multiple variables, or make multiple alternative data cleaning decisions, but only report results of a subset producing larger effect sizes or smaller p values. This phenomenon is known as *p-hacking*.

To probe for evidence of publication bias, p-hacking, or other types of systematic biases, we produced plots in which the precision of the individual effect size estimates is plotted against the estimates themselves. These plots are known as funnel plots because, when publication bias is low and true population effects are fairly homogeneous, these plots represent an inverted funnel, in which effect sizes are more tightly distributed when estimated more precisely and more widely distributed when estimated less precisely. Under unbiased conditions, the funnel should be symmetrical, with an approximately equal number of estimates above and below the meta-analytic mean. The apex of the funnel, which contains the effect sizes estimated with the highest levels of precision, should be centered close to the meta-analytic mean. The typical pattern thought to be indicative of possible publication bias is one in which lower precision estimates closer to 0 are conspic-

[4] For two-wave designs, latent difference score models can identically be expressed as growth curve models. The effect of LDS versus LGM modeling can therefore be equivalently interpreted as the effect of two versus more than two waves of data collection.

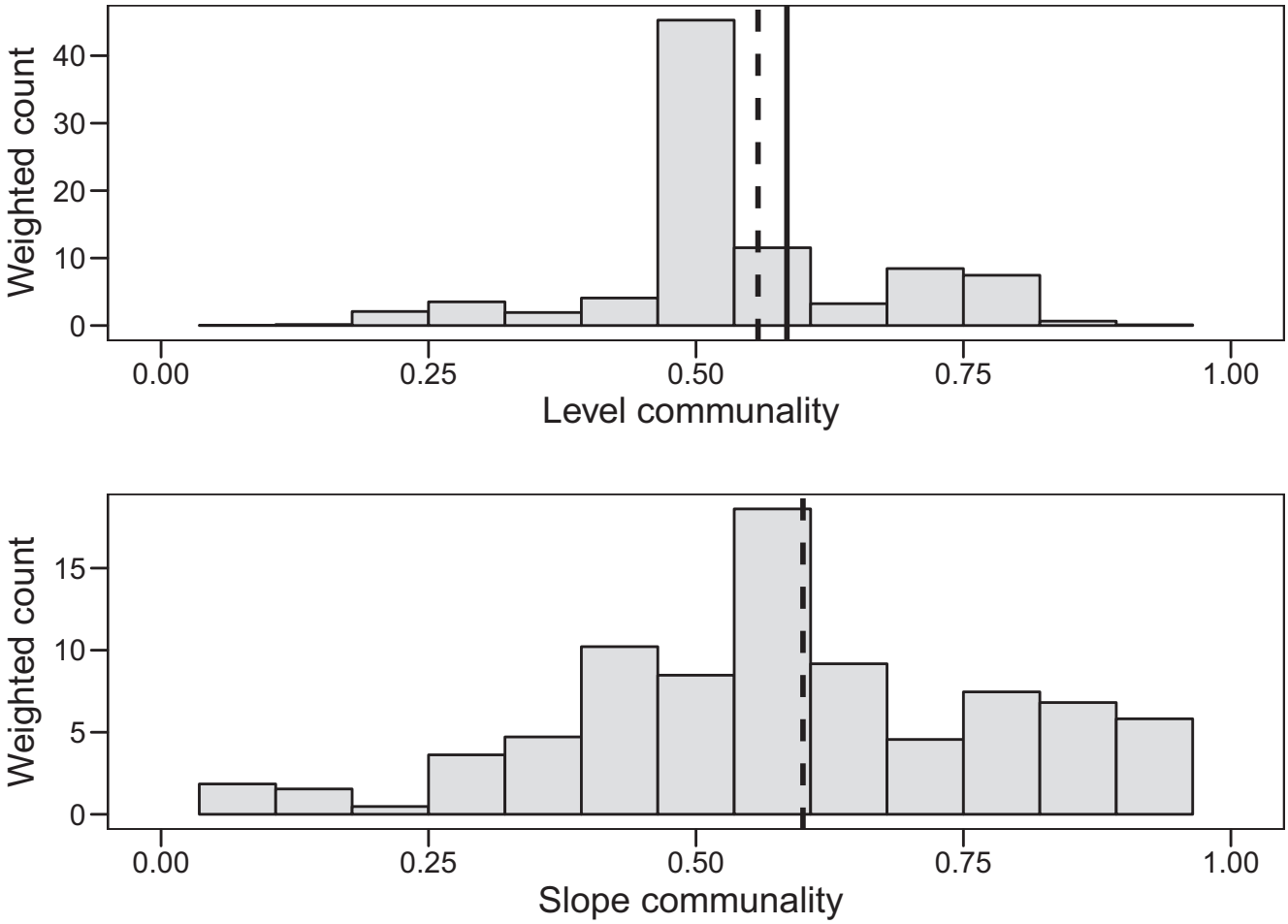

Figure 3. Histograms of level communalities and slope communalities. Histograms are weighted by the respective precision of the individual estimates and by the inverse number of effect sizes contributed by the associated dataset. In each panel, weights are scaled to sum to the total number of effect sizes (89). The dashed vertical line represents the weighted meta-analytic estimate of the mean communality for the levels (.558) and slopes (.600), respectively. To facilitate comparisons across level and slope communalities, the solid vertical line depicts the weighted meta-analytic estimate for the level communalities using the slope communality weights (.585).

uously missing, such that the bottom area of the funnel is asymmetrical.

Funnel plots for the level and slope communality estimates are depicted in Figure 4. It can be seen that both plots are symmetrical and that the most precise estimates are centered within the distributions, close to the respective meta-analytic means. To formally test funnel asymmetry (cf. Stanley & Doucouliagos, 2014) we regressed precision against communality estimates, with and without weighting by precision. For level communality, the p values for the weighted and unweighted regressions were .298 and .759, respectively. For slope communality, the p values for the weighted and unweighted regressions were .486 and .271, respectively. Thus, there was no evidence that effect size estimates were systematically associated with the precisions at which they were estimated, as might occur under conditions of publication bias or p-hacking.

Congruence of Level and Slope Structures

We were next interested in the extent to which the common dimension of change corresponded to the common dimension of levels. Alternatively put, we were interested in whether the extent to which the slope communality for a given variable was predicted by the level communality for that variable. To accomplish this, we fit a two-level metaregression model with (latent) level communalities as the independent variable and slope communalities as the dependent variable, weighted by the dependent variable precision weights and downweights, and allowing for a random regression intercept, a random regression slope, and an intercept-slope covariance. The unstandardized regression coefficient (fixed effect) was .620 ($SE = .129$, $p < .0005$). To give a further sense of this correspondence, the weighted correlation between the vector of level communalities and the vector of slope communalities was

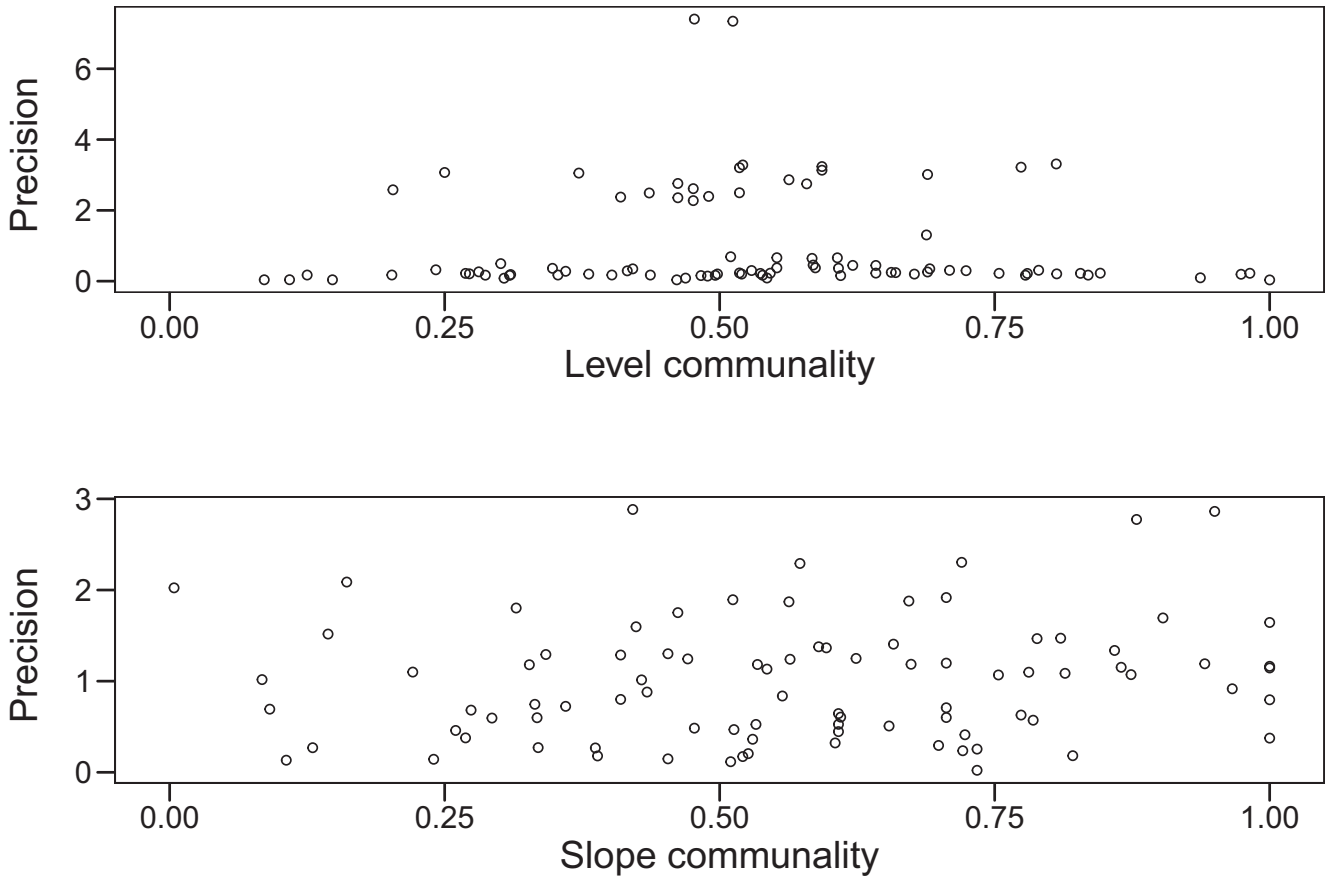

Figure 4. Funnel plots of level communalities and slope communalities. Effect size estimates are on the *x* axis and precision of the estimates are on the *y* axis. In each panel, precision values were scaled such that they sum to the total number of effect sizes (89). It can be seen that both plots are approximately symmetrical. To formally test funnel asymmetry, we regressed effect size estimates on precision, with and without weighting by precision. For level communality, the *p* values for the weighted and unweighted regressions were .298 and .759, respectively. For slope communality, the *p* values for the weighted and unweighted regressions were .486 and .271, respectively. Thus, there was no evidence that effect size estimates were systematically associated with the precisions at which they were estimated, as might occur under conditions of publication bias or p-hacking.

$r = .488$. When the vectors being correlated contained loadings (i.e., the square root of communalities), rather than communalities, the weighted correlation increased to .507. A scatterplot of the association between level communalities and slope communalities is provided in the top panel of Figure 5.

