Published in final edited form as:

Multivariate Behav Res. 2014 ; 49(1): 1–16. doi:10.1080/00273171.2013.841089.

Fighting for Intelligence: A Brief Overview of the Academic Work of John L. Horn

John J. McArdle and

Department of Psychology, University of Southern California

Scott M. Hofer

Department of Psychology, University of Victoria, British Columbia, Canada

Abstract

John L. Horn (1928–2006) was a pioneer in multivariate thinking and the application of multivariate methods to research on intelligence and personality. His key works on individual differences in the methodological areas of factor analysis and the substantive areas of cognition are reviewed here. John was also our mentor, teacher, colleague, and friend. We overview John Horn's main contributions to the field of intelligence by highlighting 3 issues about his methods of factor analysis and 3 of his substantive debates about intelligence. We first focus on Horn's methodological demonstrations describing (a) the many uses of simulated random variables in exploratory factor analysis; (b) the exploratory uses of confirmatory factor analysis; and (c) the key differences between states, traits, and trait-changes. On a substantive basis, John believed that there were important individual differences among people in terms of cognition and personality. These sentiments led to his intellectual battles about (d) Spearman's *g* theory of a unitary intelligence, (e) Guilford's multifaceted model of intelligence, and (f) the Schaie and Baltes approach to defining the lack of decline of intelligence earlier in the life span. We conclude with a summary of John Horn's unique approaches to dealing with common issues.

John L. Horn (1928–2006) was a pioneer in multivariate thinking and in the application of multivariate methods to understanding the structure and determinants of change and variation in aspects of intelligence and personality. As a leader in multivariate methodology, he tried to reach the incredible heights of his well-known mentor, Raymond B. Cattell. The title of this tribute is meant to reflect favorably on John Horn as a mentor, teacher, colleague, and friend to many of us as he consistently challenged us all to think longer, harder, and better. As well, he was a welterweight boxing champion early in his life (in the state of Colorado) and this may help to explain his fighting approach to more scholarly activities. We are well aware that his fighting style was in part due to an acerbic personality, but we emphasize here the scholarly approach that was a big part of his life. We believe he was mainly “fighting smart” and cared a lot about the appropriate collection, analysis, and interpretation of research data, and he was not “just fighting” because he liked to do so. His key academic contributions to the methodological areas of factor analysis and the analysis of change and to the substantive areas of the structure of cognitive abilities and cognitive

development and aging are reviewed. We aim to highlight the ideas and reasons that he fought for consistently across his career (see Figure 1).

To provide structure to this review, we discuss John Horn's main scholarly contributions by highlighting (a) three issues with the methods of factor analysis and (b) three substantive debates about intellectual ability theory. As we illustrate here, John believed strongly in a multivariate scientific approach, and he questioned the typical use of unweighted sum scores as if they represented the best scores of the psychological constructs of interest. We first focus on Horn's methodological demonstrations describing (a) the many uses of simulated random variables in exploratory factor analysis; (b) uses of confirmatory factor analysis procedures that are based on exploratory logic; and (c) design and analysis for distinguishing states, traits, and trait-changes. On a substantive basis, John surely believed that there were important individual differences among adults within the domains of cognition and personality. These sentiments led to (d) challenges to Spearman's *g* theory of a unitary intelligence, (e) arguments with Guilford's multifaceted model of intelligence, and (f) arguments against the prominent Schaie and Baltes approach to defining the lack of decline of intelligence earlier in the life span. There are many other areas of multivariate research in which John Horn was involved, but we think these represent some of his key contributions.

Although John Horn contributed to many areas of psychology, including wide-ranging topics in personality and alcohol abuse research and in child development later in his career, we think he is best known for his willingness to engage in debates about the methodology of factor analysis and the scientific evidence for the structure and change in cognitive abilities. Although aspects of these debates linger on, for the most part, we try to show how the written ideas of John Horn continue to prove useful in this regard. Much can be said about John as an atypical person with an atypical background, but we do not emphasize this here (see McArdle, 2007a). To say John Horn was a fighter is probably an understatement, but he certainly liked to think he was fighting for the underdog position not just for the less popular or possibly inconvenient ideas in a "knee-jerk" or "faddish" way but always from a vigorously well-reasoned and logical approach. So, following his lead, we do not simply review this work with the typical aplomb and applause that he probably deserves. Instead we review this prior work with a critical eye and consider whether his classical work is still useful nowadays.

SOME EARLY AND LATER SENTIMENTS

Some of John Horn's early comments on the methods of factor analysis are worth repeating:

The central concept of the multivariate metatheory is that of functional unity, as defined by Cattell (1950, 1957). A functional unity is indicated by a configuration—a pattern—of processes which are themselves distinct but are integral parts of a more general process. To distinguish a functional unity from a mere collection of distinct processes it is necessary to show that the processes work together—that they rise together, fall together, appear together, disappear together or, in general, covary together. The action of the heart may be taken as an example of a kind of functional unity. The processes which might indicate this function are those recorded by blood pressure measurements taken in various parts of the body,

breathing rate, skin moisture, and skin temperature. Each of the processes indicated by such measurements is distinct from the other and the intercorrelations among the process measurements will be less than unity even when corrected for unreliability. Yet the several processes work together in a way that gives evidence of the vital function and the function may be indicated by a pattern of covariation among the measurement of the separate processes. ...A functional unity is thus a rather high level of abstraction. In operational terms it can be defined in a number of ways. One of the simplest such definition is that furnished by the methods of factor analysis. (Horn, 1972, pp. 161–162)

When these ideas about a multivariate metatheory are applied to data on cognitive abilities, we can create various testable hypotheses:

Similarly, the function of intelligence may be indicated by a pattern of covariance among distinct processes of memory, reasoning, abstracting, etc. (Horn, 1972, p. 162)

This type of reasoning provides the basis for arriving at several substantive results on cognitive abilities, perhaps the most important being that

intellectual abilities are organized at a general level into two general intelligences, viz., fluid intelligence and crystallized intelligence. These represent the operation of somewhat different—i.e. independent—influences in development. On the one hand there are those influences which directly affect the physiological structure upon which intellectual processes must be constructed—influences operating through the agencies of heredity and injury: these are most accurately reflected in measures of fluid intelligence. And on the other hand there are those influences which affect physiological structure only indirectly through agencies of learnings, acculturations, etc.: crystallized intelligence is the most direct resultant of individual differences in these influences. (Horn & Cattell, 1967, p. 109)

Horn expanded on this initial work of his primary advisor, Raymond B. Cattell (see Horn, 1965a; Horn & Cattell, 1966a, 1966b, 1967), to identify additional functional unities of primary mental abilities in terms of quantitative, visual, auditory, memory, speed-of-processing kinds of intelligence (e.g., Horn, 1968; see also Horn & Hofer, 1992; Horn & Noll, 1997; McGrew, 2005, for summaries). In fact, Horn had thought of so many important cognitive processes that he started to use capital letters to indicate the many functions (i.e., G_v instead of g_v). This historical point of view is described by his own diagrams (see Figure 2; see also Horn, 1970).

What seems to be Horn's final theoretical work (Horn & Masunaga, 2000; Horn & McArdle, 2007) suggested that “within the domains of expertise, high levels of reasoning, feats of memory and speeded thinking similar to G_f are displayed by older adults” (Masunaga & Horn, 2000, p. 5). Although Ericsson and Kintsch (1995) seem to be the first to describe this type of wide-span memory, Masunaga and Horn (2000, 2001) provided further empirical tests and made it part of what they termed “extended G_f - G_c theory.” The development of these ideas challenges the typical application of typical methods. Late life G_c , for example, does not require retention of fluid abilities but creates, in the area of

expertise, a kind of wide-span memory that permits large amounts of information to be brought into immediate memory, enabling the application of this knowledge to problems at hand (see Horn & Blankson, 2005). He seemed to think that adults were far more heterogeneous than children.

Important here was the development of “adult expertise reasoning,” seemingly dependent on wide-span memory. This enables experts, such as adults in major positions of responsibility in our culture, to reason at a higher level than people who depend primarily on fluid reasoning. This part of the theory is most important for understanding why so much of our culture is in the hands of aged adults and also what we can and should do about it. It is not surprising these are not simple ideas but very broad all-encompassing issues, and they have been written about by many others in many forums (e.g., McGrew, 2005, 2009). Some of these sentiments are represented in Horn’s overview of the work of Cattell (see Horn, 1984)—here Horn demonstrates both admiration and skepticism about Cattell’s broad multivariate theories of intelligence and personality.

SIMULATION OF FACTOR ANALYSIS AS A RESEARCH TOOL

Horn became an expert in the classical techniques of common factor analysis. He generally asked why any elegant mathematical-statistical theory should be based on specific assumptions when we know these key assumptions are wrong and untestable. One of his most important papers suggested that the number of common factors should not be determined simply using the well-known “eigenvalues greater than one” or “root one” criterion defined by one of his favorite advisors, Henry Kaiser (1960; see Horn, 1965a; also see McArdle, 2007a). Instead, Horn suggested we make use of “computer-simulation” techniques, mainly based on model assumptions that were more realistic, and he advocated their use whenever possible (see Horn & McArdle, 1980).

Early in his work Horn suggested we use what he then termed “Parallel Analysis” (1965a) to determine the number of common factors by selecting the number of the eigenvalues of a correlation matrix that were greater than or equal to those provided by data simulated with known characteristics with a modern computer. On Parallel Analysis he wrote,

It is suggested that if Guttman’s latent-root-one lower bound estimate for the rank of a correlation matrix is accepted as a psychometric upper bound, following the proofs and arguments of Kaiser and Dickman, then the rank for a sample matrix should be estimated by subtracting out the component in the latent roots which can be attributed to sampling error, and least-squares capitalization on this error, in the calculation of the correlations and the roots. A procedure based on the generation of random variables is given for estimating the component which needs to be subtracted. (p. 179)

The essential basis of this work is depicted in Figure 3 (from Horn, 1965a). In this utterly simple idea, all we needed to do was generate “random data of similar size” and we could calculate the latent roots and vectors of these random data to provide a criterion tailored to the particular data set. He implemented this procedure for $N = 297$ persons from his doctoral dissertation research (see Horn, 1965b) where he thought he found evidence for between K

= 9 and $K = 16$ common factors. Of course, we now recognize the need for statistical fluctuations in these latent roots, but these kinds of calculations were done almost 50 years ago! Horn's random variable approach has more recently been found to be the most accurate for determining the number of unrotated common factors (e.g., Ledesma & Valero-Mora, 2007; Montanelli & Humphreys, 1976; Velicer, Eaton, & Fava, 2000; Zwick & Velicer, 1986; see also Dinno, 2009; Hayton, 2009, for recent evaluation and affirmation). This also marks the beginning of Horn's fascination with the use of computer-simulated data to solve the most complex problems in mathematical statistics.