As an alternative approach to indexing the correspondence between the change and level factors, we computed Tucker's congruence coefficients (Lorenzo-Seva & ten Berge, 2006), which are on the same scale as a correlation coefficient (-1 to $+1$). Whereas the correlation coefficient indexes the correspondence between relative ordering of factor loadings across solutions, the congruence coefficient additionally takes into account the absolute magnitudes of factor loadings. The unweighted congruence coefficient representing the congruence of level and slope structures was .968. When weighted using both the slope precision estimates and downweights, the congruence coefficient increased to .982. These very large values reflect the fact that, in addition to displaying similar relative orderings, the level and slope factor loadings were very similar in their overall magnitudes. This was also reflected in the earlier analysis indicating that the meta-analytic estimates for the mean slope and level communalities were very similar.

It was possible that the similarities of level and slope structures derived from heterogeneity in sample-level characteristics (e.g., quality of measurement, selection of variables, participant composition) that affect communality estimates, and did not actually reflect the extent to which level and slope communalities corresponded for individual variables within a given study. To test this possibility we reran the two-level metaregression model with level communalities as the independent variable and slope communalities as the dependent variable, first centering level communality estimates within sample. The unstandardized regression coefficient (fixed effect) from this analysis was .708 ($SE = .167$, $p < .0005$), even larger than those from the earlier analysis of uncentered data, indicating that level and slope communalities indeed tended to correspond for individual variables within each study. Moreover, the weighted correlation between the vector of level communalities and the vector of slope communalities, both centered within

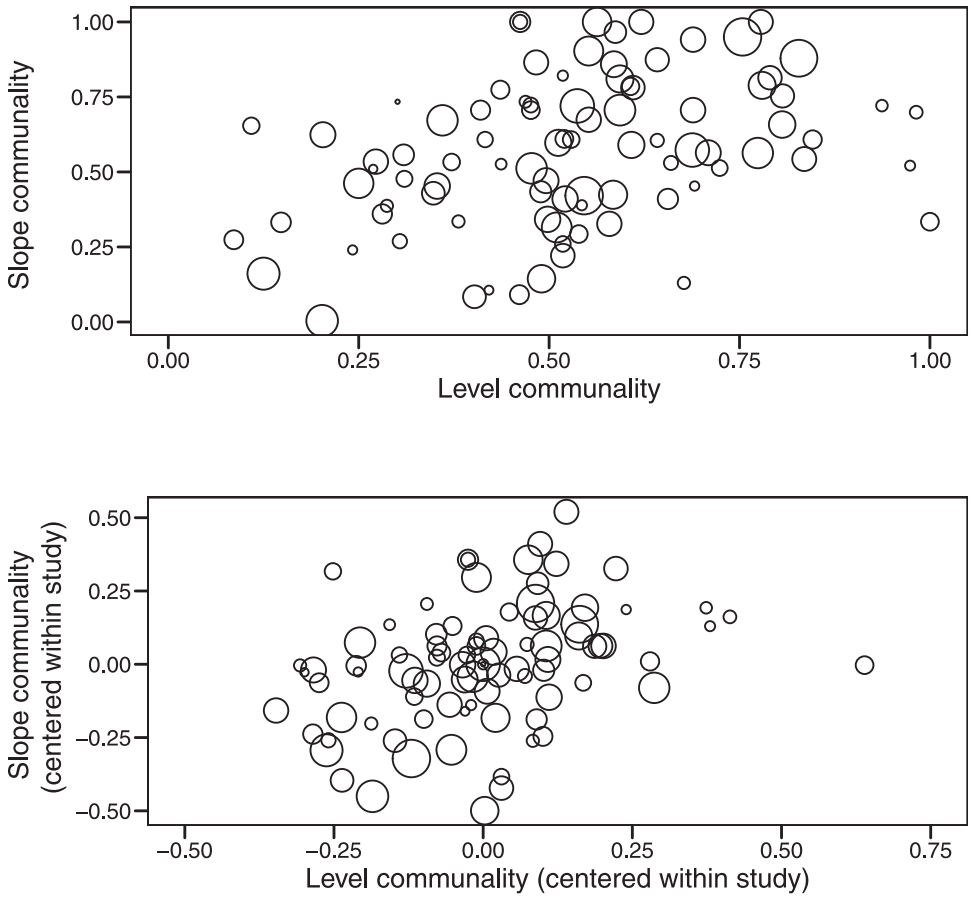

Figure 5. Scatterplots of the association between level communality estimates and slope communality estimates with (top) and without (bottom) centering estimates within dataset. The area of each point corresponds to the precision of the communality estimate, with larger points representing more precise communalities. Precision weights are scaled to sum to the total number of effect sizes (89).

sample, was $r = .486$. When the vectors being correlated contained loadings (i.e., the square root of communalities), centered within sample, the weighted correlation was $r = .462$. A scatterplot of the association between centered level communalities and centered slope communalities is provided in the bottom panel of Figure 5. Note that because centering removes information about the absolute magnitude of the factor loading, Tucker's congruence coefficients cannot be calculated from the centered data.

Domain-Specific Communality Estimates

Next, we were interested in obtaining conditional mean estimates for level and slope communalities in each of the several cognitive ability domains that were measured in the individual studies. To obtain these estimates we fit separate metaregressions to the level and slope communalities associated with the 98 outcomes in the dataset with effect-coded predictors representing six of the cognitive ability domains (episodic memory, working memory, spatial ability, reasoning, verbal knowledge, and prospective memory), with processing speed omitted as the base group (cf. Cohen, Cohen, West, & Aiken, 2003). The coefficients on each of the predictors were then combined with the regression intercepts using the delta method (Ver Hoef, 2012) to produce the conditional mean estimates for level and slope communalities in each cognitive ability domain. Metaregressions were weighted by the respective precision of the dependent variables (the individual communality estimates) and by the inverse number of effect sizes contributed by the associated dataset. To test whether communality estimates differed across domains, we assessed the decrement in fit associated with constraining all six regression parameters to zero. In other words, for both level communalities and slope communalities, we compared the fit of a model in which each of the seven cognitive ability domains was allowed to have its own mean communality to one in which the mean communality was constrained to be invariant across cognitive ability domains.

Results are reported in Table 2. Mean level communalities ranged from .491 for episodic memory to .704 for reasoning. Mean slope communality estimates ranged from .320 for prospective memory slope to .684 for spatial ability slope. Constraining all level communality estimates to be invariant across abilities resulted in a significant loss of model fit, $\chi^2(6) = 32.454$, $p < .0005$. Constraining all slope communality estimates to be invariant across abilities also resulted in a significant loss of model fit,

Table 2
Domain-Specific Effect Size Estimates

Domain	Ability-specific effect size estimate	*SE*	*p* value	Deviation from grand mean	*SE*	*p* value
Level communalities						
Processing speed	.551	.046	<.0005	−.033	.041	.411
Episodic memory	.491	.016	<.0005	−.093	.016	<.0005
Working memory	.492	.010	<.0005	−.092	.015	<.0005
Spatial ability	.640	.036	<.0005	.056	.029	.054
Reasoning	.704	.026	<.0005	.120	.033	<.0005
Verbal knowledge	.665	.071	<.0005	.081	.056	.152
Prospective memory	.546	NA	NA	−.039	NA	NA
Grand mean across domains	.584	.019	<.0005	.000	NA	NA
Longitudinal slope communalities						
Processing speed slope	.690	.050	<.0005	.127	.044	.004
Episodic memory slope	.579	.054	<.0005	.016	.045	.721
Working memory slope	.562	.025	<.0005	−.001	.025	.977
Spatial ability slope	.684	.097	<.0005	.120	.082	.140
Reasoning slope	.644	.037	<.0005	.081	.030	.007
Verbal knowledge slope	.463	.073	<.0005	−.100	.069	.145
Prospective memory slope	.320	.015	<.0005	−.243	.023	<.0005
Grand mean across domains	.563	.026	<.0005	.000	NA	NA
Longitudinal slope means (change per year)						
Processing speed slope	−.068	.008	<.0005	−.0227	.0057	<.0005
Episodic memory slope	−.046	.011	.001	−.0011	.0057	.851
Working memory slope	−.025	NA	NA	.0207	NA	NA
Spatial ability slope	−.058	.008	<.0005	−.0125	.0024	<.0005
Reasoning slope	−.057	.011	<.0005	−.0113	.0045	.012
Verbal knowledge slope	−.019	.009	.032	.0268	.0051	<.0005
Prospective memory slope	NA	NA	NA	NA	NA	NA
Grand mean across domains	−.045	.008	<.0005	.000	NA	NA

Note. Models were fit as effect-coded two-level metaregressions with a random intercept for sample. The coefficients for each of the predictors were then combined with the mean of the regression intercept using the delta method (Ver Hoef, 2012) to produce the conditional mean estimates for each cognitive ability domain. Metaregressions were weighted by the respective precision of the individual estimates and by the inverse number of effect sizes contributed by the associated dataset. Models were fit separately for Level Communalities, Slope Communalities, and Slope Means. Standard errors (*SE*) and *p* values are not reported for Prospective Memory communality, because there was only one data point for Prospective Memory in the meta-analytic dataset. Standard errors (*SE*) and *p* values are not reported for Working Memory mean slope, because there was only one mean slope estimate for Working Memory in the meta-analytic dataset. There were no mean slope estimates for Prospective Memory in the meta-analytic dataset. Constraining all level communality estimates to be invariant across abilities resulted in a significant loss of model fit, $\chi^2(6) = 32.454$, $p < .0005$. Constraining all slope communality estimates to be invariant across abilities also resulted in a significant loss of model fit, $\chi^2(6) = 15.914$, $p = .014$. Constraining all slope mean estimates to be invariant across abilities resulted in a significant loss of model fit, $\chi^2(5) = 27.256$, $p = .0001$. These results indicate that communality estimates and slope means differed across cognitive ability domains.

$\chi^2(6) = 15.914$, $p = .014$. These results indicate that communality estimates differed across cognitive ability domains. Individual post-hoc contrasts (uncorrected for false discovery) between the ability-specific *level* communalities and the simple grand mean indicated that episodic memory and working memory had significantly lower level-communality estimates than the grand mean estimate across domains, and reasoning had a significantly higher level-communality estimate than the grand mean estimate. Individual post-hoc contrasts (also uncorrected for false discovery) between the ability-specific *slope* communalities and the grand mean indicated that processing speed and reasoning had significantly higher slope-communality estimates than the grand mean estimate. Note, however, that because of variability in the number of effect sizes associated with the individual domains, significance levels do not correspond closely to effect sizes.

Figure 6 is a path diagram representing the key results from the domain-specific analyses. This path diagram depicts a single common factor of levels and a single common factor of slopes, with levels and slopes of the seven individual cognitive ability domains loading on the respective factors. Superimposed on the paths from the common factors to the individual domains are standardized loadings, which are computed as the square root of the domain-specific communality estimates.

Domain-Specific Estimates of Mean Change

Next, we were interested in obtaining conditional mean estimates for the mean rates of longitudinal change in each of the ability domains. To obtain these estimates we fit separate metaregressions to the level and slope communalities associated with the 75 available outcomes in the dataset with effect-coded predictors representing five of the cognitive ability domains (episodic memory, working memory, spatial ability, reasoning, verbal knowledge), with processing speed omitted as the base group (cf. Cohen et al., 2003), and additional parameters derived from the primary parameters using the delta method (Ver Hoef, 2012). Note that no mean change estimates were available for prospective memory. Metaregressions were weighted by the inverse sampling variance of the individual estimates and by the inverse number of effect

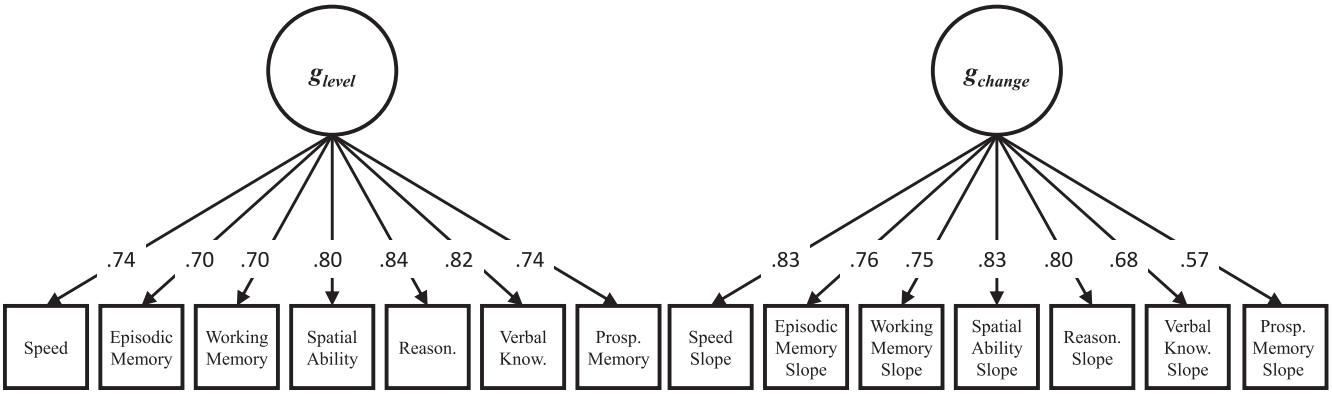

Figure 6. Path diagram representing meta-analytic estimates for standardized factor loadings of levels of individual cognitive abilities on a general factor of levels (left) and standardized factors loadings of longitudinal slopes of individual cognitive abilities on a general factor of changes (right). Variances were omitted from the diagram. Standardized factor loadings were calculated by taking the square root of the respective communalities. Reason. = Reasoning; Verbal Know. = Verbal knowledge; Prosp. Memory = Prospective memory.

sizes contributed by the associated dataset to the complete meta-analytic dataset.

Results are reported in the bottom portion of Table 2. Mean change was −.045 standard deviations per year on average across abilities. Processing speed displayed particularly steep declines, with a mean estimate of −.068 standard deviations per year. In contrast, verbal knowledge displayed particularly shallow declines, with a mean estimate of −.019 standard deviations per year. The estimate for working memory change was also shallow, but because this estimate was only derived from one effect size it is unlikely to be very reliable. Constraining all slope mean estimates to be invariant across abilities resulted in a significant loss of model fit, $\chi^2(5) = 27.256$, $p = .0001$. Individual contrasts indicated that processing speed, spatial ability, and reasoning displayed significantly more decline (more negative) than the grand mean estimate across domains, and verbal knowledge displayed significantly less (less negative) decline than the grand mean estimate.

Age and Other Moderators of Communality Effect Sizes

We went on to test a number of additional moderators of effect size estimates for both slope communality and level communality. We were particularly interested in the static and dynamic versions of the age dedifferentiation hypothesis, which respectively predict increasing level communalities and increasing slope communalities with age. An important consideration for testing these hypotheses concerns how the basis coefficients for the growth curve slopes were parameterized. Centering relative to a constant (e.g., subtracting 65 years) changes the interpretation of growth curve levels but does not change the interpretation of growth curve slopes (Biesanz et al., 2004). Thus, age-related differences in the covariance structure of levels are dependent on the age at which the basis coefficients are centered, whereas age-related differences in the covariance structure of the slopes are dependent on the age composition of the sample, but do not depend on the age at which the basis coefficients are centered. The most appropriate test of the static dedifferentiation hypothesis therefore involves testing age at level as the moderator, whereas the most appropriate test of the dynamic dedifferentiation hypothesis involves testing the actual age composition of the sample as the moderator. As an index of the age composition of the sample, we rely on mean age at baseline.

Age moderation was estimated with two-level metaregression models with communalities as the dependent variable, weighted by the dependent variable precision weights and downweights, and allowing for a random regression intercept. Because mean age at baseline is a between-sample characteristic, we did not estimate random regression slopes. Results are reported in the upper portion of Table 3. We did not find evidence consistent with the static age dedifferentiation hypothesis. Both age at level and mean age at baseline were unrelated to level communalities ($b = .001$, $SE = .002$, $p = .587$; $b = .000$, $SE = .002$, $p = .911$, respectively). However, we did find considerable evidence for the dynamic age dedifferentiation hypothesis. Mean age at baseline was positively related to slope communalities ($b = .005$, $SE = .002$, $p = .001$). As expected, the association between age at level and slope communalities was weaker and nonsignificant ($b = .003$, $SE = .002$, $p = .060$).

To further visualize the moderation of slope communalities by age, we produced a scatterplot of the relation between mean age at baseline and slope communalities, and the metaregression implied a linear relation between these two variables. This plot can be found in the left panel of Figure 7. According to the two-level metaregression model, mean slope communalities increased from approximately 45% at age 35 years to approximately 70% at age 85 years. As can be seen, however, there were very few data points associated with mean baseline ages lower than 50 years. To ensure that the association between mean baseline age and slope communalities was not simply driven by high leverage exerted by these data points, we reran the metaregression model only including effect sizes associated with mean baseline ages greater than 50 years. As reported in

Table 3
Parameter Estimates for Additional Moderators of Meta-Analytic Effect Sizes

Moderator	Level communality			Slope communality		
	Coefficient (unstandardized)	*SE*	*p*	Coefficient (unstandardized)	*SE*	*p*
Mean age at baseline (years)	.000	.002	.911	.005	.002	.001
Mean age at baseline (age > 50 years only)	.002	.002	.218	.004	.002	.013
Age at level (years)	.001	.002	.587	.003	.002	.060
Time lag (years)	.007	.007	.358	.002	.008	.818
Number of waves	.000	.008	.950	.017	.014	.236
LDS (1) versus LGM (0)	−.056	.069	.415	−.184	.069	.008
Age-based (1) versus Time-based (0)	−.086	.066	.194	.070	.063	.270
Broad (1) versus Narrow (0) ability	−.022	.053	.677	.107	.053	.044
Mean rate of change	−.363	.960	.705	−1.156	1.362	.396

Note. Moderators were tested individually. Two-level metaregressions were weighted by the respective precision of the individual communality estimates and by the inverse number of effect sizes contributed by the associated dataset, with random intercepts estimated. Because of substantial variability within datasets for Mean rate of change, we also allowed for a random regression slope and an intercept-slope covariance. *SE* = standard error; LDS = Latent difference score model; LGM = Latent growth curve model.

Table 3, results were nearly identical ($b = .004$, $SE = .002$, $p = .013$) to those obtained from analyses of the entire meta-analytic sample ($b = .005$, $SE = .002$, $p = .001$).

Other moderators tested included the time lag of the longitudinal study, the number of waves, whether the statistical analysis was a latent difference score model or a latent growth curve model, whether a broad or narrow cognitive ability was measured, and whether the change was modeled as occurring as a function of age or time. Parameter estimates, standard errors, and *p* values are reported in Table 3. It can be seen that none of these moderators was significantly related to level communality estimates. Two moderators were significantly related to slope communalities. First, slope communalities derived from latent difference score models tended to be lower than those obtained

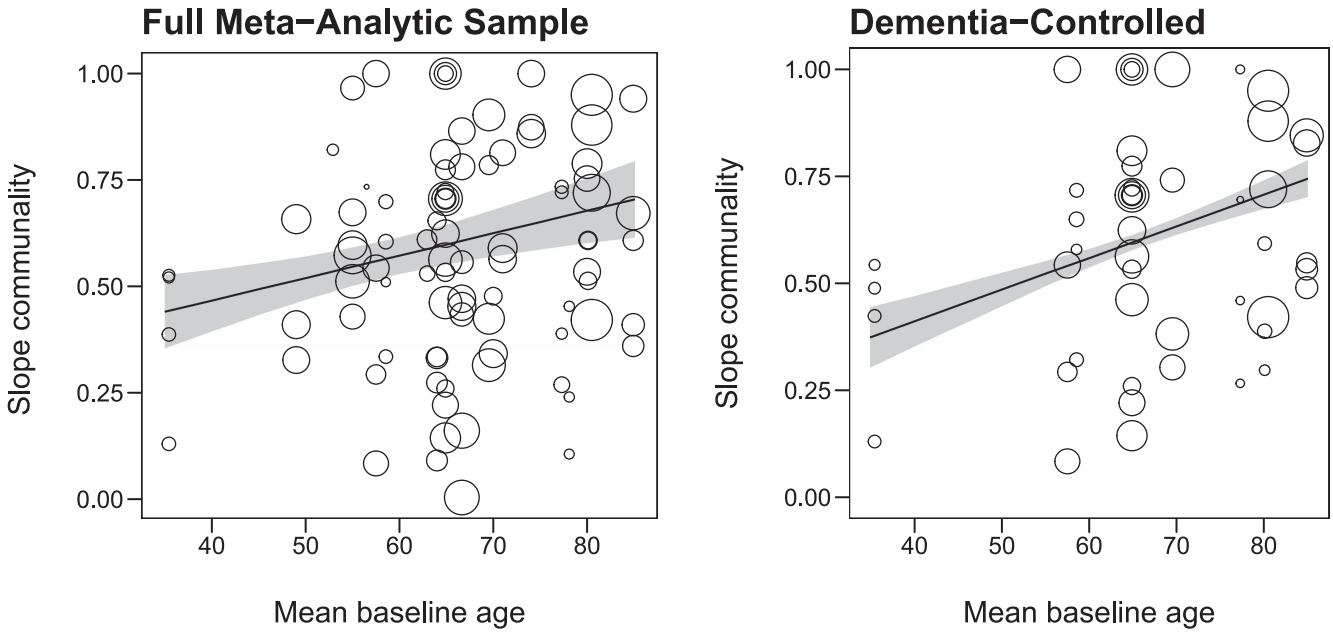

Figure 7. Slope communality plotted as a function of mean age at baseline for full meta-analytic dataset (left) and for dementia-controlled sensitivity analyses (right). The area of each point corresponds to the precision of the slope communality estimate, with larger points representing more precise communalities. The overlaid regression lines represent the metaregression model-implied linear trends and their 95% confidence intervals. For the full dataset, the positive association between slope communality and mean age at baseline remained ($p = .013$) when restricting analysis to estimates derived from mean ages at baseline that were greater than 50 years. For the dementia-controlled sensitivity analyses, the positive association between slope communality and mean age at baseline also remained ($p < .0005$) when restricting analysis to estimates derived from mean ages at baseline that were greater than 50 years.

from growth curve models. Second, slope communalities for broad abilities tended to be higher than those for narrow abilities.

Finally, we tested whether the mean rate of change in a variable was related to its level or slope communality estimate. One might expect that individual differences in changes in variables that exhibit greater mean change are more strongly shared with individual differences in changes with other variables. This was not the case. Mean rate of change in a variable was unrelated to either its slope or level communality. In other words, variables that exhibited steeper mean rates of longitudinal aging-related decline were not more likely to share individual differences in longitudinal change more strongly with other variables.

Simultaneous Analysis of Multiple Moderators

Next, we sought to determine whether each of the associations identified above persisted in a simultaneous model. Thus, we fit a two-level metaregression model in which level communalities, mean age at baseline, whether a latent difference score (vs. growth curve) model was used, whether the outcome was a broad ability (vs. a narrow ability), and effect-coded indicators of cognitive ability domains were predictors of slope communalities. Models were weighted by the respective precision of the dependent variable and by the inverse number of effect sizes contributed by the associated dataset. To maintain consistency with the above-described univariate analyses, we allowed for a random intercept, a random slope for the regression effect of level communalities on slope communalities, and a covariance between the random intercept and the random slope.