Other features of factor analysis that Horn investigated with simulation procedures included the well-known "rotation" problem (see Horn, 1967b). Here he questioned the use of popular forms of factor rotation that were clearly founded on "substantive judgment" but that seemed to be considered "objective" largely because they were "blind" (p. 813). He knew that

a sample of 300 is fairly large relative to samples used in substantive experiments, and the correlations expected by chance in this sample are correspondingly small; hence, the factor loadings in these analyses can be small, as obtained. But if the N were reduced, factor loadings would average larger and thus appear to be more "respectable." (p. 819)

Horn generated entirely random variables ($M = 74$) for this size sample ($N = 300$), and he found a clear willingness of famous faculty members (presumably several expresidents of the American Psychological Association at the University of Illinois) to assign what seemed to be reasonable labels to several common factors that were no more than random rotations of simulated or "random" data. Rather than create more controversy, the reader was largely left to judge what this meant about the available factor analysis procedures and what should be done next.

Horn wrote a great deal about his simulation studies of factor analysis, and he focused considerable attention on "factor score estimates" (Horn, 1965a; Horn & Miller, 1966; Wackwitz & Horn, 1971). He knew that factor analysis was based on observed scores and he tried to provide a vehicle to calculate the unobserved (common factor) scores from these. Several factor score estimation approaches were examined using "exceedingly simple" (Wackwitz & Horn, 1971, p. 406) simulated data. These authors somewhat surprisingly concluded that "inexact procedures," such as the use of "unit-weighted salient variables" (but not just all unit-weighted variables) led to the most replicable estimates for common factor scores. Although others have suggested further practical advances (e.g., DiStefano, Zhu, & Mindlna, 2009; Grice, 2001; Grice & Harris, 1998), there has become much less doubt about the procedures we can use effectively. Indeed, the suggested use of a simple common factor score has become a part of our understanding of optimal factor scores today.

This computer-simulation approach has certainly increased in popularity in psychometric and statistical research, and of course Horn was not alone in understanding these topics (e.g., Ciesla, Cole, & Steiger, 2007; L. K. Muthén & Muthén, 2005). This approach has been extended to estimation of standard errors (see Efron, 1979) and Bayesian estimation (e.g., Markov Chain Monte Carlo; see Ntzoufras, 2009) and even structural equation models (see

Horn & McArdle, 1980). It appears that many scientists now agree that statistical analysis using simulations of otherwise highly complex systems is a viable approach to data analysis.

EXPLORATORY USES OF CONFIRMATORY FACTOR ANALYSIS

In order to understand the popular move toward what Tucker and Lewis (1973) termed “confirmatory analysis,” Horn first related his early work on Parallel Analysis to the development of the well-known chi-square test of the number of factors (Horn & Engstrom, 1979). Here he showed that his simulation approach matched the formal basis of Bartlett’s (1950) chi-square test, and he was pleased. Next he took on the arbitrary use of rotations in confirmatory factor analysis (CFA), and it was clear that he was opposed to the term “confirmatory” for models that were “exploratory” at best (see McArdle, 2012a). He suggested the idea of simulating complex systems that would not be possible to uniquely identify due to the selection of variables and persons (as in Horn & McArdle, 1980). Although he did not provide clear solutions to these problems, he was basically trying to point out that there were no available solutions!

Horn’s main work on this topic can be summarized by the $M = 25$ variable example presented initially by McArdle and Horn (1981). These data were collected on $N = 297$ persons as part of his doctoral dissertation (at the University of Illinois under R. B. Cattell). The explicit intention was to represent at least five broad factors of intellectual ability, including Fluid Intelligence (i.e., Gf ; reasoning in novel situations), Crystallized Intelligence (i.e., Gc ; using information from the dominant culture to solve problems), General Visualization (i.e., Gv ; seeing solutions to visual problems as part of immediate awareness), General Speed of Processing (i.e., Gs ; solving problems quickly), and General Retrieval (i.e., Gr ; extracting fundamentals of solutions from memory). Although the first two factors, Gf and Gc , were clearly based on the work of Cattell (1941, 1971), the others included were novel features of the factorial hypothesis at the time, and these are depicted in Figure 4 (from Horn, Donaldson, & Engstrom, 1981). Basically, he was determined to show that Cattell was incorrect because he did not go far enough. That is, there were more than two common factors needed to describe these data.

In this new analysis, an interesting debate was created between McArdle and Horn about “which direction the arrows were supposed to go” to mimic standard factor analysis and “should there be one higher order factor?” For McArdle, the arrows in a factor analysis were supposed to begin at the latent variable and end at the observed variables, mainly because this is how the LISREL model was fitted. But for Horn, the arrows in a factor analysis were supposed to go up the chain from observation to hypothesis because this is how information actually flowed in the cognitive system. Figure 4, from Horn et al. (1981), shows Horn’s view regarding the direction of the arrows. When the first model was fit to the original dissertation data for presentation at the 1981 American Psychological Association Convention (see <http://kiptron.usc.edu>), using the then newly available LISREL-II program (Jöreskog & Sörbom, 1979), the basic model results were obtained suggesting a good fit to the observed data. Incredibly (at least to McArdle) the key debate about proper arrow diagramming seemed to end rather quickly. Incidentally, all models presented in this section

have recently been refit using current programs (e.g., Mplus 6.0; L. K. Muthén & Muthén, 1998–2012), and all numerical results are essentially the same.

But several higher order aspects of these results are worth reconsidering. A five-factor highly restricted model fit was deemed acceptable ($X^2 = 180$, $df = 80$). Because the probability basis of this model required various assumptions that did not seem testable or tenable (i.e., normality of uniquenesses), no probability value was offered at the time (1981). Recall also that at this time, other nonprobability indicators of misfit were not yet popular (i.e., RMSEA; Browne & Cudeck, 1989, 1993; Steiger, 1990; Steiger & Lind, 1980). Nevertheless, this initial solution was contrasted to a model where there was only a single common factor for the same data ($X^2 = 685$, $df = 90$), and this simpler model was soundly rejected. Seven additional factor loadings were added based on earlier work by Horn (1965a) and, although this made the factorial solution more complex, this also seemed to make the fit better ($X^2 = 224$, $df = 7$). This subjective difference in this kind of model comparison eventually led to the consideration of better ways to examine alternative arrays of models (e.g., McArdle & Nesselroade, 1994, 2012).

Of course, CFA-type analyses could have stopped at this point, but Horn was curious—very curious—about exactly what could be mathematically identified and what could not be mathematically identified. That is, he wanted to know what was different about these CFA models beyond what was already known in exploratory factor analysis (EFA) work (e.g., Lederman's [1937] limits). For example, he asked if the common factor G could be represented as a second-order factor (as in Thurstone, 1947). To examine this hypothesis, McArdle and Horn fit a model where the correlations among the common factors had second-order factor loadings. Of course, there was no obvious way to fit such a model using the standard LISREL notation, so he simply pushed his coworkers to develop and use a new structural notation (RAM; see Horn & McArdle, 1980; McArdle, 2005; McArdle & McDonald, 1984). The second-order factor model was fit simultaneously with the same first-order pattern and compared with the first-order model, and it still left something to be desired ($X^2 = 55$, $df = 5$).

Any ordinary scientist today might stop at this point, but Horn certainly was not an ordinary scientist. Instead he asked if a second-order G factor would be reasonable if the first-order factors were not as rigid in their orientation. In this way, he asked if we could use the second-order factor pattern as a rotational device for the first-order factor pattern. To our knowledge this kind of approach had not been considered before. Although it was a bit difficult to see exactly where Horn was going with this logic, several exactly identified models were fit, and the overall fit was certainly improved. This started with an unrestricted (exactly identified) five-factor model that achieved very good fit ($X^2 = 60$, $df = 40$). This continued with an unrestricted (exactly identified) two-factor higher order solution that seemed very similar in fit ($X^2 = 68$, $df = 45$). A variety of other lower order and higher order combinations were tried, including an overidentified second-order factor model with an exactly identified first-order model.

Horn thought that this type of solution was much closer to his original EFA. Relationships among many demographic variables were added, including possible nonlinear relations

among the common factors (e.g., McArdle & Prescott, 1992), and the numerical results suggested some limits to the kinds of relations that could be fitted (i.e., external relationships to the higher order factor eliminated the possibility of one relationship to the lower order factor). These results were reported at the time. Horn's early ideas about exploratory uses of confirmatory techniques have recently become much more popular (e.g., Marsh et al., 2010; Marsh et al., 2009; McArdle & Cattell, 1994; B. Muthén & Asparouhov, 2012).

Horn finished his work on the CFA approach with a great regard for the computer programs and techniques (e.g., LISREL) but with a continued skepticism for those that claimed the basic ideas for a factorial hypothesis preceded any common factor analysis. He seemed to advocate the use of EFA, partly because it was a more honest approach than the CFA (see McArdle, 2012a). More important, in Horn's view, researchers were typically in the position of exploring multivariate data with restricted structures but were trying to express these misfits using the probabilistic statements recently made available for a priori hypotheses (see McArdle, 2012a). Although Horn certainly advocated the features of restriction across factor structures, he found that most tests of *metric factorial invariance over groups* (see Meredith, 1993) were carried out in a far less than scientific fashion, so he favored "configural" analysis as a practical outcome (see Horn & McArdle, 1992). This paper represented another unscheduled debate by Horn and McArdle, but here each author simply contributed his own subsection (Horn was first, McArdle second). In an important sense, Horn was fighting against the publication of theoretical models based on an exploration of data with factor analysis as if these were confirmatory in nature, just because it looked so much better and was easier to understand (see McArdle, 2012a). This skepticism can also be seen in his penultimate work on longitudinal factor analysis (see Meredith & Horn, 2001).