Parameter estimates from the simultaneous analysis are reported in Table 4. Level communality and mean age at baseline remained significant moderators of slope communalities. Whether slope communalities were derived latent difference score versus growth curve models and whether slope communalities were for broad versus narrow abilities were no longer significant predictors.

Table 4
Parameter Estimates From Simultaneous Analysis Predicting Slope Communalities

Predictor	Parameter	*SE*	*p* value
Level communality	.535	.151	<.0005
LDS (1) vs. LGM (0)	−.100	.061	.104
Broad (1) versus Narrow (0) ability	.061	.045	.169
Mean age at baseline	.003	.001	.007
Processing speed	.105	.036	.004
Episodic memory	.039	.045	.377
Working memory	.070	.023	.002
Spatial ability	.007	.089	.936
Reasoning ability	.034	.038	.375
Verbal knowledge	−.094	.059	.112
Prospective memory	−.161	NA	NA

Note. Two-level metaregressions were weighted by the respective precision of the individual communality estimates and by the inverse number of effect sizes contributed by the associated dataset. We allowed for a random intercept, a random slope for the regression effect of level communalities on slope communalities, and a covariance between the random intercept and the random slope. Cognitive ability domains were effect coded, with processing speed omitted as the base group. All parameters for cognitive ability domains therefore represent the deviation of the mean communality estimate for the respective ability domain from the simple grand mean across domains. A parameter representing the deviation of the processing speed mean from the simple grand mean across all ability domains was derived from the other parameters using the delta method (Ver Hoef, 2012). Standard errors (*SE*) and *p* values are not reported for prospective memory, because there was only one data point for prospective memory in the meta-analytic dataset. LDS = Latent difference score model; LGM = Latent growth curve model.

Dementia-Controlled Sensitivity Analyses

The more restrictive dataset only containing effect sizes from studies in which dementia status was carefully controlled contained 49 effect sizes representing shared variance in change in 49 outcomes from nine unique samples. Because the more restrictive dataset did not contain bivariate approaches (all were multivariate), the number of total shared variance estimates is equal to the number of total variables. All estimates came from growth curve models. Across the 49 estimates of shared variance, the average age at baseline wave ranged from 35.42 years to 84.92 years, with a median of 64.90, a mean of 67.01 years and a standard deviation of 12.92 years. Of the 49 outcomes analyzed, 12 indexed processing speed, 17 indexed episodic memory, 8 indexed spatial ability, 6 indexed reasoning, 6 indexed verbal knowledge, and 1 indexed prospective memory. Twenty-four outcomes were classified as broad ability composites or factors and 25 were classified as specific measures.

Level communality estimates ranged from .20 to 1.0. In an unconditional two-level metaregression model, weighted by the level precision weights and downweights, the mean level communality was .516 (SE = .017, p < .0005), and .573 (SE = .037, p < .0005) when weighted by the slope precision weights and downweights. Slope communality estimates ranged from .084 to 1.0. In an unconditional two-level metaregression model, weighted by the slope precision weights and downweights, the mean slope communality was .651 (SE = .037, p < .0005). Because the estimates for mean level and slope communalities from this more restrictive sensitivity analysis are very similar to those from the full meta-analytic dataset, it can be inferred that the key finding that approximately half or more of the variance in aging-related cognitive changes are shared across domains is not a simple artifact of confounds associated with dementia status.

We tested whether the earlier-reported association between mean age at baseline and slope communality persisted in this more restrictive dataset that only contained effect sizes from studies in which dementia was carefully controlled. In a two-level metaregression model, weighted by the slope communality precision weights and downweights, and allowing for a random regression intercept, mean age at baseline was positively related to slope communalities (b = .007, SE = .001, p < .0005; results from the full sample reported earlier were b = .005, SE = .002, p = .001). Moreover, when we reran the metaregression model only including effect sizes associated with mean baseline ages greater than 50 years, results were nearly identical (b = .008, SE = .002, p < .0005). The right panel of Figure 7 provides a scatterplot of the relation between mean age at baseline and slope communalities, and the linear association implied by the metaregression. The fact that this dynamic dedifferentiation pattern of age-related increases in shared variance in change was present in this more restrictive

dataset indicates that the pattern is not a simple epiphenomenon of the increased prevalence of dementia at later ages.

Discussion

Common factor theories of human cognitive aging have been popular for some time. Cross-sectional approaches to testing such theories, although capable of capturing information about overlapping mean age trends, are not able to directly gauge the extent to which interindividual differences in different cognitive abilities change in tandem. For more than 15 years, researchers have used longitudinal approaches to estimate covariation among individual differences in rates of aging-related change in different abilities over time. The goal of the current meta-analysis was to systematically compile and meta-analyze results of these longitudinal studies to provide an estimate of the overall magnitude of shared variance in aging-related cognitive changes and test for moderators of the magnitude of shared variance in aging-related cognitive changes.

A primary finding of this meta-analysis is that individual differences in longitudinal changes in different cognitive abilities changes are moderately to strongly correlated with one another. A model in which individual differences in longitudinal cognitive changes are specified to load on a common change factor indicates that an average of 60% of the variance in aging-related cognitive changes is explained by the common factor. This relatively high estimate indicates that individuals who decline precipitously in, for example, processing speed relative to their peers, are also likely to be declining precipitously in, for example, reasoning and episodic memory relative to their peers. Moreover, even though verbal knowledge exhibited relatively shallow mean rates of longitudinal change, individual differences in verbal knowledge change loaded together with individual differences in changes in other cognitive abilities. This indicates that individuals who show greater decline in a cognitive ability showing strong mean decline are less likely to show positive change or stability in verbal knowledge, for which mean decline is quite shallow.

Remarkably, the magnitude of variance in cognitive ability *levels* that was explained by a general intelligence factor was 56%, extremely similar to the proportion of shared variance in rates of *change*. Moreover, there was a moderate-to-strong correspondence between a variable's loading on the general intelligence factor and the extent to which changes in that variable loaded on a general cognitive change factor. When indexed with a correlation coefficient, which quantifies the relative ordering of communality estimates across indicators, the correspondence between level and slope communalities was moderate ($r = .49$). When indexed with a congruence coefficient, which takes into account the absolute magnitudes of communality estimates in addition to their relative orderings, this correspondence was strong (congruence coefficient = .98). The positive correspondence between level and slope communalities was evident even within datasets (i.e., even after centering estimates at study-specific communality means), suggesting that it is not an artifact of unobserved sources of between-study heterogeneity (e.g., the demographic composition of participants or specific aspects of the longitudinal design) that could systematically affect factor loadings. Importantly, however, differences in communality estimates across cognitive ability domain classifications were rather small, and there was not an apparent correspondence between mean level and slope communality estimates across domain classifications. Thus the positive association between level and slope communality estimates for individual study variables may not be driven by correspondence of effect sizes within domains, but instead be attributable to other features of the variables, such as their construct validity or aspects of measurement.

It is useful to compare results from the current meta-analysis to those obtained from three studies that we did not include in the meta-analysis because of the unavailability of information necessary for calculating precision weights. These three studies can therefore be treated as opportunities for out-of-meta-analytic-sample cross-validation. First, Christensen et al. (2004) and Hofer et al. (2002) estimated correlations between longitudinal growth curve slopes for cognitive measures in the Canberra Longitudinal Study. Christensen et al. (2004) report correlations of .42 for memory slope–RT slope, .71 for memory slope–processing speed slope, and .70 for reaction time slope–processing speed slope. Hofer et al. (2002) similarly report correlations of .67 for memory slope–verbal slope, .65 for memory slope–speed slope, and .46 for verbal slope–speed slope. These estimates are all very similar to the mean estimate of 60% shared variance in change from the current meta-analysis. Second, Anstey et al. (2003) reported a correlation between growth curve slopes for memory and processing speed of .62 in an unadjusted model, and .50 in a model that excluded individuals with possible cognitive impairment and adjusted for a host of covariates. Again, these estimates are very similar to those from the current meta-analysis. Third, the analysis of longitudinal data from the Religious Orders Study by Wilson et al. (2002) was of particularly high quality. Wilson et al. (2002) used growth curve (random coefficient) modeling to produce a correlation matrix of individual differences in longitudinal slopes for seven different cognitive variables, including measures of working memory, visual spatial ability, perceptual speed, fluency, episodic memory, and verbal knowledge. Wilson et al. (2002) found that a single principal component accounted for 61.6% of the variance in individual differences in cognitive changes, an estimate strikingly close to the estimate of 60% of shared variance in cognitive change obtained in the current meta-analysis. Importantly, Wilson et al.'s (2002) analysis found that this proportion was nearly identical (61.8%) after accounting for practice effects. A handful of other studies (e.g., Ferrer et al., 2005; Tucker-Drob, 2011a) have also reported that shared variance among aging-related changes persists after controlling for practice effects. However, because most studies did not include sufficient information regarding the role of practice effects, we were not well positioned to formally test their role in the current meta-analysis.

Because the general factor of individual differences in cognitive abilities is moderately stable beginning in middle childhood (Deary, 2014; Humphreys & Davey, 1988; Tucker-Drob & Briley, 2014), static individual differences in *adult* cognitive abilities may substantially reflect processes that have unfolded over child development. Thus, the finding of similarly strong common factors of levels and slopes indirectly suggests that cognitive decline may operate along a similar general dimension as does cognitive development. As Juan-Espinosa et al. (2002) have suggested, the structure of life span changes in

cognitive abilities may be invariant in much the same way that the structure of changes in human anatomy are invariant: Just as age-related growth and shrinkage of the human bones is organized by the anatomical structure of the human skeleton, individual differences in human cognitive abilities may have an inherent structure along which growth and decline naturally occur (see also Baltes, Lindenberger, & Staudinger, 2006; Reinert, 1970; Schaie, 1962; Tetens, 1777; Werner, 1948). Consistent with this proposal, Rhemtulla and Tucker-Drob (2011) reported evidence for a general factor of longitudinal changes across cognitive, psychomotor, and preacademic domains in more than 8,000 children followed between ages 3 and 7 years. Moreover, Tucker-Drob (2009) reported consistent evidence for a general intelligence factor across the age range from 4 to 101 years. Gignac (2014) reported similar results across the age range from 2.5 to 90 years. Cheung, Harden, and Tucker-Drob (2016) found consistent evidence for a general intelligence factor across the range from 0 to 6 years. Moreover, Cheung et al. (2016) found that the general intelligence factor, but not the domain-specific factors, became increasingly heritable with age, suggesting that a "generalist" genetic architecture (Kovas & Plomin, 2006) may undergird substantial portions of individual differences in child cognitive development.

The present findings add an important qualification to two-component theories of adult intellectual development (Kühn & Lindenberger, 2016; Lindenberger, 2001; Tetens, 1777), such as the Cattell-Horn theory of fluid and crystallized intelligence (Gf/Gc theory; Cattell, 1971; Horn, 1989) or the mechanics versus pragmatics theory of cognition (Baltes, 1987). These theories build on the observation that cognitive abilities diverge in their associations with age, presumably reflecting differences in the relative importance of biological and cultural influences. In this meta-analysis, we replicated the well-known pattern of flat or shallow mean declines in verbal knowledge and steep mean declines in fluid/mechanic abilities, such as processing speed, episodic memory, and reasoning. Despite these pronounced differences in patterns of mean decline, all seven cognitive abilities, including verbal knowledge, showed strong and rather uniform loadings on a common factor of change. Indeed, we did not find any indication that the mean change in an outcome was related to the extent to which individual differences in change in that outcome were shared with other outcomes. What this means is that individuals who *decline less* in abilities such as perceptual speed or reasoning are likely to *improve more* on verbal knowledge relative to others. In other words, in spite of relatively stark differences between mechanic and pragmatic abilities in their patterns of mean aging-related declines, individual differences in longitudinal changes in mechanic and pragmatic abilities were moderately coupled. Methodologically, this finding adds weight to the assertion that cross-sectional methods, which are dominated by the contribution of ability differences in mean age trends, do not adequately reflect the implications of models involving covariance of change (Hofer & Sliwinski, 2001; Kalveram, 1965; Lindenberger et al., 2011; Lindenberger & Pötter, 1998).