STATE, TRAIT, AND TRAIT-CHANGE DIMENSIONS OF COMMON FACTORS

One of the major questions asked by Horn concerned the optimal way to deal with repeated measures of multivariate phenomena. A number of papers on this topic (e.g., Horn & Little, 1966) were based on the methodological concepts of "level, scatter, and shape," originally recognized by yet another of his academic mentors, Lee J. Cronbach (see Cronbach & Gleser, 1953). But, by far and away, the best statement of his conclusions about this topic, or any other for that matter, was presented in his article "State, Trait, and Change Dimensions of Intelligence" (Horn, 1972).

After a fairly elaborate but worthwhile introduction to multivariate metatheory for both cognition and personality, Horn described an investigation using an intensive measurement design in which he intended to make analytic sense out of Cattell's "Data Box" (1966; depicted in Horn, 1972; see Figure 5). But Horn did a lot more than just draw this picture—Horn operationalized this generic approach by measuring $N \sim 100$ persons for $T = 10$ consecutive days on $M = 10$ cognitive tests.

Horn immediately recognized that these kinds of repeated measures data were difficult data to collect, so he carried out his experiments with a highly cooperative set of captive participants (i.e., prisoners in Colorado). But the methods proposed to analyze these repeated measures data were entirely novel, and this is what turns it into a classic John L.

Horn investigation. He calculated the usual R-type “trait” factors in terms of *Gf-Gc* theory from the averages of scores over all 10 occasions—the person-trait **T** matrix. Next he calculated the common factors of the variations of people over 10 occasions—the within-person **W** matrix—and he called these State factors. In contrast to the approach now popularized and advocated by others (e.g., Steyer, Ferring, & Schmitt, 1992; Steyer, Schmitt, & Eid, 1999; cf. McArdle, 2009), these States are supposedly factors that are independent of the Traits (i.e., not the current State of the organism). That is, the State was strictly defined as the part of the Trait that randomly fluctuates, and this is consistent with the definitions of Cattell (1950a, 1966) and later used by McArdle and Woodcock (1997).

At the same time Horn also defined the “Trait-Changes” as that part of the Trait that displays systematic variation over time. In later work, we tried to recognize these changes using slope factors (see McArdle, 1988, 1991), but Horn (1972) estimated these as trait-changes in the common factors of the $\mathbf{B} = \mathbf{W}^{-1}\mathbf{T}$ as well as a pure **S** Trait-State matrix. In this work, he clearly recognized that this **B** matrix was not necessarily symmetric, so he created novel factorial methods to overcome the resulting computation problems. The definitions and calculations subsequently used by Steyer et al. (1992), Steyer et al. (1999), and McArdle and Woodcock (1997) are a bit different although the broad goals seem very similar.

Horn also recognized that the common factors of these separate matrices were not exactly “restricted to be identical,” and this raises another fundamental issue: are the common factors that describe the overall differences between persons identical to those that describe the differences within a person? (a principle used in McArdle, 2007b; McArdle, Fisher, & Kadlec, 2007; McArdle & Nesselrode, 1994, 2012). Nevertheless, the correlation of these common factor scores (Horn, 1972, Table 3, page 178) seemed to be the only way to examine the identity of these common factors at the time, so “congruence coefficients,” and not strict “equality of loadings,” were what was calculated by Horn. It is now known that the factor loading equality is necessary to establish factorial identification (see Meredith & Horn, 2001), and now this is easy to apply to this problem (as described by McArdle & Nesselrode, 2012). So exactly what was done by Horn (1972) is not the currently optimal way to do this analysis but, once again, Horn was way ahead of his time!

Horn’s results are still novel. He created these four matrices and applied principal components analysis to each. He listed the unrestricted factor patterns for all matrices as well as the correlations among the resulting factor scores using extension analysis techniques (see Horn, 1972, p. 178). Because the results are so similar, he concluded these common factors of different source matrices represent the same constructs. This allowed him to decompose the common factors of *Gf-Gc* theory in terms of (a) State, (b) Trait, and (c) Trait-Change variance. For example, the *Gc* factor was estimated to be 80% Trait, 10% State, and 10% Trait-Change. In contrast, the *Gf* factor was estimated to be 60% Trait, 5% State, and 30% Trait-Change. Indeed, the different patterns of change are a key reason these should be considered as different constructs. At the extreme, the *Gv* factor was estimated to be the most fluctuating—50% Trait, 30% State, and 20% Trait-Change (for further elaboration, see McArdle & Woodcock, 1997).

It is fairly obvious that this kind of analytic treatment of repeated measures does not come from the same source as a more standard MANOVA of Repeated Measures (see Bock, 1975) or even the newer and now popular Latent Curve approach (McArdle, 2009; McArdle & Nesselroade, 2012; Meredith & Tisak, 1990). Horn thought much more could be obtained from individual differences mixed with group differences (e.g., Horn & Little, 1966). But, as far as we know, his difficult-to-collect repeated-measures measurement design had not really been repeated until very recently (e.g., Schmiedek, Lövdén, & Lindenberger, 2010; see also Stawski, Sliwinski, & Hofer, 2013). Nevertheless, these ideas about the novel treatment of repeated measures data are prominent in the current literature (see Boker, Molenaar, & Nesselroade, 2009; McArdle & Woodcock, 1997; Sliwinski, Hoffman, & Hofer, 2010b; cf. Steyer et al, 1999).

THE NUMBER OF DIMENSIONS OF COGNITION AND SPEARMAN'S "G"

As just described, a primary focus of Horn's career had to do with the organization of cognitive capabilities, their development, and the impact of life span factors and aging-related change on different abilities. Horn, in contrast to Spearman (1904) and many other notable researchers (see Gottfredson, 1997; Jensen, 1998; Jensen & Weng, 1994; Johnson, Bouchard, Krueger, McGue, & Gotesman, 2004), did not believe the evidence pointed toward a single factor of intelligence "g"—a unitary construct that is manifest in every form of cognitive functioning (Horn & McArdle, 2007). Note that Spearman's theory is distinct from "collection-of-ability" theories of Binet and Simon (1905) and Jensen (1998) because it requires that the factor be the only common factor and not just the first principal component (Horn & McArdle, 2007). This critical distinction makes Spearman's theory of "g" somewhat different—it is a scientifically testable theory—and this can be contrasted to the current majority acceptance of general intelligence as simply a weighted sum of a set of test scores (i.e., first principal component). Of course, Horn did not believe that the "majority vote rules" in science, and he often quoted the Ptolemaic "geo-centric view" as the largest vote getter of all time (even though this theory was proven to be incorrect by Kepler's elliptical model). In this context, he said, "Spearman's theory requires that that factor be the only common factor, not just the first principal factor" (Horn & McArdle, 2007, p. 206), permitting no covariance among any specific factors in measures used to indicate different kinds of mental effort. Although this strict test of general intelligence had the benefit of being falsifiable, it was indeed falsified as it did not account for the data, although these results are typically overlooked even now (e.g., Gottfredson, 1997; Jensen, 1998; Johnson et al., 2004; but see McArdle, 2012b).

Spearman developed a somewhat different theory in 1927, arguing for the "universality of g" based on the "theorem of indifference of the indicator." This seems like a very reasonable idea. But, as Horn & McArdle (2007, p. 219) noted, "In accordance with this theorem, evidence of positive correlations among abilities was sufficient to support his theory." This was actually a significant step down from the previously falsifiable scientific theory of g and probably provides the basis for the current and prevalent view of general intelligence. It should also be noted that the positive manifold among ability tests can be due to a number of unadjusted artifactual influences due to sample differences in age, maturational level, social class, race, and sex (Hofer, Flaherty, & Hoffman, 2006; Horn, 1965b; Kelley, 1928;

Meredith & Horn, 2001). Although the currently accepted model of general intelligence, as the first principal component of a set of cognitive ability tests, is not falsifiable in the factor analytic sense, there are a number of arguments against a theory of a single factor of intelligence. Horn was consistently critical of the construct of *g*-theory, particularly when considered from a life span developmental perspective and the use of single summary scores from collections of tests that measure different mixtures of cognitive constructs.

Of course, a correlation of Fluid (*Gf*) and Crystallized (*Gc*) intelligence was routinely allowed, and it was estimated to be both positive and relatively high ($r > .60$). For this reason, many people assumed that this was primary evidence of a single factor (*g*). Not John Horn. Although it is clear that one factor could cause the other, and this could lead to a strong correlation between them (see McArdle, 1989), Horn rigorously believed that “independent” did not mean two constructs were required to have a zero correlation. After all, this was thought to be true for other independently and well-measured variables, such as Height and Weight, or even the key measurements of Personality processes (e.g., Revelle, 1995). These are also scores that are highly positively correlated (see Sargent, 1963) but which have independent predictors and independent outcomes but are correlated due to being based on living samples (e.g., where height and weight must be correlated; see McArdle, 2012b). What we usually require is that these factors, assuming there exist more than one factor, behave in different ways from one another. This behavior is largely external to the factor analysis so it not determined in this way. But we would be very skeptical if subsequent models for both of these variables led to the same conclusions (they usually do not).

Incidentally, Thurstone (1947) showed that a part of the correlation among factors can emerge from nonrandom samplings of persons. What seems a bit unusual about this cognitive literature is that researchers in this substantive domain seem to ignore these kinds of arguments and simply continue to advocate *g* at a higher level until no restrictive test is possible (e.g., Carroll, 1993; Jensen, 1980). If so, we will always generate a single factor (see Humphreys, Davey, & Park, 1985).

Horn (1965b, 2005; Horn & Blankson, 2005) argued that given the evidence of different developmental trajectories (found in both between-person age differences and within-person age changes) across different cognitive abilities and tests, and their differential associations to distinct risk and protective factors, a consistent estimate of general intelligence is not likely or useful. Although always calculable from the positive correlations among tests, the general factor computed on one collection of tests has been shown to be different from general factors computed on other collections of tests. In a recent example, Floyd, Clark, and Shadish (2008) demonstrated that aggregate scores based on several well-established intelligence test batteries (e.g., KABC; WAIS, WJ III) resulted in scores that were not sufficiently exchangeable for many purposes in which they are used, noting that “1 in 4 individuals taking an intelligence test battery will receive an IQ more than 10 points higher or lower when taking another battery” (p. 414).

Horn’s major argument against the principle of general intelligence, however, is that different cognitive abilities have different construct validities and developmental patterns

(Horn, 1985, 1986, 1988, 1989, 1991). He seemed to make a similar argument with respect to the broad domain of Personality (see Revelle, 1995). The first principal components of collections of cognitive tests generally account for approximately 50% of the reliable variance (depending on the selection and breadth of such tests). Therefore, he reasoned, this first principal component analysis does not account for the other 50% of reliable individual differences in a variety of measured abilities, and something else must be measurable. Horn thought that such individual differences should not be ignored and aggregated given the evidence for distinct developmental patterns and important individual differences across particular intellectual abilities.

The *Gf-Gc* structural theory derived from studies of patterns of change related to development and aging as well as studies of how particular abilities, particularly *Gf* abilities, are vulnerable to lifestyle differences and brain damage whereas *Gc* abilities are in general maintained (Cattell & Horn, 1978; Horn, 1965b, 1967a). Thus, *Gf-Gc* theory is primarily based on second-order factor analysis of abilities represented in Thurstone's primary mental ability system and the pattern of findings have been replicated in a number of studies (e.g., Carroll, 1993; McArdle & Woodcock, 1997; Undheim & Gustaffson, 1987; Woodcock, 1990). In turn, *Gf* and *Gc* abilities have distinct characteristics but are related to other types of cognitive capabilities (Horn, 1991) including Processing Speed (*Gs*), Episodic Memory (*Gsm*), Long-Term Storage and Retrieval (*Glr*), Spatial Reasoning (*Gv*), Auditory Processes (*Ga*), and Quantitative Reasoning (*Gq*; see Figure 4). Horn constantly argued that Processing Speed measures were questionable and that Broad Concentration was the ability we needed to measure (e.g., see Horn, 2005). Based on Horn's efforts and clear theoretical critiques of general intelligence, the distinction between *Gf* and *Gc* abilities has become well established (see Blair, 2006; Nisbett et al., 2012; but see Johnson et al., 2004).

But Horn also proposed a unique view on this situation (see Horn, 2005; Horn & McArdle, 2007). This multifaceted phenomenon he envisioned was thought to exhibit no age-related differences when higher levels of individual expertise were obtained, and he recognized this concept probably requires a much different kind of dynamic assessment model (see Ferrer & McArdle, 2004, 2010; Hamagami & McArdle, 2007; McArdle, 2009). That is, standard factor analysis models would always yield a hierarchy of cognitive abilities as some have suggested (e.g., McGrew, 2005, 2009), but what we need is a specific cognitive profile where some indicators (i.e., *Gc*) would require a more complete understanding of areas of expertise within an individual (but see Woodcock, 1998). When the optimal measurement of people requires us to know if they "are good at anything" the standard factor analysis models of specific variables would not be indicated, and some form of individual mixture modeling is probably required (see McArdle & Nesselroade, 2012). To our knowledge, this type of alternative analysis of "extended *Gf-Gc* theory" has not yet been fully developed, but it certainly illustrates John L. Horn's scientific imagination.

THE MULTIFACETED STRUCTURE OF COGNITION AND GUILFORD'S SOI MODEL

As stated earlier, Horn did not believe the evidence pointed toward a single factor of intelligence, but he believed that a relatively small number of broad common factors could be introduced that would sufficiently account for the observed data:

Results from a great number of factor analytic studies done in this century add up to suggest a paramorphic organization that has become known as $g_f g_c$ theory (Cattell, 1957, 1971; Horn, 1968, 1988; Horn & Cattell, 1966). Such results indicate that if, in a substantial and heterogeneous sample of people, a researcher measures a variety of cognitive capabilities sampled to represent a simple structure of the abilities specified in $g_f g_c$ theory, the results from testing to see if the model fits the intercorrelations will indicate that, indeed, such a model does represent the data. It means, too, that a great number of other, seemingly plausible, models do not provide good approximations to the data. (Horn & Hofer, 1992, p. 50)

Horn's approach was based on what he demanded of others—a broad and adequate sampling of different kinds of ability tests that were administered to samples heterogeneous in terms of developmental influences and factors associated with impairment in function. Horn emphasized a developmentally based factor model of multiple intelligences—the theory of fluid and crystallized intelligence. The central idea is that exposure to different environments will lead to a differentiation of functioning into separate, but related, types of abilities early in the life span and that these developmental trends would carry forward into older age. That these distinct abilities exhibit different age-change curves is another principle argument against a single ability construct.

Thurstone's studies (1938, 1947) of the structure of primary mental abilities found that a number of factors were required to account for reliable individual differences obtained with different tests designed to measure important features of intelligence. However, Thurstone also supported a hierarchical organization of abilities, with g at the pinnacle. A technique for the rejection of this higher order factor model was never presented. Incidentally, a similar higher order approach was used by a student of Thurstone, J. B. Carroll (1993), and this classical hierarchical approach is still quite popular.

In what seemed to be different class of cognitive research, Guilford (1967) developed a theory referred to as the *structure-of-intellect* model (SOI) based on results of studies of primary mental abilities. Guilford took a taxonomic approach and based his theory on a three-dimensional construction of separate abilities based on mental operations (cognition, memory, divergent production, convergent production, and evaluation), contents (figural, symbolic, semantic, and behavioral), and types of products (units, classes, relations, systems, transformations, and implications). This organizational system implied 120 distinct mental abilities. It was never clear if "distinct" implied "independence" here either, but orthogonal rotations were quite typical of factorial solutions available at that time. Needless to say, orthogonal rotations were not a favorite of Horn.

Horn and others (Horn & Knapp, 1973, 1974; Undheim & Horn, 1977) argued that, although it was useful to consider new ways to construct tests, Guilford's SOI system did not correspond well to distinct human abilities on the basis of structural models and developmental processes. He also noted several problems with the methods used to factor analyze the data. Further interaction between Guilford and Horn focused on criticism of the theory of fluid and crystallized intelligence (Guilford, 1980; Horn & Cattell, 1982) and demonstrates again the positive impact of rigorous debate on science and theory development for which Horn is remembered. So, although Horn certainly appreciated the creation of novel cognitive measurements, Guilford's elegant SOI was not close to what Horn considered a methodological breakthrough. That is, John Horn seemed to think that factor analyses based on orthogonal rotations were great mathematical devices but not useful for scientific purposes. Guilford (1980) clearly did not view it this same way.

SCHAIIE AND BALTES' PERSPECTIVE ON THE LACK OF AGE-RELATED DECLINE IN COGNITION

Horn was deeply interested in understanding individual differences in the pattern and structure of cognitive functioning over the life span, including the question "When does age-related decline begin?" His studies were largely based on age-heterogeneous cross-sectional samples and his approach was to partial out individual differences in a variety of characteristics (e.g., persistence, carefulness) to understand the influences that modify the age-change curve.

Fundamentally, Horn's view was that both *Gf* and *Gc* abilities can be learned and carry forward across the life span, with different trends observed across different types of cognitive abilities (Horn & Cattell, 1967). In this theory, fluid ability (*Gf*) is age vulnerable with population average age-related decline seen in early to middle age. *Gf* is indicated by measures that tap the ability to solve novel problems (i.e., tasks that cannot be performed automatically) including mental operations such as drawing inferences; forming concepts; classifying, generating, and testing hypothesis; identifying relations; and comprehending implications. On the other hand, crystallized knowledge (*Gc*) abilities are relatively age maintained and show increases over much of the life span. Thus, and in contrast, *Gc* abilities refer to acquired knowledge of the language, information, and concepts specific to the dominant culture.

Although the general pattern of *Gf* declining before *Gc* had been demonstrated in both cross-sectional (between-person age differences) and longitudinal (within-person age changes) designs, these designs typically produce quite different findings regarding the age at which decline in abilities begins in the population. For example, cross-sectional studies (e.g., Salthouse, 2009) report decline in *Gf* in the mid-20s and *Gc* abilities after age 60. However, longitudinal studies typically find that *Gf* is maintained until the mid-50s and *Gc* abilities until later in the life span. These differences in results associated with research designs have been a long-standing issue and made for a heated debate in the 1970s.

At the heart of this debate, which took place in the *American Psychologist* between Schaie and Baltes (Baltes & Schaie, 1974; Baltes & Schaie, 1976; Schaie, 1974; Schaie & Baltes,

1977) and Horn and Donaldson (1976, 1977), was a claim by Baltes and Schaie (1974, p. 39) that “intellectual decline is largely a myth.” Baltes and Schaie described the general findings of “age-related cognitive decline” based on comparisons of individuals differing in age as a myth. Recall too that persons on both sides of this debate were considered leaders in the field of cognitive aging, so all eyes were on them. Baltes and Schaie’s evidence was presented in the diagram of their original paper. Through the back-and-forth discussion of this claim, one can clearly see the critical reasoning and exuberance that Horn brought to every discussion. He said he could not obtain the raw data from Baltes and Schaie, so he calculated all values of change (from their figures) on his own personal devices (e.g., a protractor). In subsequent work, the same basic concepts were used to generate the simulated data of Figure 6 (from Horn & McArdle, 1980).

Numerous methodological issues were discussed, such as whether dimensions of chronological age, birth cohort, and period of measurement could be disambiguated to provide a clear estimate of age-related trends. For example, Donaldson and Horn (1992) were highly critical of the potential for disambiguating age, cohort, and time, highlighting the inherent confounds associated with statistically disambiguating age and cohort without making strong assumptions. The Age by Cohort by Time decomposition was also considered in structural equation terms by Horn & McArdle (1980). Horn and McArdle (1980) suggested that assumptions be replaced by measurements and theory of birth cohort and period effects as a potential way forward in resolving this design confound. Complicating this approach is the likelihood for differential sample selectivity (i.e., birth and death rates) that operate in complex ways across individuals sampled at different ages and from different cohorts. So a final resolution remains to be seen but this is unlikely given the nature of the problem.