The finding that older mean baseline age was associated with slope communalities is consistent with what has been termed the *dynamic dedifferentiation hypothesis.* Motivated by a general constraint theory of neurodegeneration and cognitive decline, de Frias et al. (2007), predicted this precise pattern of increasing shared variance in change with advancing adult age. In the meta-analytic dataset, the dynamic dedifferentiation pattern was appreciable. According to the linear moderation model that we fit, the mean expected slope communality at age 35 years is 42%, increasing to 72% by age 85 years. This result is consistent with the hypothesis that "an ensemble of common sources *increasingly* dominates development of intellectual abilities" (de Frias et al., 2007, p. 382).

In contrast, we did not find evidence supporting the *static dedifferentiation hypothesis*, which predicts that a global sources of change should give rise increasing shared variance among ability levels with advancing adult age (cf. Hofer & Sliwinski, 2001; Tucker-Drob, 2009). In other words, neither mean age at baseline nor age at level was associated with level communalities. Increasing correlations among static individual differences with age are predicted to arise when the shared variance in change is larger than the shared variance in levels. Here we find that these proportions are nearly identical. This may explain why it appears that the covariance structure of individual differences is stationary, or homeostatic, over much of adult age. However, particularly in light of the positive evidence for dynamic dedifferentiation, it is possible that had more samples with large proportions of older individuals been included, a static dedifferentiation pattern would have emerged.

One necessary limitation of the meta-analytic dataset is that different studies employed different modalities of measuring the constructs and of modeling change over time. Some studies employed growth curve models whereas other studies employed latent difference score models. Those employing growth curve models differed from one another in their coding of time (e.g., whether change was assumed to occur as a function of age at measurement or time since baseline measurement), and in whether they included nonlinear (e.g., quadratic) components of change in addition to linear components. We tested for differences in effect sizes across age- versus time-based approaches as potential moderators. However, it would have been particularly informative had the original studies analyzed data using a standardized set of modeling strategies. Similarly, studies differed in both the specific assessments used and in the cognitive abilities measured. To address this, we restricted the meta-analytic dataset to only those effects that represented associations across (and not within) cognitive ability domains. According to Spearman's (1927) theorem of "indifference of the indicator," as long as a sufficiently diverse array of cognitive measures is used (Little, Lindenberger, & Nesselroade, 1999), the same latent factor may be triangulated on across different sets of measures. This theorem has been validated empirically (Johnson, Bouchard, Krueger, McGue, & Gottesman, 2004).

It is important to reiterate our statement from earlier in this article that the finding that a single common factor accounts for upward of half of the variance in individual differences in age-related changes in different cognitive abilities indicates that a large proportion of variation in cognitive aging can be organized by a single cognitive dimension, but does not imply that a single social, genetic, or neurobiological cause of cognitive aging is likely. Ghisletta et al. (2012) wrote:

> While we presented strong evidence that the dimensionality of cognitive aging is low, we cannot, based on the present behavioral

> evidence, draw strong conclusions about the number and nature of the underlying driving factors. (p. 267)

Tucker-Drob, Reynolds, et al. (2014) similarly wrote:

> [O]ur finding that a global dimension can account for large proportions of variation in aging-related cognitive changes in older adulthood indicates that late-life cognitive aging is manifest in a largely global pattern of change across multiple variables, but does not indicate that a single cause is responsible for global changes. It is very possible, if not likely, that many thousands of genetic and environmental causes of cognitive aging exist. What the current findings indicate is that . . . these many causes tend to operate at very broad levels to affect many forms of cognition. (p. 164)

Indeed, for the cross-domain coupling among aging-related cognitive changes documented here to represent a meaningful property of the cognitive aging process, rather than a simple epiphenomenon of a less interesting mechanism, we would expect the finding to persist after controlling for key socioeconomic and medical variables that are known to be associated with cognitive abilities. One important potential confound that we considered was dementia status. Because dementia is associated with impairment across multiple cognitive abilities, it is conceivable that correlated changes across cognitive abilities are induced by mean differences in declines between demented and nondemented groups (Harrington et al., 2018; but see Boyle et al., 2013 and Sibbett, Russ, Pattie, Starr, & Deary, 2018), but not a more general characteristic of covariation among cognitive declines within the respective groups. Our sensitivity analyses indicated that this was not the case. Even when we performed the meta-analysis on a highly restricted dataset only containing studies that excluded data from person-waves at which dementia was present, controlled for dementia status as a time-varying covariate, or provided an estimate of a low rate of dementia in the sample, the key finding of cross-domain coupling among aging-related cognitive changes persisted. In fact, inspection of results of each of these individual studies indicates that the overall pattern is present in each study. Some individual studies included a wider range of control variables beyond dementia, and continued to document evidence for a general factor of aging-related cognitive change. For instance, Tucker-Drob (2011b) reported substantial coupling among aging-related changes in reasoning, processing speed, and episodic memory even after controlling for age, sex, years of educational attainment, MMSE score, and baseline performance levels. Tucker-Drob, Briley, et al. (2014) similarly reported that strong evidence for a general factor of longitudinal cognitive change persisted even after controlling for a host of carefully measured demographic, physical health, and medical variables, including forced expiratory volume, walk time, grip strength, smoking status, cardiovascular disease status, hypertension status, diabetes diagnosis, along with early life IQ, educational attainment, sex, age, and time lag. Lindenberger and Ghisletta (2009) reported that after controlling for age, time to death, and dementia risk, a general factor of change went from accounting for 60% of the variance to 65% of the variance in cognitive changes. Overall then, the evidence is very consistent with the conclusion that the factor structure of cognitive aging is more than an epiphenomenon of a small and obvious set of simple confounding variables. Rather, a common factor of cognitive change may be a more fundamental description of the cognitive aging process.

Even accepting the conclusion that coupled aging-related cognitive changes represent a meaningful property of the cognitive aging process, the question remains as to whether the common factor of aging identified in the current meta-analysis represents a coherent entity that is directly affected by biological and contextual etiological factors for cognitive aging and affects changes in individual cognitive domains. An equally logical possibility is that the common factor of change represents an emergent property of dynamical systems processes that occur more directly between etiological factors and ability domains. In the Introduction to this article, we refrained from taking a position on this issue. Complex systems approaches, such as graph-theoretic network models, for representing interrelations among individual differences in aging-related changes may be able to faithfully, or perhaps even more accurately, present the patterns of change interrelations that were captured by the factor analytic approach taken here (cf. van der Maas et al., 2006). When summarizing general patterns in complex networks, global metrics may be used (Borsboom, Cramer, Schmittmann, Epskamp, & Waldorp, 2011). In this sense, common factors and global network metrics capture general patterns of covariance in the data, while eliding nuance. We believe that both approaches are useful, albeit imperfect, insofar as they convey the most salient and robust patterns present in the data.

In summary, we found that more than half of the variance in cognitive changes is shared across cognitive abilities. Specifically, the meta-analytic estimate of average change communality was 60%, which is very similar to the estimate of 56% for shared variance in levels. Moreover, we found that shared variance in changes increased with age, from approximately 40% at age 35 years to approximately 70% at age 85 years. These patterns persisted at full strength in a sensitivity analysis based on studies that carefully controlled for dementia. These results together provide strong evidence for a general factor of cognitive aging that strengthens with advancing adult age.

References

References marked with an asterisk indicate studies included in the meta-analysis.

Anstey, K. J., Hofer, S. M., & Luszcz, M. A. (2003). A latent growth curve analysis of late-life sensory and cognitive function over 8 years: Evidence for specific and common factors underlying change. *Psychology and Aging, 18,* 714–726. http://dx.doi.org/10.1037/0882-7974.18.4.714

Asparouhov, T. (2008). *Scaling of sampling weights for two level models in Mplus 4.2*. Los Angeles, CA: Muthén & Muthén.

Baltes, P. B. (1987). Theoretical propositions of life-span developmental psychology: On the dynamics between growth and decline. *Developmental Psychology, 23,* 611–626.

Baltes, P. B., Cornelius, S. W., Spiro, A., Nesselroade, J. R., & Willis, S. L. (1980). Integration vs. differentiation of fluid-crystallized intelligence in old age. *Developmental Psychology, 16,* 625–635. http://dx.doi.org/10.1037/0012-1649.16.6.625

Baltes, P. B., Lindenberger, U., & Staudinger, U. M. (2006). Life-span theory in developmental psychology. In W. Damon & R. M. Lerner (Eds.), *Handbook of Child Psychology: Theoretical models of human development* (6th ed., Vol. 1, pp. 569–664). New York, NY: Wiley.

Baltes, P. B., Nesselroade, J. R., & Cornelius, S. W. (1978). Multivariate antecedents of structural change in development: A simulation of cu-

mulative environmental patterns. *Multivariate Behavioral Research, 13,* 127–152. http://dx.doi.org/10.1207/s15327906mbr1302_1

Baltes, P. B., Reese, H. W., & Nesselroade, J. R. (1977). *Life-span developmental psychology: Introduction to research methods*. Monterey, CA: Brooks.

Bartholomew, D. J., Deary, I. J., & Lawn, M. (2009). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review, 116,* 567–579. http://dx.doi.org/10.1037/a0016262

Biesanz, J. C., Deeb-Sossa, N., Papadakis, A. A., Bollen, K. A., & Curran, P. J. (2004). The role of coding time in estimating and interpreting growth curve models. *Psychological Methods, 9,* 30–52. http://dx.doi.org/10.1037/1082-989X.9.1.30

Birren, J. E. (1964). *The psychology of aging*. Englewood Cliffs, NJ: Prentice Hall.

Borsboom, D., Cramer, A. O., Schmittmann, V. D., Epskamp, S., & Waldorp, L. J. (2011). The small world of psychopathology. *PLoS ONE, 6,* e27407. http://dx.doi.org/10.1371/journal.pone.0027407

Boyle, P. A., Wilson, R. S., Yu, L., Barr, A. M., Honer, W. G., Schneider, J. A., & Bennett, D. A. (2013). Much of late life cognitive decline is not due to common neurodegenerative pathologies. *Annals of Neurology, 74,* 478–489. http://dx.doi.org/10.1002/ana.23964

Brandmaier, A. M., von Oertzen, T., Ghisletta, P., Hertzog, C., & Lindenberger, U. (2015). LIFESPAN: A tool for the computer-aided design of longitudinal studies. *Frontiers in Psychology, 6,* 272. http://dx.doi.org/10.3389/fpsyg.2015.00272

Brandmaier, A. M., von Oertzen, T., Ghisletta, P., Lindenberger, U., & Hertzog, C. (2018). Precision, reliability, and effect size of slope variance in latent growth curve models: Implications for statistical power analysis. *Frontiers in Psychology, 9,* 294. http://dx.doi.org/10.3389/fpsyg.2018.00294

Brandmaier, A. M., Wenger, E., Bodammer, N. C., Kühn, S., Raz, N., & Lindenberger, U. (2018). Assessing reliability in neuroimaging research through intra-class effect decomposition (ICED). *eLife, 7,* e35718. http://dx.doi.org/10.7554/eLife.35718

Carroll, J. B. (1989). Factor analysis since Spearman: Where do we stand? What do we know? In R. Kanfer, P. L. Ackerman, & R. Cudeck (Eds.), *Abilities, motivation, and methodology: The Minnesota Symposium on Learning and Individual Differences* (Vol. 10, pp. 43–70). London, UK: Routledge.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. United Kingdom: Cambridge University Press. http://dx.doi.org/10.1017/CBO9780511571312

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston, MA: Houghton Mifflin.