These same arguments over results from cross-sectional and longitudinal studies and their inherent design confounds continue today (McArdle & Nesselroade, 2012). Proponents of the findings regarding early life declines (e.g., Salthouse, 2009) have typically focused on the effect of repeated testing as the reason earlier age differences are not observed in longitudinal designs. Many researchers have studied design issues and statistical adjustment of retest effects (see Ferrer, Salthouse, McArdle, Stewart, & Schwartz, 2005; Hoffman, Hofer, & Sliwinski, 2011; McArdle, Ferrer-Caja, Hamagami, & Woodcock, 2002; McArdle & Woodcock, 1997; Sliwinski et al., 2010b; Thorvaldsson, Hofer, Berg, & Johansson, 2006). Interestingly, these authors all seem to conclude that retesting effects are unlikely to account for the major differences in results across these designs (McArdle & Woodcock, 1997; Sliwinski, Hoffman, & Hofer, 2010a).

Social change, improvements in the social and cultural environments over the last century, remains a potential reason for between-person differences in population change indicating cognitive decline. Flynn (1984, 1987; but see Teasdale & Owen, 2008) provided evidence for a steady increase in intelligence scores (approximately 3 IQ points or one fifth of a standard deviation per decade) across birth cohorts in the United States and many European countries. Although there is no clear consensus as to the reasons for these increases, potential causes include schooling, test sophistication, nutrition, stimulating environment, fertility patterns, and infectious diseases (e.g., Ang, Rodgers, & Wanstrom, 2010; Neisser et

al., 1998). The important point here is that these differences in cognitive test scores across birth cohorts may explain between-person age differences (e.g., Lee, Gorsuch, Saklofske, & Patterson, 2008) as these same cohorts are those that comprise a majority of our current cross-sectional and longitudinal studies of aging.

Although there continue to be methodological and substantive disagreements regarding some of these basic questions, sufficiently systematic results across study designs have been found regarding potential moderating factors associated with cognitive outcomes. A number of risk factors associated with age-related cognitive decline have been identified and replicated (for older summary, see Horn & McArdle, 1980; for recent summary, see Williams, Plassman, Burke, Holsinger, & Benjamin, 2010). Factors associated with increased risk for decline, particularly in fluid abilities, include dementia, head injury, cardiovascular disease, diabetes, APO $\epsilon 4$ genotype, depression, vision and hearing deficits, and smoking. Decreased risk of decline is associated with physical activity, cognitive engagement, social interaction, Mediterranean diet, educational attainment, socioeconomic status, and higher childhood cognitive ability.

The question “When does age-related decline begin?” is still unresolved and continues to be a topic of informed debate (e.g., Salthouse, 2009; Schaie, 2009) but with only minimal progress toward resolution. The general issues remain the same and focus on confounds related to differences in design. The inferential and statistical issues that complicate the analysis of longitudinal data (e.g., retest gains related to test exposure) are often cited as limiting factors for the development of theories and models of life span development and aging. However, longitudinal data contain information necessary to address population selection (e.g., attrition, mortality) and within-person change processes that is ignored and inaccessible in cross-sectional data. These methodological debates are likely to continue but will encourage us toward developing and using measurement designs to resolve some of these confounding factors (e.g., Hoffman et al., 2011; Sliwinski et al., 2010a, 2010b).

Although the focus of much of Horn’s research was on general between-person age trends, there are known limitations to the use of cross-sectional data and analysis for understanding associations and predictors of individual life span change. A common finding in cross-sectional cognitive aging studies is that the effects of chronological age differences are not statistically independent but are shared among different types of cognitive and noncognitive variables (for review, see Hofer, Berg, & Era, 2003). Indeed, for many years, research on understanding changes in cognition with age was based predominantly on comparison of individuals differing in age (i.e., cross-sectional designs and analysis). Although the between-person comparative approach has led to highly consistent findings regarding the interdependency of age-related changes, these findings are likely to be, at least partly, spurious and related to confounds with average between-person trends (e.g., Hofer et al., 2003; Hofer et al., 2006; Hofer & Sliwinski, 2001; Kraemer, Yesavage, Taylor, & Kupfer, 2000; McArdle & Woodcock, 1997). Inference from cross-sectional designs are additionally confounded by birth cohort effects (e.g., Schaie, 2011) and mortality selection (e.g., Kurland, Johnson, Eggleston, & Diehr, 2009), which limit the opportunities to understand patterns and determinants of individual and population change (but, for such corrections, see McArdle, Small, Backman, & Fratiglioni, 2005).

A key challenge then is to describe and explain changes that occur within developing and aging individuals and this requires longitudinal follow-up. In general, this process orientation is underdeveloped in current theories because most studies treat “age” as a between-person characteristic—as an age “difference.” For example, an implicit assumption is that continuous and smooth population age trends observed in cross-sectional samples reflect continuous and smooth cognitive changes occurring within aging individuals. However, depending on the cause, as Horn described in terms of brain injury and abrupt loss of function, cognitive loss with aging may not be continuous but rather stochastic and discontinuous. The pattern and understanding of such individual differences must be addressed by the direct observation of intraindividual change in the context of longitudinal designs. Indeed, as we discussed earlier, John L. Horn’s innovative work on intensive measurement designs and disentangling state, trait, and change dimensions (i.e., Horn, 1972) is still very much at the cutting edge of current research activities for understanding variation and change in cognition.

DISCUSSION

We have presented some of the work on factor analysis and on the structure and development of cognitive abilities initially created by John L. Horn. Of course these bodies of work are held together tightly, and each informs the other. Because these are based on multivariate metatheory, it would be difficult to separate one from the other, and Horn would probably not want us to do it anyway. Attempts to answer these questions regarding the development and aging of cognitive functions across the life span have contributed to a rich array of substantive and methodological debates and advances. Recent advances include application of innovative developmental designs (e.g., Stawski et al., 2013) in dynamic modeling and population inference conditional on mortality (see McArdle, Grimm, Hamagami, Bowles, & Meredith, 2009; McArdle et al., 2005; Muniz, van den Hout, Piccinin, Matthews, & Hofer, 2013) and measurement and analysis issues related to replication and comparisons across birth cohorts and cultures (Hofer & Piccinin, 2009; Hoffman et al., 2011; Sliwinski et al., 2010a). Although we may not yet have definitive answers to many of these fundamental questions, the field of investigation is increasingly emphasizing the need to focus on the structure and determinants of within-person change and variation in cognition, health, and aging (e.g., Alwin & Hofer, 2011; Hofer & Alwin, 2008). We hope this is the kind of work Horn would like to see.

John L. Horn also knew it was very difficult to answer these questions from any particular study or data set, so he advocated a cooperative and interdisciplinary merging of all kinds of data (as in Hofer & Piccinin, 2010), encouraging particular use of developments in analysis with “planned” incomplete data (Graham, Hofer, & MacKinnon, 1996; McArdle, 1994; McArdle & Hamagami, 1992; McArdle & Woodcock, 1997). We have realized the many benefits to collaborative endeavors related to longitudinal studies on aging, most notably the opportunity for simultaneous evaluation of longitudinal data to test, replicate, and extend prior findings on aging (Gallacher & Hofer, 2011; Hofer & Piccinin, 2009, 2010; McArdle, 2009; Piccinin & Hofer, 2008; Piccinin et al., 2013). Given the key aim of cross-study comparison, measurement equivalence of constructs (as in Horn & McArdle, 1992; Meredith & Horn, 2001) and use of similar statistical models do seem critical, as are the

evaluation of alternative models on the same data to permit sensitivity analysis of results across models and the determination of best approaches for answering particular questions (e.g., Piccinin, Muniz, Sparks, & Bontempo, 2011).

It is in this very sense that we think the work of John L. Horn was a bright light in the field of multivariate experimental psychology. His early work on multivariate computer simulation, his subsequent work on the number of factors or the optimal calculation of factor scores, and his early work on repeated measures factoring have all been useful to others using more modern terms. Likewise, his substantive conclusions remain important as well—his refutation of the single factor model of cognition, along with his rigid rejection of the creation of multiple indicators based on multiple orthogonal rotations, and his work on the underlying processes and lifestyle factors associated with downhill trends in cognition over age have all been picked up by others in the field and used well.

For these reasons we think that John L. Horn's major contributions to psychology, only some of which have been discussed here, continue to be ahead of his time and have a profound influence on our thinking and critical approach to answering complex questions. His contributions to factor analysis and the structure of intelligence; the important methodological debates of the 1970s and 1980s regarding age and cohort effects and related issues of sample selectivity; the innovative ideas underlying his approach to evaluating state, trait, and trait-change; and his willingness and encouragement to engage in critical evaluation of fundamental ideas and accepted scientific approaches (i.e., the *g* theory) are and will remain important contributions. Through his research and teaching he forced people to question popular assumptions, evaluate all the data available, and he consistently challenged us to think longer, harder, and better. His work will continue to inspire important research in the fields of multivariate analysis and human cognitive abilities for many decades to come. John L. Horn is obviously still "fighting for intelligence" and through his many writings and numerous colleagues we expect he will continue to do so for many years to come.

Acknowledgments

We thank Carol Prescott, Andrea Piccinin, Stephen West, Keith Widaman, and the reviewers for helpful comments on the draft.

FUNDING

We thank the National Institute on Aging (AG0713720 to the first author and AG026453 to the second author) for support of this research. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institute on Aging or the National Institutes of Health.

References

- Alwin DF, Hofer SM. Health and cognition in aging research. *The Journals of Gerontology Series B: Psychological Sciences*. 2011; 66(Suppl 1):i9–i16.
- Ang S, Rodgers JL, Wanstrom L. The Flynn Effect within subgroups in the U.S.: Gender, race, income, education, and urbanization differences in the NLSY-Children data. *Intelligence*. 2010; 38:367–384. [PubMed: 20657802]
- Baltes PB, Schaie KW. The myth of the twilight years. *Psychology Today*. 1974; 40:35–38.