Cheung, A. K, Harden, K. P., & Tucker-Drob, E. M. (2016). Multivariate behavioral genetic analysis of parenting in early childhood. *Parenting: Science and Practice, 16,* 257–283.

Cheung, M. W. L. (2015). *Meta-analysis: A structural equation modeling approach*. Chichester, UK: Wiley. http://dx.doi.org/10.1002/9781118957813

Christensen, H., Mackinnon, A., Jorm, A. F., Korten, A., Jacomb, P., Hofer, S. M., & Henderson, S. (2004). The Canberra Longitudinal Study: Design, aims, methodology, outcomes and recent empirical investigations. *Aging, Neuropsychology and Cognition, 11,* 169–195. http://dx.doi.org/10.1080/13825580490511053

Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2003). *Applied multiple correlation/regression analysis for the behavioral sciences*. London, UK: Taylor & Francis.

Cox, S. R., Ritchie, S. J., Tucker-Drob, E. M., Liewald, D. C., Hagenaars, S. P., Davies, G., . . . Deary, I. J. (2016). Ageing and brain white matter structure in 3,513 U.K. Biobank participants. *Nature Communications, 7,* 13629. http://dx.doi.org/10.1038/ncomms13629

Craik, F. I. M. (1983). On the transfer of information from temporary to permanent memory. *Philosophical Transactions of the Royal Society of London Series B, Biological Sciences, 302,* 341–359. http://dx.doi.org/10.1098/rstb.1983.0059

Cronbach, L. J., & Furby, L. (1970). How we should measure "change": Or should we? *Psychological Bulletin, 74,* 68–80. http://dx.doi.org/10.1037/h0029382

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin, 52,* 281–302. http://dx.doi.org/10.1037/h0040957

Deary, I. J. (2000). *Looking down on human intelligence: From psychometrics to the brain* (Vol. 34). New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780198524175.001.0001

Deary, I. J. (2014). The stability of intelligence from childhood to old age. *Current Directions in Psychological Science, 23,* 239–245. http://dx.doi.org/10.1177/0963721414536905

Deary, I. J., Cox, S. R., & Ritchie, S. J. (2016). Getting Spearman off the skyhook: One more in a century (since Thomson, 1916) of attempts to vanquish *g. Psychological Inquiry, 27,* 192–199. http://dx.doi.org/10.1080/1047840X.2016.1186525

Deater-Deckard, K., & Mayr, U. (2005). Cognitive change in aging: Identifying gene-environment correlation and nonshared environment mechanisms. *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 60,* 24–31. http://dx.doi.org/10.1093/geronb/60.Special_Issue_1.24

de Frias, C. M., Lövdén, M., Lindenberger, U., & Nilsson, L. G. (2007). Revisiting the dedifferentiation hypothesis with longitudinal multicohort data. *Intelligence, 35,* 381–392. http://dx.doi.org/10.1016/j.intell.2006.07.011

Dickens, W. T. (2007). What is g? Retrieved May 1, 2017, from https://www.brookings.edu/wp-content/uploads/2016/06/20070503.pdf

*Ferrer, E., Salthouse, T. A., McArdle, J. J., Stewart, W. F., & Schwartz, B. S. (2005). Multivariate modeling of age and retest in longitudinal studies of cognitive abilities. *Psychology and Aging, 20,* 412–422. http://dx.doi.org/10.1037/0882-7974.20.3.412

*Ghisletta, P., Rabbitt, P., Lunn, M., & Lindenberger, U. (2012). Two thirds of the age-based changes in fluid and crystallized intelligence, perceptual speed, and memory in adulthood are shared. *Intelligence, 40,* 260–268. http://dx.doi.org/10.1016/j.intell.2012.02.008

Gignac, G. E. (2014). Dynamic mutualism versus g factor theory: An empirical test. *Intelligence, 42,* 89–97. http://dx.doi.org/10.1016/j.intell.2013.11.004

Gignac, G. E. (2016). Residual group-level factor associations: Possibly negative implications for the mutualism theory of general intelligence. *Intelligence, 55,* 69–78. http://dx.doi.org/10.1016/j.intell.2016.01.007

Gould, S. J. (1981). *The Mismeasure of Man*. New York, NY: W. W. Norton & Company.

Harrington, K. D., Schembri, A., Lim, Y. Y., Dang, C., Ames, D., Hassenstab, J., . . . the AIBL Research Group. (2018). Estimates of age-related memory decline are inflated by unrecognized Alzheimer's disease. *Neurobiology of Aging, 70,* 170–179. http://dx.doi.org/10.1016/j.neurobiolaging.2018.06.005

Hertzog, C. (1985). An individual differences perspective. Implications for cognitive research in gerontology. *Research on Aging, 7,* 7–45. http://dx.doi.org/10.1177/0164027585007001002

*Hertzog, C., Dixon, R. A., Hultsch, D. F., & MacDonald, S. W. (2003). Latent change models of adult cognition: Are changes in processing speed and working memory associated with changes in episodic memory? *Psychology and Aging, 18,* 755–769. http://dx.doi.org/10.1037/0882-7974.18.4.755

Hertzog, C., Lindenberger, U., Ghisletta, P., & von Oertzen, T. (2006). On the power of multivariate latent growth curve models to detect correlated change. *Psychological Methods, 11,* 244–252. http://dx.doi.org/10.1037/1082-989X.11.3.244

Hofer, S. M., Christensen, H., Mackinnon, A. J., Korten, A. E., Jorm, A. F., Henderson, A. S., & Easteal, S. (2002). Change in cognitive functioning

associated with apoE genotype in a community sample of older adults. *Psychology and Aging, 17,* 194–208. http://dx.doi.org/10.1037/0882-7974.17.2.194

Hofer, S. M., Flaherty, B. P., & Hoffman, L. (2006). Cross-sectional analysis of time-dependent data: Mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohort samples. *Multivariate Behavioral Research, 41,* 165–187. http://dx.doi.org/10.1207/s15327906mbr4102_4

Hofer, S. M., & Sliwinski, M. J. (2001). Understanding ageing: An evaluation of research designs for assessing the interdependence of ageing-related changes. *Gerontology, 47,* 341–352. http://dx.doi.org/10.1159/000052825

Horn, J. L. (1970). Organization of data on life-span development of human abilities. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 423–466). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-0-12-293850-4.50022-4

Horn, J. L. (1989). Models of intelligence. In R. L. Linn (Ed.), *Intelligence: Measurement, theory, and public policy* (pp. 29–73). Urbana: University of Illinois Press.

Humphreys, L. G., & Davey, T. C. (1988). Continuity in intellectual growth from 12 months to 9 years. *Intelligence, 12,* 183–197. http://dx.doi.org/10.1016/0160-2896(88)90015-3

Johnson, W., Bouchard, T. J., Jr., Krueger, R. F., McGue, M., & Gottesman, I. I. (2004). Just one g: Consistent results from three test batteries. *Intelligence, 32,* 95–107. http://dx.doi.org/10.1016/S0160-2896(03)00062-X

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. http://dx.doi.org/10.1016/S0160-2896(02)00092-2

Kalveram, K. T. (1965). Die Veränderung von Faktorenstrukturen durch "simultane Überlagerung" [Alterations in factorial structure through "simultaneous interference"]. *Archiv für die gesamte Psychologie, 117,* 296–305.

Kievit, R. A., Brandmaier, A. M., Ziegler, G., van Harmelen, A.-L., de Mooij, S. M. M., Moutoussis, M., . . . Dolan, R. J. (2018). Developmental cognitive neuroscience using latent change score models: A tutorial and applications. *Developmental Cognitive Neuroscience, 33,* 99–117. http://dx.doi.org/10.1016/j.dcn.2017.11.007

Kievit, R. A., Lindenberger, U., Goodyer, I. M., Jones, P. B., Fonagy, P., Bullmore, E. T., . . . the Neuroscience in Psychiatry Network. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science, 28,* 1419–1431. http://dx.doi.org/10.1177/0956797617710785

Kovacs, K., & Conway, A. R. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry, 27,* 151–177. http://dx.doi.org/10.1080/1047840X.2016.1153946

Kovas, Y., & Plomin, R. (2006). Generalist genes: Implications for the cognitive sciences. *Trends in Cognitive Sciences, 10,* 198–203. http://dx.doi.org/10.1016/j.tics.2006.03.001

Kühn, S., & Lindenberger, U. (2016). Research on human plasticity in adulthood: A lifespan agenda. In K. W. Schaie & S. L. Willis (Eds.), *Handbook of the psychology of aging* (8th ed., pp. 105–123). San Diego, CA: Elsevier.

*Lemke, U., & Zimprich, D. (2005). Longitudinal changes in memory performance and processing speed in old age. *Aging, Neuropsychology and Cognition, 12,* 57–77. http://dx.doi.org/10.1080/13825580590925116

Lindenberger, U. (2001). Lifespan theories of cognitive development. In N. J. Smelser & P. B. Baltes (Eds.), *International encyclopedia of the social and behavioral sciences* (pp. 8848–8854). Oxford, UK: Elsevier Science.

*Lindenberger, U., & Ghisletta, P. (2009). Cognitive and sensory declines in old age: Gauging the evidence for a common cause. *Psychology and Aging, 24,* 1–16. http://dx.doi.org/10.1037/a0014986

Lindenberger, U., Li, S.-C., & Bäckman, L. (2006). Delineating brain-behavior mappings across the lifespan: Substantive and methodological advances in developmental neuroscience. *Neuroscience and Biobehavioral Reviews, 30,* 713–717. http://dx.doi.org/10.1016/j.neubiorev.2006.06.006

Lindenberger, U., Mayr, U., & Kliegl, R. (1993). Speed and intelligence in old age. *Psychology and Aging, 8,* 207–220. http://dx.doi.org/10.1037/0882-7974.8.2.207

Lindenberger, U., & Pötter, U. (1998). The complex nature of unique and shared effects in hierarchical linear regression: Implications for developmental psychology. *Psychological Methods, 3,* 218–230. http://dx.doi.org/10.1037/1082-989X.3.2.218

Lindenberger, U., Singer, T., & Baltes, P. B. (2002). Longitudinal selectivity in aging populations: Separating mortality-associated versus experimental components in the Berlin Aging Study (BASE). *The Journals of Gerontology Series B, Psychological Sciences and Social Sciences, 57,* 474–482. http://dx.doi.org/10.1093/geronb/57.6.P474

Lindenberger, U., von Oertzen, T., Ghisletta, P., & Hertzog, C. (2011). Cross-sectional age variance extraction: What's change got to do with it? *Psychology and Aging, 26,* 34–47. http://dx.doi.org/10.1037/a0020525

Little, T. D., Lindenberger, U., & Nesselroade, J. R. (1999). On selecting indicators for multivariate measurement and modeling with latent variables: When "good" indicators are bad and "bad" indicators are good. *Psychological Methods, 4,* 192–211. http://dx.doi.org/10.1037/1082-989X.4.2.192

Lorenzo-Seva, U., & ten Berge, J. M. F. (2006). Tucker's congruence coefficient as a meaningful index of factor similarity. *Methodology, 2,* 57–64. http://dx.doi.org/10.1027/1614-2241.2.2.57

Lövdén, M., Ghisletta, P., & Lindenberger, U. (2004). Cognition in the Berlin Aging Study (BASE): The first 10 years. *Aging, Neuropsychology and Cognition, 11,* 104–133. http://dx.doi.org/10.1080/13825580490510982

Maxwell, S. E., & Cole, D. A. (2007). Bias in cross-sectional analyses of longitudinal mediation. *Psychological Methods, 12,* 23–44. http://dx.doi.org/10.1037/1082-989X.12.1.23

McArdle, I. J., & Nesselroade, J. R. (1994). Using multivariate data to structure developmental change. In S. H. Cohen & H. W. Reese (Eds.), *Life-span developmental psychology: Methodological contributions* (pp. 223–267). Hillsdale, NJ: Erlbaum.