- Baltes PB, Schaie KW. On the plasticity of intelligence in adulthood and old age: Where Horn and Donaldson fail. *American Psychologist*. 1976; 31:720–725. [PubMed: 1008332]
- Bartlett MS. Tests of significance in factor analysis. *British Journal of Psychology*. 1950; 3:77–85.
- Binet, A.; Simon, T. New methods for the diagnosis of the intellectual level of subnormals. In: Goddard, HH., editor; Kite, ES., translator. *Development of intelligence in children (the Binet-Simon Scale)*. Baltimore, MD: Williams & Wilkins; 1905/1916. p. 37-90.
- Blair C. How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*. 2006; 29:109–125. [PubMed: 16606477]
- Bock, RD. *Multivariate statistical methods in behavioral research*. New York, NY: McGraw-Hill; 1975.
- Boker SM, Molenaar PCM, Nesselroade JR. Issues in intraindividual variability: Individual differences in equilibria and dynamics over multiple time scales. *Psychology and Aging*. 2009; 24:858–862. [PubMed: 20025401]
- Browne MW, Cudeck R. Single sample cross-validation indices for covariance structures. *British Journal of Mathematical and Statistical Psychology*. 1989; 37:62–83. [PubMed: 6733054]
- Browne, MW.; Cudeck, R. Alternative ways of assessing model fit. In: Bollen, KA.; Long, JS., editors. *Testing structural equation models*. Beverly Hills, CA: Sage; 1993. p. 136-162.
- Carroll, JB. *Human cognitive abilities: A survey of factor-analytical studies*. New York, NY: Cambridge University Press; 1993.
- Cattell RB. Some theoretical issues in adult intelligence testing. *Psychological Bulletin*. 1941; 38:592.
- Cattell, RB. *Personality: A systematic theoretical and factual study*. New York, NY: McGraw-Hill; 1950.
- Cattell, RB. *Personality and motivation structure and measurement*. New York, NY: World Books; 1957.
- Cattell, RB. Psychological theory and scientific method. In: Cattell, RB., editor. *Handbook of multivariate experimental psychology*. Chicago, IL: Rand McNally & Co; 1966. p. 1-18.
- Cattell, RB. *Abilities: Their structure, growth and action*. Boston, MA: HoughtonMifflin; 1971.
- Cattell RB, Horn JL. A check on the theory of fluid and crystallized intelligence with description of new subtest designs. *Journal of Educational Measurement*. 1978; 15:139–164.
- Ciesla JA, Cole DA, Steiger JH. Extending the trait-state-occasion model: How important is within-wave measurement equivalence? *Structural Equation Modeling*. 2007; 14:77–97.
- Cronbach LJ, Gleser GC. Assessing similarity between profiles. *Psychological Bulletin*. 1953; 50:456–473. [PubMed: 13112334]
- Dinno A. Exploring the sensitivity of Horn's parallel analysis to the distributional form of random data. *Multivariate Behavioral Research*. 2009; 44:362–388. [PubMed: 20234802]
- DiStefano C, Zhu M, Mindola D. Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research, & Evaluation*. 2009; 14(20):1–11. Available online at <http://pareonline.net/getvn.asp?v=14&n=20>.
- Donaldson G, Horn JL. Age, cohort, and time developmental muddles: Easy in practice, hard in theory. *Experimental Aging Research*. 1992; 18(4):213–222. [PubMed: 1459165]
- Efron B. Bootstrap methods: Another look at the jackknife. *Annals of Mathematical Statistics*. 1979; 7:1–26.
- Ericsson KA, Kintsch W. Long-term working memory. *Psychological Review*. 1995; 105:211–245. [PubMed: 7740089]
- Ferrer E, McArdle JJ. An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*. 2004; 40:935–952. [PubMed: 15535749]
- Ferrer E, McArdle JJ. Longitudinal modeling of developmental changes in psychological research. *Current Directions in Psychological Science*. 2010; 19:149–154.
- Ferrer E, Salthouse TA, McArdle JJ, Stewart WF, Schwartz B. Multivariate modeling of age and retest effects in longitudinal studies of cognitive abilities. *Psychology and Aging*. 2005; 20(3):412–442. [PubMed: 16248701]

- Floyd RG, Clark MH, Shadish WR. The exchangeability of IQs: Implications for professional psychology. *Professional Psychology: Research and Practice*. 2008; 39(4):414–423.
- Flynn JR. The mean IQ of Americans: Massive gains 1932 to 1978. *Psychological Bulletin*. 1984; 95:29–51.
- Flynn JR. Massive IQ gains in 14 nations: What IQ tests really measure. *Psychological Bulletin*. 1987; 101:171–191.
- Gallacher J, Hofer SM. Generating large-scale longitudinal data resources for aging research. *Journals of Gerontology: Psychological Sciences*. 2011; 66B:i72–i79.
- Gottfredson LS. Intelligence and social policy [Special issue]. *Intelligence*. 1997; 24(1):1–320.
- Graham JW, Hofer SM, MacKinnon DP. Maximizing the usefulness of data obtained with planned missing value patterns: An application of maximum likelihood procedures. *Multivariate Behavioral Research*. 1996; 31:197–218.
- Grice JW. A comparison of factor scores under conditions of factor obliquity. *Psychological Bulletin*. 2001; 6:67–83.
- Grice JW, Harris RJ. A comparison of regression and loading weights for the computation of factor scores. *Multivariate Behavioral Research*. 1998; 33:221–247.
- Guilford, JP. *The nature of human intelligence*. New York, NY: McGraw-Hill; 1967.
- Guilford JP. Fluid and crystallized intelligences: Two fanciful concepts. *Psychological Bulletin*. 1980; 88:406–412.
- Hamagami, F.; McArdle, JJ. Dynamic extensions of latent difference score models. In: Boker, SM.; Wegner, ML., editors. *Data analytic techniques for dynamical systems*. Mahwah, NJ: Erlbaum; 2007. p. 47–85.
- Hayton JC. Commentary on “Exploring the Sensitivity of Horn’s Parallel Analysis to the Distributional Form of Random Data. *Multivariate Behavioral Research*. 2009; 44:389–395.
- Hofer, SM.; Alwin, DF. The future of cognitive aging research: Interdisciplinary perspectives and integrative science. In: Hofer, SM.; Alwin, DF., editors. *Handbook on cognitive aging: Interdisciplinary perspectives*. Thousand Oaks, CA: Sage; 2008. p. 662–672.
- Hofer SM, Berg S, Era P. Evaluating the interdependence of aging-related changes in visual and auditory acuity, balance, and cognitive functioning. *Psychology and Aging*. 2003; 18:285–305. [PubMed: 12825777]
- Hofer SM, Flaherty BP, Hoffman L. Cross-sectional analysis of time-dependent data: Mean-induced association in age-heterogeneous samples and an alternative method based on sequential narrow age-cohorts. *Multivariate Behavioral Research*. 2006; 41:165–187.
- Hofer SM, Piccinin AM. Integrative data analysis through coordination of measurement and analysis protocol across independent longitudinal studies. *Psychological Methods*. 2009; 14:150–164. [PubMed: 19485626]
- Hofer SM, Piccinin AM. Toward an integrative science of lifespan development and aging. *Journals of Gerontology: Psychological Sciences*. 2010; 65B:269–278.
- Hofer SM, Sliwinski MJ. Understanding ageing: An evaluation of research designs for assessing the interdependence of ageing-related changes. *Gerontology*. 2001; 47:341–352. [PubMed: 11721149]
- Hoffman L, Hofer SM, Sliwinski MJ. On the confounds among retest gains and age-cohort differences in the estimation of within-person change in longitudinal studies: A simulation study. *Psychology and Aging*. 2011; 26:778–791. [PubMed: 21639642]
- Horn JL. A rationale and test for the number of factors in factor analysis. *Psychometrika*. 1965a; 30:179–185. [PubMed: 14306381]
- Horn, JL. Doctoral dissertation. University of Illinois; 1965b. Fluid and crystallized intelligence: A factor analytic study of the structure among primary mental abilities.
- Horn JL. Intelligence—Why it grows, why it declines. *Society*. 1967a; 5:23–31.
- Horn JL. On subjectivity in factor analysis. *Educational and Psychological Measurement*. 1967b; 27:811–820.
- Horn JL. Organization of abilities and the development of intelligence. *Psychological Review*. 1968; 75:242–259. [PubMed: 4875815]

- Horn, J.L. Organisation of data on life-span development of human abilities. In: Goulet, L.R.; Baltes, P.B., editors. *Life-span developmental psychology: Research and theory*. New York, NY: Academic Press; 1970. p. 423-466.
- Horn J.L. State, trait, and change dimensions of intelligence. *British Journal of Educational Psychology*. 1972; 42:159-185.
- Horn J.L. To set the stage. *Multivariate Behavioral Research*. 1984; 19:115-120.
- Horn, J.L. Remodeling old models of intelligence. In: Wolman, B.B., editor. *Handbook of intelligence*. New York, NY: Wiley; 1985. p. 267-300.
- Horn, J.L. Intellectual ability concepts. In: Sternberg, R.J., editor. *Advances in the psychology of human intelligence*. Vol. 3. Hillsdale, NJ: Erlbaum; 1986. p. 35-77.
- Horn, J.L. Thinking about human abilities. In: Nesselroade, J.R.; Cattell, R.B., editors. *The handbook of multivariate experimental psychology*. Vol. 2. New York, NY: Plenum Press; 1988. p. 561-614.
- Horn, J.L. Models of intelligence. In: Linn, R., editor. *Intelligence, measurement & public policy*. Urbana: University of Illinois Press; 1989. p. 29-73.
- Horn, J.L. Measurement of intellectual capabilities: A review of theory. In: McGrew, K.S.; Werder, J.K.; Woodcock, R.W., editors. *Woodcock-Johnson technical manual*. Allen, TX: DLM Teaching; 1991. p. 197-245.
- Horn, J.L. Spearman, expertise, and the nature of human cognitive capabilities. In: Kyllonen, P.C.; Roberts, R.D.; Stankov, L., editors. *Extending intelligence: Enhancement and new constructs*. New York, NY: Taylor & Francis; 2005. p. 185-230.
- Horn, J.L.; Blankson, N. Foundation for better understanding cognitive abilities. In: Flanagan, D.P.; Harrison, P., editors. *Contemporary intellectual assessment*. 2. New York, NY: Guilford Press; 2005. p. 41-68.
- Horn J.L, Cattell R.B. Age differences in primary mental ability factors. *Journal of Gerontology*. 1966a; 21:210-220. [PubMed: 5930515]
- Horn J.L, Cattell R.B. Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology*. 1966b; 57:253-270. [PubMed: 5918295]
- Horn J.L, Cattell R.B. Age differences in fluid and crystallized intelligence. *Acta Psychologica*. 1967; 26:107-129. [PubMed: 6037305]
- Horn J.L, Cattell R.B. Whimsy and misunderstanding of gf-gc theory: A comment on Guilford. *Psychological Bulletin*. 1982; 91:623-633.
- Horn J.L, Donaldson G. On the myth of intellectual decline in adulthood. *American Psychologist*. 1976; 31:701-729. [PubMed: 1008331]
- Horn J.L, Donaldson G. Faith is not enough: A response to the Baltes-Schaie claim that intelligence does not wane. *American Psychologist*. 1977; 32:369-373.
- Horn J.L, Donaldson G, Engstrom R. Apprehension, memory, and fluid intelligence decline in adulthood. *Research on Aging*. 1981; 3:33-84.
- Horn J.L, Engstrom R. Cattell's scree test in relation to Bartlett's chi-square test and other observations on the number of factors problem. *Multivariate Behavioral Research*. 1979; 14:283-300.
- Horn, J.L.; Hofer, S.M. Major abilities and development in the adult period. In: Sternberg, R.J.; Berg, C.A., editors. *Intellectual development*. New York, NY: Cambridge University Press; 1992. p. 44-99.
- Horn J.L, Knapp J.R. On the subjective character of the empirical base of Guilford's structure-of-intellect model. *Psychological Bulletin*. 1973; 80(1):33-43.
- Horn J.L, Knapp J.R. Thirty wrongs do not make a right: Reply to Guilford. *Psychological Bulletin*. 1974; 81(8):502-504.
- Horn J.L, Little K.B. Isolating change and invariance in patterns of behavior. *Multivariate Behavioral Research*. 1966; 1:219-228.
- Horn, J.L.; Masunaga, H. New directions for research into aging and intelligence: The development of expertise. In: Perfect, T.J.; Maylor, E.A., editors. *Models of cognitive aging: Debates in psychology*. New York, NY: Oxford University Press; 2000. p. 125-159.

- Horn, J.L.; McArdle, J.J. Perspectives on mathematical and statistical model building (MASMOB) in research on aging. In: Poon, L., editor. *Aging in the 1980's: Psychological Issues*. Washington, DC: American Psychological Association; 1980. p. 503-541.
- Horn J.L, McArdle J.J. A practical guide to measurement invariance in aging research. *Experimental Aging Research*. 1992; 18(3):117-144. [PubMed: 1459160]
- Horn, J.L.; McArdle, J.J. Understanding human intelligence since Spearman. In: Cudeck, R.; MacCallum, R., editors. *Factor analysis at 100 years*. Mahwah, NJ: Erlbaum; 2007. p. 205-247.
- Horn J.L, Miller W.C. Evidence on problems in estimating common factor scores. *Educational and Psychological Measurement*. 1966; 26:617-622.
- Horn, J.L.; Noll, J. Human cognitive capabilities: Gf-Gc theory. In: Flanagan, D.P.; Genshaft, J.L.; Harrison, P.L., editors. *Contemporary intellectual assessment: Theories, tests, and issues*. New York, NY: Guilford Press; 1997. p. 53-91.
- Humphreys L.G, Davey T.C, Park R.K. Longitudinal correlation analysis of standing height and intelligence. *Child Development*. 1985; 56:1465-1478. [PubMed: 4075869]
- Jensen, A.R. *Bias in mental testing*. New York, NY: Free Press; 1980.
- Jensen, A.R. *The g factor: The science of mental ability*. Westport, CT: Praeger; 1998.
- Jensen A.R, Weng L.J. What is a good g? *Intelligence*. 1994; 18:231-258.
- Johnson W, Bouchard T.J, Krueger R.F, McGue M, Gotesman I.I. Just one g: Consistent results from three test batteries. *Intelligence*. 2004; 32:95-107.
- Jöreskog, K.G.; Sörbom, D. *Advances in factor analysis and structural equation models*. Cambridge, MA: Abt; 1979.
- Kaiser H.F. The application of electronic computers to factor analysis. *Educational and Psychological Measurement*. 1960; 20:141-151.
- Kelley, T.L. *Crossroads in the mind of man*. Stanford, CA: Stanford University Press; 1928.
- Kraemer H.C, Yesavage J.A, Taylor J.L, Kupfer D. How can we learn about developmental processes from cross-sectional studies, or can we? *American Journal of Psychiatry*. 2000; 157:163-171. [PubMed: 10671382]
- Kurland B.F, Johnson L.L, Egleston B.L, Diehr P.H. Longitudinal data with follow-up truncated by death: Match the analysis method to research aims. *Statistical Science*. 2009; 24(2):211-222.10.1214/09-STS293 [PubMed: 20119502]
- Lederman W. On the rank of the reduced correlational matrix in multiple factor analysis. *Psychometrika*. 1937; 2:85-93.
- Ledesma R, Valero-Mora P. Determining the number of factors to retain in EFA: An easy-to-use computer program for carrying out Parallel Analysis. *Practical Assessment, Research & Evaluation*. 2007; 12(1):1-11.
- Lee H.F, Gorsuch R.L, Saklofske D.H, Patterson C.A. Cognitive differences for ages 16 to 89 years (Canadian WAIS-III): Curvilinear with Flynn and processing speed corrections. *Journal of Psychoeducational Assessment*. 2008; 26:382-394.
- Marsh H.W, Ludtke O, Muthén B, Asparouhov T, Morin A.J.S, Trautwein U, Nagengast B. A new look at the big five factor structure through exploratory structural equation modeling. *Psychological Assessment*. 2010; 22:471-491. [PubMed: 20822261]
- Marsh H.W, Muthén B, Asparouhov T, Ludtke O, Robitzsch A, Morin A.J.S, Trautwein U. Exploratory structural equation modeling, integrating CFA and EFA: Application to students' evaluations of university teaching. *Structural Equation Modeling*. 2009; 16:439-476.
- Masunaga H, Horn J.L. Characterizing mature human intelligence: Expertise development. *Learning and Individual Differences*. 2000; 12:5-33.
- Masunaga H, Horn J.L. Expertise and age-related changes in components of intelligence. *Psychology and Aging*. 2001; 16:293-311. [PubMed: 11405317]
- McArdle, J.J. Dynamic but structural equation modeling of repeated measures data. In: Nesselroade, J.R.; Cattell, R.B., editors. *The handbook of multivariate experimental psychology*. Vol. 2. New York, NY: Plenum Press; 1988. p. 561-614.

- McArdle, JJ. Structural modeling experiments using multiple growth functions. In: Ackerman, P.; Kanfer, R.; Cudeck, R., editors. *Learning and individual differences: Abilities, motivation, and methodology*. Hillsdale, NJ: Erlbaum; 1989. p. 71-117.
- McArdle, JJ. Comments on "Latent Variable Models for Studying Difference and Changes". In: Collins, L.; Horn, JL., editors. *Best methods for the analysis of change*. Washington, DC: APA Press; 1991. p. 164-169.
- McArdle JJ. Structural factor analysis experiments with incomplete data. *Multivariate Behavioral Research*. 1994; 29:409-454.
- McArdle, JJ. The development of RAM rules for latent variable structural equation modeling. In: Madeau, A.; McArdle, JJ., editors. *Contemporary advances in psychometrics*. Mahwah, NJ: Erlbaum; 2005. p. 225-273.
- McArdle JJ. John L. Horn (1928–2006). *American Psychologist*. 2007a; 62:596–597. [PubMed: 17874900]
- McArdle, JJ. Five steps in the structural factor analysis of longitudinal data. In: Cudeck, R.; MacCallum, R., editors. *Factor analysis at 100 years*. Mahwah, NJ: Erlbaum; 2007b. p. 99-130.
- McArdle JJ. Latent variable modeling of longitudinal data. *Annual Review of Psychology*. 2009; 60:577–605.
- McArdle, JJ. Ethical issues in factor analysis. In: Panter, AT.; Sterba, SK., editors. *Handbook of ethics in quantitative methodology*. New York, NY: Taylor & Francis; 2012a. p. 313-340.
- McArdle JJ. Testing the idea of general intelligence. *F&M Scientist*. 2012b; 1:27–66.
- McArdle JJ, Cattell RB. Structural equation models of factorial invariance in parallel proportional profiles and oblique confactor problems. *Multivariate Behavioral Research*. 1994; 29:63–113.
- McArdle JJ, Ferrer-Caja E, Hamagami F, Woodcock RW. Comparative longitudinal multilevel structural analyses of the growth and decline of multiple intellectual abilities over the life-span. *Developmental Psychology*. 2002; 38(1):115–142. [PubMed: 11806695]
- McArdle JJ, Fisher GG, Kadlec KM. Latent variable analysis of age trends in tests of cognitive ability in the health and retirement survey, 1992–2004. *Psychology and Aging*. 2007; 22(3):525–545. [PubMed: 17874952]
- McArdle JJ, Grimm KJ, Hamagami F, Bowles RP, Meredith W. Modeling life-span growth curves of cognition using longitudinal data with multiple samples and changing scales of measurements. *Psychological Methods*. 2009; 14:126–149. [PubMed: 19485625]
- McArdle JJ, Hamagami E. Modeling incomplete longitudinal and cross-sectional data using latent growth structural models. *Experimental Aging Research*. 1992; 18(3):145–166. [PubMed: 1459161]
- McArdle, JJ.; Horn, JL. Structural equation models of Gf-Gc intelligence theory. Paper presented at the annual meetings of the American Psychological Association; Los Angeles, CA. 1981 Aug.
- McArdle JJ, McDonald RP. Some algebraic properties of the Reticular Action Model for moment structures. *The British Journal of Mathematical and Statistical Psychology*. 1984; 37:234–251. [PubMed: 6509005]
- McArdle, JJ.; Nesselroade, JR. Using multivariate data to structure developmental change. In: Cohen, SH.; Reese, HW., editors. *Life-span developmental psychology: Methodological innovations*. Hillsdale, NJ: Erlbaum; 1994. p. 223-267.
- McArdle, JJ.; Nesselroade, JR. A workshop in longitudinal structural equation modeling. Washington, DC: APA Books; 2012.
- McArdle JJ, Prescott CA. Age-based construct validation using structural equation models. *Experimental Aging Research*. 1992; 18(3):87–115. [PubMed: 1459166]
- McArdle JJ, Small BJ, Backman L, Fratiglioni L. Longitudinal models of growth and survival applied to the early detection of Alzheimer's disease. *Journal of Geriatric Psychiatry and Neurology*. 2005; 18(4):234–241. [PubMed: 16306246]
- McArdle JJ, Woodcock RW. Expanding test-retest designs to include developmental time-lag components. *Psychological Methods*. 1997; 2(4):403–435.
- McGrew, KS. The Cattell-Horn-Carroll theory of cognitive abilities: Past, present, and future. In: Flanagan, DP.; Genshaft, JL.; Harrison, PL., editors. *Contemporary intellectual assessment: Theories, tests, and issues*. New York, NY: Guilford Press; 2005. p. 136-182.