McArdle, J. J. (1988). Dynamic but structural equation modeling of repeated measures data. In J. R. Nesselroade & R. B. Cattell (Eds.), *The handbook of multivariate experimental psychology* (Vol. 2, pp. 561–614). New York, NY: Plenum Press. http://dx.doi.org/10.1007/978-1-4613-0893-5_17

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology, 38,* 115–142. http://dx.doi.org/10.1037/0012-1649.38.1.115

McArdle, J. J., & Nesselroade, J. R. (2003). *Handbook of psychology: Vol. 3. Growth curve analysis in contemporary psychological research* (pp. 44–480). Chichester, UK: Wiley. http://dx.doi.org/10.1002/0471264385.wei0218

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Author.

Panizzon, M. S., Vuoksimaa, E., Spoon, K. M., Jacobson, K. C., Lyons, M. J., Franz, C. E., . . . Kremen, W. S. (2014). Genetic and environmental influences of general cognitive ability: Is *g* a valid latent construct? *Intelligence, 43,* 65–76. http://dx.doi.org/10.1016/j.intell.2014.01.008

Rabbitt, P. (1993). Does it all go together when it goes? The Nineteenth Bartlett Memorial Lecture. *The Quarterly Journal of Experimental Psy-*

chology A: Human Experimental Psychology, 46, 385–434. http://dx.doi.org/10.1080/14640749308401055

*Rast, P., & Hofer, S. M. (2014). Longitudinal design considerations to optimize power to detect variances and covariances among rates of change: Simulation results based on actual longitudinal studies. *Psychological Methods, 19,* 133–154. http://dx.doi.org/10.1037/a0034524

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models: Applications and data analysis methods* (Vol. 1). London, UK: Sage.

R Core Team. (2016). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Reinert, G. (1970). Comparative factor analytic studies of intelligence throughout the human life span. In L. R. Goulet & P. B. Baltes (Eds.), *Life-span developmental psychology: Research and theory* (pp. 467–484). New York, NY: Academic Press. http://dx.doi.org/10.1016/B978-0-12-293850-4.50023-6

Rhemtulla, M., & Tucker-Drob, E. M. (2011). Correlated longitudinal changes across linguistic, achievement, and psychomotor domains in early childhood: Evidence for a global dimension of development. *Developmental Science, 14,* 1245–1254. http://dx.doi.org/10.1111/j.1467-7687.2011.01071.x

Ritchie, S. J., Tucker-Drob, E. M., Cox, S. R., Corley, J., Dykiert, D., Redmond, P., . . . Deary, I. J. (2016). Predictors of ageing-related decline across multiple cognitive functions. *Intelligence, 59,* 115–126. http://dx.doi.org/10.1016/j.intell.2016.08.007

*Ritchie, S. J., Tucker-Drob, E. M., Starr, J. M., & Deary, I. J. (2016). Do cognitive and physical functions age in concert from age 70 to 76? Evidence from the Lothian Birth Cohort 1936. *The Spanish Journal of Psychology, 19,* E90. http://dx.doi.org/10.1017/sjp.2016.85

Rönnlund, M., Nyberg, L., Bäckman, L., & Nilsson, L. G. (2005). Stability, growth, and decline in adult life span development of declarative memory: Cross-sectional and longitudinal data from a population-based study. *Psychology and Aging, 20,* 3–18. http://dx.doi.org/10.1037/0882-7974.20.1.3

Rovine, M. J., & Molenaar, P. C. M. (1998). The covariance between level and shape in the latent growth curve model with estimated basis vector coefficients. *Methods of Psychological Research, 3,* 95–107.

Salthouse, T. A. (1988). Resource-reduction interpretations of cognitive aging. *Developmental Review, 8,* 238–272. http://dx.doi.org/10.1016/0273-2297(88)90006-8

Salthouse, T. A. (1990). Working memory as a processing resource in cognitive aging. *Developmental Review, 10,* 101–124.

Salthouse, T. A. (1991). *Theoretical perspectives on cognitive aging*. Hillsdale, NJ: Erlbaum.

Salthouse, T. A. (1994). How many causes are there of aging-related decrements in cognitive functioning? *Developmental Review, 14,* 413–437. http://dx.doi.org/10.1006/drev.1994.1016

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103,* 403–428. http://dx.doi.org/10.1037/0033-295X.103.3.403

Salthouse, T. A. (2004a). What and when of cognitive aging. *Current Directions in Psychological Science, 13,* 140–144. http://dx.doi.org/10.1111/j.0963-7214.2004.00293.x

Salthouse, T. A. (2004b). Localizing age-related individual differences in a hierarchical structure. *Intelligence, 32,* 541–561. http://dx.doi.org/10.1016/j.intell.2004.07.003

Salthouse, T. A. (2009). When does age-related cognitive decline begin? *Neurobiology of Aging, 30,* 507–514. http://dx.doi.org/10.1016/j.neurobiolaging.2008.09.023

Salthouse, T. A. (2016). Little relation of adult age with cognition after controlling general influences. *Developmental Psychology, 52,* 1545–1554. http://dx.doi.org/10.1037/dev0000162

Sattler, C., Wahl, H. W., Schröder, J., Kruse, A., Schönknecht, P., Kunzmann, U., . . . Rammelsberg, P. (2015). Interdisciplinary Longitudinal Study on Adult Development and Aging (ILSE). In N. A. Pachana (Ed.), *Encyclopedia of geropsychology*. Singapore: Springer. http://dx.doi.org/10.1007/978-981-287-082-7

Schaie, K. W. (1962). A field-theory approach to age changes in cognitive behavior. *Vita Humana, 5,* 129–141.

Schaie, K. W. (1994). The course of adult intellectual development. *American Psychologist, 49,* 304–313. http://dx.doi.org/10.1037/0003-066X.49.4.304

Sibbett, R. A., Russ, T. C., Pattie, A., Starr, J. M., & Deary, I. J. (2018). Does incipient dementia explain normal cognitive decline determinants? Lothian Birth Cohort 1921. *Psychology and Aging, 33,* 674–684. http://dx.doi.org/10.1037/pag0000241

*Sliwinski, M., & Buschke, H. (2004). Modeling intraindividual cognitive change in aging adults: Results from the Einstein Aging Studies. *Aging, Neuropsychology and Cognition, 11,* 196–211. http://dx.doi.org/10.1080/13825580490511080

*Sliwinski, M. J., Hofer, S. M., & Hall, C. (2003). Correlated and coupled cognitive change in older adults with and without preclinical dementia. *Psychology and Aging, 18,* 672–683. http://dx.doi.org/10.1037/0882-7974.18.4.672

Spearman, C. (1904). "General intelligence," objectively determined and measured. *The American Journal of Psychology, 15,* 201–292. http://dx.doi.org/10.2307/1412107

Spearman, C. (1927). *The abilities of man*. London, UK: Macmillan.

Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods, 5,* 60–78. http://dx.doi.org/10.1002/jrsm.1095

Tetens, J. N. (1777). *Philosophische Versuche über die menschliche Natur und ihre Entwicklung* [Philosophical essays on human nature and its development]. Leipzig, Germany: Weidmanns Erben und Reich.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago, IL: University Press.

Tucker-Drob, E. M. (2009). Differentiation of cognitive abilities across the life span. *Developmental Psychology, 45,* 1097–1118. http://dx.doi.org/10.1037/a0015864

*Tucker-Drob, E. M. (2011a). Global and domain-specific changes in cognition throughout adulthood. *Developmental Psychology, 47,* 331–343. http://dx.doi.org/10.1037/a0021361

*Tucker-Drob, E. M. (2011b). Neurocognitive functions and everyday functions change together in old age. *Neuropsychology, 25,* 368–377. http://dx.doi.org/10.1037/a0022348

Tucker-Drob, E. M. (2013). How many pathways underlie socioeconomic differences in the development of cognition and achievement? *Learning and Individual Differences, 25,* 12–20. http://dx.doi.org/10.1016/j.lindif.2013.01.015

Tucker-Drob, E. M., & Briley, D. A. (2014). Continuity of genetic and environmental influences on cognition across the life span: A meta-analysis of longitudinal twin and adoption studies. *Psychological Bulletin, 140,* 949–979. http://dx.doi.org/10.1037/a0035893

Tucker-Drob, E. M., Briley, D. A., & Harden, K. P. (2013). Genetic and environmental influences on cognition across development and context. *Current Directions in Psychological Science, 22,* 349–355. http://dx.doi.org/10.1177/0963721413485087

Tucker-Drob, E. M., Briley, D. A., Starr, J. M., & Deary, I. J. (2014). Structure and correlates of cognitive aging in a narrow age cohort. *Psychology and Aging, 29,* 236–249. http://dx.doi.org/10.1037/a0036187

Tucker-Drob, E. M., Johnson, K. E., & Jones, R. N. (2009). The cognitive reserve hypothesis: A longitudinal examination of age-associated declines in reasoning and processing speed. *Developmental Psychology, 45,* 431–446. http://dx.doi.org/10.1037/a0014012

*Tucker-Drob, E. M., Reynolds, C. A., Finkel, D., & Pedersen, N. L. (2014). Shared and unique genetic and environmental influences on

aging-related changes in multiple cognitive abilities. *Developmental Psychology, 50,* 152–166. http://dx.doi.org/10.1037/a0032468

Tucker-Drob, E. M., & Salthouse, T. A. (2011). Individual differences in cognitive aging. In T. Chamorro-Premuzic, S. von Stumm, & A. Furnham (Eds.), *The Wiley-Blackwell handbook of individual differences* (Vol. 3, pp. 242–267). Chichester, UK: Wiley. http://dx.doi.org/10.1002/9781444343120.ch9

Unverzagt, F. W., Guey, L. T., Jones, R. N., Marsiske, M., King, J. W., Wadley, V. G., . . . Tennstedt, S. L. (2012). ACTIVE cognitive training and rates of incident dementia. *Journal of the International Neuropsychological Society, 18,* 669–677.

van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113,* 842–861. http://dx.doi.org/10.1037/0033-295X.113.4.842

Verhaeghen, P. (2013). *The elements of cognitive aging: Meta-analyses of age-related differences in processing speed and their consequences*. New York, NY: Oxford University Press. http://dx.doi.org/10.1093/acprof:oso/9780195368697.001.0001

Verhaeghen, P., & Salthouse, T. A. (1997). Meta-analyses of age-cognition relations in adulthood: Estimates of linear and nonlinear age effects and structural models. *Psychological Bulletin, 122,* 231–249. http://dx.doi.org/10.1037/0033-2909.122.3.231

Ver Hoef, J. M. (2012). Who invented the delta method? *The American Statistician, 66,* 124–127. http://dx.doi.org/10.1080/00031305.2012.687494

von Oertzen, T., & Brandmaier, A. M. (2013). Optimal study design with identical power: An application of power equivalence to latent growth curve models. *Psychology and Aging, 28,* 414–428. http://dx.doi.org/10.1037/a0031844

Welford, A. T. (1965). Performance, biological mechanisms, and age: A theoretical sketch. In A. T. Welford & J. E. Birren (Eds.), *Behavior, aging, and the nervous system* (pp. 3–20). Springfield, IL: Thomas.