- McGrew KS. CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*. 2009; 37(1):1–10.
- Meredith W. Measurement invariance, factor analysis and factorial invariance. *Psychometrika*. 1993; 58(4):525–543.
- Meredith, W.; Horn, J. The role of factorial invariance in modeling growth and change. In: Collins, LM., editor. *New methods for the analysis of change*. Washington, DC: American Psychological Association; 2001. p. 203–240.
- Meredith W, Tisak J. Latent curve analysis. *Psychometrika*. 1990; 55:107–122.
- Montanelli RG Jr, Humphreys LG. Latent roots of random data correlation matrices with squared multiple correlations on the diagonal: A Monte Carlo study. *Psychometrika*. 1976; 41:341–348.
- Muniz GM, van den Hout A, Piccinin AM, Matthews F, Hofer SM. Investigating terminal decline: Results from a UK population-based study of ageing. *Psychology and Aging*. 2013; 28:377–385. [PubMed: 23276221]
- Muthén B, Asparouhov T. Bayesian SEM: A more flexible representation of substantive theory. *Psychological Methods*. 2012; 17:313–335. [PubMed: 22962886]
- Muthén, LK.; Muthén, BO. *Mplus user's guide*. 7. Los Angeles, CA: Author; 1998–2012.
- Muthén LK, Muthén BO. How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling: A Multidisciplinary Journal*. 2005; 9(4):599–620.
- Neisser, U. *The rising curve: Long-term gains in IQ and related measures*. Washington, DC: American Psychological Association; 1998.
- Nisbett RE, Aronson J, Blair C, Dickens W, Flynn J, Halpern D, Turkheimer E. *Intelligence: New findings and theoretical developments*. *American Psychologist*. 2012; 67:130–159. [PubMed: 22233090]
- Ntzoufras, I. *Bayesian modeling using WinBUGS*. New York, NY: Wiley; 2009.
- Piccinin, AM.; Hofer, SM. Integrative analysis of longitudinal studies on aging: Collaborative research networks, meta-analysis, and optimizing future studies. In: Hofer, SM.; Alwin, DF., editors. *Handbook on cognitive aging: Interdisciplinary perspectives*. Thousand Oaks, CA: Sage; 2008. p. 446–476.
- Piccinin AM, Muniz G, Clouston S, Reynolds C, Thorvaldsson V, Deary I, Hofer SM. Coordinated analysis of age, sex, and education effects on change in MMSE scores. *The Journals of Gerontology, Series B: Psychological and Social Sciences*. 2013; 68:374–390.
- Piccinin AM, Muniz G, Sparks C, Bontempo DE. An evaluation of analytical approaches for understanding change in cognition in the context of aging and health. *The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences*. 2011; 66B(S1):i35–i48. 10.1093/geronb/gbr038
- Revelle W. Personality processes. *Annual Review of Psychology*. 1995; 46:295–328.
- Salthouse TA. When does age-related cognitive decline begin? *Neurobiology of Aging*. 2009; 30:507–514. [PubMed: 19231028]
- Sargent D. Weight-height relationship of young men and women. *American Journal of Clinical Nutrition*. 1963; 13:318–325. [PubMed: 14080506]
- Schaie KW. Translations in gerontology—From lab to life: Intellectual functioning. *American Psychologist*. 1974; 29:802–807. [PubMed: 4440934]
- Schaie KW. When does age-related cognitive decline begin? Salt-house again reifies the “cross-sectional fallacy. *Neurobiology of Aging*. 2009; 30:528–533. [PubMed: 19231029]
- Schaie, KW. Historical influences on aging and behavior. In: Schaie, KW.; Willis, SL., editors. *Handbook of the psychology of aging*. 7. San Diego, CA: Elsevier; 2011. p. 41–55.
- Schaie KW, Baltes PB. Some faith helps to see the forest: A final comment on the Horn and Donaldson myth of the Baltes-Schaie position on adult intelligence. *American Psychologist*. 1977; 32:1118–1120.
- Schmiedek F, Lövdén M, Lindenberger U. Hundred days of cognitive training enhance broad cognitive abilities in adulthood: Findings from the COGITO study. *Frontiers in Aging Neuroscience*. 2010; 2:1–10. 10.3389/fnagi.2010.00027 [PubMed: 20552041]

- Sliwinski MJ, Hoffman L, Hofer SM. Evaluating convergence of within-person change and between-person age differences in age-heterogeneous longitudinal studies. *Research in Human Development*. 2010a; 7:45–60. [PubMed: 20671986]
- Sliwinski, MJ.; Hoffman, L.; Hofer, SM. Modeling retest and aging effects in a measurement burst design. In: Molenaar, PCM.; Newell, KM., editors. *Individual pathways of change: Statistical models for analyzing learning and development*. Washington, DC: American Psychological Association; 2010b. p. 37-50.
- Spearman CE. "General intelligence," objectively determined and measured. *American Journal of Psychology*. 1904; 15:201–293.
- Stawski R, Sliwinski MJ, Hofer SM. Between person and within person associations among speed, attention switching and working memory in younger and older adults. *Experimental Aging Research*. 2013; 39:194–214. [PubMed: 23421639]
- Steiger JH. Structural model evaluation and modification: an interval estimation approach. *Multivariate Behavioral Research*. 1990; 25:173–180.
- Steiger, JH.; Lind, J. Statistically-based tests for the number of common factors. Paper presented at the annual spring meeting of the Psychometric Society; Iowa City, IA. 1980 May.
- Steyer R, Ferring D, Schmitt M. States and traits in psychological assessment. *European Journal of Psychological Assessment*. 1992; 8:79–98.
- Steyer R, Schmitt M, Eid M. Latent state-trait theory and research in personality and individual differences. *European Journal of Personality*. 1999; 13(5):389–408.
- Teasdale TW, Owen DR. Secular declines in cognitive test scores: A reversal of the Flynn Effect. *Intelligence*. 2008; 36:121–126.
- Thorvaldsson V, Hofer SM, Berg S, Johansson B. Effects of repeated testing in a longitudinal age-homogeneous study of cognitive aging. *Journal of Gerontology: Psychological Sciences*. 2006; 61B:348–354.
- Thurstone, LL. *Psychometric Monographs*, No. 1. Chicago, IL: University of Chicago Press; 1938. Primary mental abilities.
- Thurstone, LL. *Multiple-factor analysis*. Chicago, IL: University of Chicago Press; 1947.
- Tucker LR, Lewis C. The reliability coefficient for maximum likelihood factor analysis. *Psychometrika*. 1973; 38:1–10.
- Undheim JO, Gustaffson JE. The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations. *Multivariate Behavioral Research*. 1987; 22:149–171.
- Undheim JO, Horn JL. Critical evaluation of Guilford's structure-of-intellect theory. *Intelligence*. 1977; 1:65–81.
- Velicer, WF.; Eaton, CA.; Fava, JL. Construct explication through factor or component analysis: A review and evaluation of alternative procedures for determining the number of factors or components. In: Goffen, RD.; Helms, E., editors. *Problems and solutions in human assessment: Honoring Douglas N. Jackson at seventy*. New York, NY: Springer; 2000. p. 41-71.
- Wackwitz JH, Horn JL. On obtaining the best estimates of factor scores within an ideal simple structure. *Multivariate Behavioral Research*. 1971; 6:389–408.
- Williams, JW.; Plassman, BL.; Burke, J.; Holsinger, T.; Benjamin, S. Preventing Alzheimer's disease and cognitive decline: Evidence report/technology assessment No. 193. Rockville, MD: Agency for Healthcare Research and Quality; 2010. (Prepared by the Duke Evidence-based Practice Center under Contract No. HHS 290-2007-10066-I.) AHRQ Publication No. 10-E005
- Woodcock RW. Theoretical foundations of the WJ-R measures of cognitive ability. *Journal of Psychoeducational Assessment*. 1990; 8:231–258.
- Woodcock, RW. Extending :Gf-Gc theory into practice. In: McArdle, JJ.; Woodcock, RW., editors. *Human abilities in theory and practice*. Mahwah, NJ: Erlbaum; 1998. p. 137-156.
- Zwick WR, Velicer WF. Comparison of five rules for determining the number of components to retain. *Psychological Bulletin*. 1986; 99:432–442.

FIGURE 1.
John L. Horn solving a multivariate problem (photograph courtesy of J. Wackwitz, 1983).

DEVELOPMENTAL CURVES FOR FLUID AND CRYSTALLISED INTELLIGENCE.
(after Horn, 1970)

G = Overall Intelligence Ability = $G_c + G_f$.
 G_c = Crystallised Intelligence.
 G_f = Fluid Intelligence.
 E = Acculturational Exposure.
 M = Maturational Growth and Decline.
 I = Injury to Neural Structure.

FIGURE 2.

G_f - G_c declines (from Horn, 1970, 1972).

FIGURE 3.

Using random variables to determine number of common factors (from Horn, 1965a).

INTERPRET ARROWS AS ONLY VERY ROUGHLY REPRESENTING THE "FLOW" OF FUNCTION (I.E., ONE MUST DETECT BEFORE STORAGE, STORE BEFORE ORGANIZING, ETC.) AND DEVELOPMENT (I.E., SENSORIMOTOR CAPACITIES APPEAR BEFORE ASSOCIATIVE ORGANIZATIONS--CONCRETE OPERATIONAL REPRESENTATION--WHICH IN TURN APPEAR BEFORE ABSTRACT THINKING).

FIGURE 4.

A generic structural model of cognition (from Horn, Donaldson, & Engstrom, 1981).

THE DATA BOX: A REPRESENTATION OF MEASUREMENTS OBTAINED IN THE STUDY OF
CHANGE AND INVARIANCE.

FIGURE 5.
Cattell's databox (from Horn, 1972).

FIGURE 6.
 Plots of simulated data to mimic Baltes and Schaie (1974; Horn & McArdle, 1980).