Werner, H. (1948). *Comparative psychology of mental development*. New York, NY: International Universities Press.

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

*Zelinski, E. M., & Stewart, S. T. (1998). Individual differences in 16-year memory changes. *Psychology and Aging, 13,* 622–630. http://dx.doi.org/10.1037/0882-7974.13.4.622

*Zimprich, D., & Martin, M. (2002). Can longitudinal changes in processing speed explain longitudinal age changes in fluid intelligence? *Psychology and Aging, 17,* 690–695. http://dx.doi.org/10.1037/0882-7974.17.4.690

(*Appendices follow*)

Appendix A

Multivariate Models of Longitudinal Changes

Here we provide an overview of the statistical basis of the studies meta-analyzed. These can generally be classified as multivariate growth curve models and multivariate latent difference score models. As described below, both statistical models take similar forms.

Multivariate Growth Curve Models

A multivariate growth curve model for a set of cognitive outcomes measured repeatedly over time can be written as:

$$Y[t]_{\mathrm{w,n}} = i_{w,n} + A[t]_{w,n} \cdot s_{w,n} + e[t]_{w,n}, \tag{A1}$$

where $Y[t]_{\mathrm{w,n}}$ is the score Y of person n on variable w at time t; $i_{w,n}$ is the level for person n on variable w; $s_{w,n}$ is the longitudinal slope for person n on variable w, and $e[t]_{w,n}$ is a disturbance for person n on variable w at time t. The term $A[t]_{w,n}$ represents a set of growth curve basis coefficients that define the shape of the longitudinal changes over time. In the common case of linear growth curve modeling, these basis coefficients can be set to the amount of time (e.g., in years) that has passed between baseline measurement (e.g., the first wave of assessment) and each subsequent assessment for person n on variable w, or they can be set to the age of individual n at each assessment on variable w (e.g., in years, often centered relative to a meaningful early age e.g., 21 years for early adulthood or 65 years for the beginning of old age). For time-based modeling, it is common for the basis coefficients $A[t]$ to be specified using the average or target time between assessment waves rather than the person-specific time-lags. The basis coefficients need not be specified to be linear over time or age. Longitudinal changes may be better represented as occurring nonlinearly, in which case the basis coefficients $A[t]$ may be specified as nonlinear functions of time or age. Additional growth curve slopes, for example, slopes representing quadratic components of change over time or retest effects may also be specified using separate sets of basis coefficients. In the current meta-analysis, we were constrained by the original modeling decisions of the authors of the primary studies (e.g., whether to specify basis coefficients as functions of age or time, or whether to include additional growth curve slopes beyond those representing linear change).[5]

Multivariate Latent Difference Score Curve Models

A multivariate latent difference score model, in which several sets of cognitive outcomes are each measured at both baseline and a single follow-up wave, can be written in a similar form as the multivariate growth model above:

$$Y[t]_{\mathrm{w,z,n}} = \lambda_{w,z} \cdot i_{w,n} + t \cdot \lambda_{w,z} \cdot s_{w,n} + e[t]_{w,z,n}, \tag{A1b}$$

where $Y[t]_{\mathrm{w,z,n}}$ is the score Y of person n on indicator z of latent variable w at time t (0 for baseline and 1 for follow-up wave; or to scale the difference score in annual units, 0 for baseline and the average or target number of years since baseline for follow-up wave), and $\lambda_{w,z}$ is a time-invariant factor loading relating the latent variable w to indicator z. As in the multivariate growth model, $i_{w,n}$ is the level for person n on (now latent) variable w, and $s_{w,n}$ is the longitudinal change over time for person n on variable w. Whereas error in the growth curve model is separated from true change by virtue of being modeled as disturbance from a systematic function of change over time, error in the latent difference score model is separated from true scores psychometrically by specifying indicators to be functions of common factor w and time-specific indicator-specific errors, $e[t]_{w,z,n}$, which represent a combination of indicator-specific true score variability and error of measurement.

Level and Change Covariance Structure

Regardless of whether longitudinal changes are modeled using a multivariate growth curve approach or a multivariate latent difference score approach, the person-specific levels and slopes for each variable can be allowed to freely covary with one another as follows:

$$\mathrm{cov}(i_{w,n}, s_{w,n}) = \begin{bmatrix} \sigma^2_{i_{1,n}} & & & & & & & \\ \sigma_{i_{1,n},i_{2,n}} & \sigma^2_{i_{2,n}} & & & & & & \\ \vdots & \vdots & \ddots & & & & & \\ \sigma_{i_{1,n},i_{w,n}} & \sigma_{i_{2,n},i_{w,n}} & \cdots & \sigma^2_{i_{w,n}} & & & & \\ \sigma_{i_{1,n},s_{1,n}} & \sigma_{i_{2,n},s_{1,n}} & \cdots & \sigma_{i_{w,n},s_{1,n}} & \sigma^2_{s_{1,n}} & & & \\ \sigma_{i_{1,n},s_{2,n}} & \sigma_{i_{2,n},s_{2,n}} & \cdots & \sigma_{i_{w,n},s_{2,n}} & \sigma_{s_{1,n},s_{2,n}} & \sigma^2_{s_{2,n}} & & \\ \vdots & \vdots & & \vdots & \vdots & \vdots & \ddots & \\ \sigma_{i_{1,n},s_{w,n}} & \sigma_{i_{2,n},s_{w,n}} & \cdots & \sigma_{i_{w,n},s_{w,n}} & \sigma_{s_{1,n},s_{w,n}} & \sigma_{s_{2,n},s_{w,n}} & & \sigma^2_{s_{w,n}} \end{bmatrix} \tag{A2a}$$

This covariance matrix can be written more compactly as a joint covariance matrix of levels and slopes, as follows:

$$\mathrm{cov}(i_{w,n}, s_{w,n}) = \Sigma_{i_{w,n},s_{w,n}} = \begin{bmatrix} \Sigma_i & \Sigma_{is}^{\mathrm{T}} \\ \Sigma_{is} & \Sigma_s \end{bmatrix} \tag{A2b}$$

[5] There was a small number of papers included in the current meta-analysis that specified growth curves involving nonlinear slopes in addition to linear slopes. For instance, Ghisletta et al. (2012) estimated linear and quadratic growth curve slopes, but because the quadratic slopes were not significant, they did not estimate their covariances across abilities. Additionally, both Ghisletta et al. (2012) and Tucker-Drob (2011a) estimated additional nonlinear slopes representing practice effects, which were not a direct focus of the current meta-analysis. Many studies simply estimated linear growth curve models. Thus, we exclusively coded the relevant data from the linear components of longitudinal changes.

(*Appendices continue*)

Key components of this covariance matrix are: a submatrix Σ_i that includes level variances on its diagonal and level-level covariances off its diagonal, a submatrix Σ_s that includes slope variances on its diagonal and slope-slope covariances off its diagonal, and a submatrix Σ_{is} that includes within-variable level-slope covariances on its diagonal and cross-variable level-slope covariances off its diagonal. When this covariance matrix is freely estimated, it is sometimes referred to as a "parallel process" model.

Higher-Order Factors of Curves

Rather than allowing the person-specific levels and slopes for each variable to freely covary, as in Equations A2a and A2b above, the interrelations among levels and among slopes can be approximated by common factors. The factor portions of such a "factor of curves" model can be written as:

$$i_{w,n} = \tau_{i_w} + \lambda_{i_w} \cdot F_{i,n} + u_{i_{w,n}} \tag{A3a}$$

and

$$s_{w,n} = \tau_{s_w} + \lambda_{s_w} \cdot F_{s,n} + u_{s_{w,n}}, \tag{A3b}$$

where τ_{i_w} and τ_{s_w} are the mean level and slope for variable w, λ_{i_w} is the loading of the person-specific level of variable w on the common factor of the levels, $F_{i,n}$; λ_{s_w} is the loading of the person-specific slope of variable w on the common factor of the slopes, $F_{s,n}$; $u_{i_{w,n}}$ is a person-specific unique factor for the level of variable w, and $u_{s_{w,n}}$ is a person-specific unique factor for the slope of variable w. Typically, within-variable covariances between level and slope unique factors are freely estimated; all remaining covariances among uniqueness are fixed to 0, and the level and slope common factors are assumed to have means of 0 and are allowed to freely covary with one another.

The approximation of level and slope covariances by the higher-order factor of curves model is given in matrix notation as:

$$\Sigma_{i_{w,n}, s_{w,n}} \approx \Lambda \Psi \Lambda' + \Theta, \tag{A4}$$

where Λ is a $2w \times m$ matrix of loadings for the levels and slope of each variable w on m common factors (typically two factors: one common factor of levels and one common factor of slopes), Ψ is an $m \times m$ covariance matrix of the common factors, and Θ is a $2w \times 2w$ covariance matrix of level and slope unique factors (typically with unique factor variances on the diagonals, within-variable level-slope covariances freely estimated, and between-variable level-slope covariances fixed to zero).

Communality is the term used in factor analysis for the proportion of variance in a variable that is explained by a common factor. In the context of a factor of curves model, we are specifically interested in the proportions of variance in the growth curve levels and slopes that are explained by the common factors of levels and slopes respectively. Communality can be computed as

$$\text{Communality} = \frac{\text{shared variance}}{\text{total variance}} = \frac{\text{shared variance}}{\text{shared variance} + \text{unique variance}} = \frac{\lambda^2 \cdot \sigma_F^2}{\lambda^2 \cdot \sigma_F^2 + \sigma_u^2}, \tag{A5}$$

where λ is the common loading of the variable-specific level or slope on the common factor of levels or slopes, σ_F^2 is the variance of that common factor, and σ_u^2 is the variance of the unique factor of the variable-specific level or slope. For cases in which the factor loading has already been standardized (such that the standardized variance of the factor is 1 and the total variance in the outcome is 1), the communality is simply calculated as the square of that standardized factor loading. Level and slope communalities are the primary outcomes of the current meta-analysis.

Appendix B

Computation of Precision Weights

The raw meta-analytic precision weights, w_i, were computed as:

$$w_i = r_{i,x} \cdot r_{i,y} \cdot N_i, \tag{B1}$$

where N_i is the complete sample size (at the first wave), and $r_{i,x}$ and $r_{i,y}$ are the reliabilities of the i-th pair of growth curves (or difference scores, respectively) analyzed. The reliabilities were each downweighted by the observed missing data patterns using the theorem in Appendix B of von Oertzen and Brandmaier (2013) to properly account for attrition. By definition, the weights are always positive. Raw weights were converted to scaled weights, such that they sum to the total number of effect sizes included in the metaregression model.

As described in the Method section (under Analytic Approach: Multilevel Metaregression Models), precision weights were implemented in combination with downweights that adjusted for the number of effect sizes contributed per sample. Precision weights were specified at the level of the individual effect sizes (i.e., as *within*-cluster weights) and downweights were specified at the level of the contributing samples (i.e., as *between*-cluster weights). For multilevel metaregression models, between-cluster weights were scaled such that the products of the within-cluster and between-cluster weights summed to the total number of effect sizes in the meta-analytic dataset.

Received August 24, 2017
Revision received August 22, 2018
Accepted October 8, 2018 ■