

Research Articles

More Than “Just 1 g ”? The General Intelligence Paradox and Its Solution

Matthew A. Sarraf¹, Michael A. Woodley of Menie¹, John G.R. Fuerst², Mateo Peñaherrera-Aguirre³

¹ Independent Sector, ² Independent Researcher, ³ School of Animal and Comparative Sciences, University of Arizona

Keywords: CHC Model, formative vs. reflective, G_c , G_f , g paradox

<https://doi.org/10.65550/001c.155379>

Intelligence & Cognitive Abilities

Vol. 2, Issue 1, 2026

Cattell–Horn–Carroll models of intelligence frequently show that, at the group-factor level, G_f is most strongly related to g , whereas at the subtest level, G_c -associated measures exhibit the highest g loadings. One proposed solution to this “ g paradox” holds that Stratum-III g and Stratum-II G_f are identical, and that the sizeable g loadings of crystallized subtests merely reflect the investment of G_f into learning. Investment theory is weakly evidenced, however. We argue that the “ g paradox” results from subtests measuring facets of G_f exhibiting pronounced specificity for cognitive entities. Capturing everything that goes into G_f is difficult on a single-measure basis, hence lower Stratum-I g loadings. The G_f group factor is nonetheless *reflective* of the composite of these entities and therefore is uniquely (at Stratum II) associated with g . Subtests measuring G_c broadly index the quality of global systems involving many cognitive processes, not entities, and so relate to factors that have *formative* effects on g , which are Stratum-I specific. We posit the existence of two distinguishable sources of general covariance: a formative g (associated primarily with G_c) and a reflective g (associated primarily with G_f), with the latter hierarchically superordinate to the former. Network analysis of “pure” psychometric measures of the G_c and G_f domains indicates that the former exhibits significantly greater network integrity than the latter, consistent with this formative/reflective model. Random effects meta-analysis of SEM contrast parameters, derived from four large genetically informed studies, finds that subtests assigned to a “ G_c ” category are associated with higher-magnitude direct (formative) genetic paths relative to those in a “ G_f ” category, suggesting a weak but discriminable and broad Stratum-I g in the residual covariance structure. Given the theorized phylogenetic histories of these two g s, we term the formative (“bottom-up”) g “proto g ” (g_p), and the reflective (“top-down”) g “neo g ” (g_n).

Introduction

In the psychometrics literature, there is (seemingly) conflicting evidence regarding which cognitive abilities¹ most strongly associate with general intelligence (g ; or general cognitive ability, GCA), an issue that has substantial implications for the nature of g itself. A popular theory of the structure of intelligence originates with psychologist Raymond B. Cattell (R. Cattell, 1963), and decomposes mental abilities into two broad categories: *fluid* (g_f or G_f) and *crystallized* (g_c or G_c). Fluid abilities are characterized by

their relative lack of dependence on an individual’s existing knowledge, and instead are thought to underlie capacities for abstract reasoning and problem-solving;² crystallized abilities are characterized by their relatively high dependence on an individual’s existing knowledge, acquisition of which is strongly culturally mediated, with crystallized measures thought to indicate one’s store of declarative knowledge and ability to effectively apply it to meet various challenges (see, e.g., Kan et al., 2011). Later developments in the theory of fluid and crystallized intelligence, based on

1 The reader should understand that in using the word “ability” and cognates we do not mean to denote a latent factor. We use “ability” as a very broad term for any cognitive variable reasonably interpreted as a capacity that supports performance on at least one type of intelligence-demanding task. Some abilities, but not all, are best modeled as latent factors, and not all latent factors should be understood as entities. See the section in the Introduction, *Formative vs. reflective models: A new level of explanation* for more on the latter point.

2 Certain fluid tests are more sensitive to the Flynn effect than crystallized ones because of their rule dependence; thus performance on such fluid tests is conditioned (to a degree) by diachronic changes in cultural knowledge concerning, e.g., specific problem-solving strategies (Armstrong & Woodley, 2014).

foundational contributions from psychologists John L. Horn and John B. Carroll (1993), in addition to Cattell, gave rise to the Cattell–Horn–Carroll (CHC) model of intelligence (Keith & Reynolds, 2010). Higher-order CHC models generally feature three levels of psychometric aggregation: the level of the general factor of cognitive ability (g); the level of group factors denoting clusters of conceptually similar abilities that are necessarily less general than the general factor, such as G_f and G_c , but also, for example, G_s (general processing speed) and G_{sm} (general short-term memory);³ and the level of subtests.⁴ Hierarchical CHC models contrast with others, such as Spearman’s model of g , which features only two levels of psychometric aggregation: the g factor and the subtests onto which the former loads.⁵

It is apparent that there are two, perhaps somewhat independent, “streams” of psychometric research offering (superficially but not necessarily) contradictory perspectives on whether crystallized or fluid intelligence relates more closely to g . The apparently standard view, in light of CHC models, holds that G_f best captures g , perhaps even being isomorphic with this factor.⁶ One impressively large multi-battery study specifies six group factors, including G_c and G_f , and states of the resulting model that “ G_f had the strongest loading on g and a non-significant residual, supporting previous research that these two constructs are not statistically distinguishable” (Caemmerer et al., 2020, p. 9). Nevertheless, there is clear, high-quality evidence that at the *subtest level*, g loading, crystallization, and culture loading (see footnote 5 for a definition of this term) are positively associated (Kan et al., 2013). Although Arthur

Jensen is often associated with a fluid-ability conception of g , he did not regard it as problematic that vocabulary tests are consistently among the most g -loaded subtests (Jensen, 1973, 2001).⁷ This notwithstanding, Kan (2011) argues that it is “contrary” to Jensen’s (and J. Philippe Rushton’s) “biological g theory” that “cultural loading and g loading may well be intrinsically positively related”—this is “contrary” to Jensen’s theory because, according to Kan, “[i]n biological g theory . . . [c]ulture-reduced [more fluid] tests have higher g loadings . . . than culture loaded [more crystallized] tests” (p. 54; emphasis in original). But it seems quite possible for biological g theory to accommodate this observation if it is posited that achieving mastery of and retaining (especially over the long term) knowledge received through the culture in which one is embedded are very cognitively demanding endeavors, with success in them strongly conditioned by the global biological quality of the brain.

The findings of these two “streams” of psychometric research are, *prima facie*, paradoxical. One would expect that if the culture loading/“crystallization” of subtests is “intrinsically positively related” to g , then if any group factor is to show virtual isomorphism with g , it ought to be G_c and certainly not G_f . Yet, as Jensen (1998) and others have noted, the opposite pattern holds at the group factor level, against the expectation that emerges from the subtest-level association between crystallization/culture loading and g loading. Kaufman et al. (2012), for example, note that “[a]lthough G_f was the strongest [Broad Ability] measure of g for both test batteries [considered in Kaufman

3 A typical number of group factors in a CHC model seems to be four or five (see, e.g., Reynolds et al., 2013; Weiss et al., 2013), but many models specify more (see, e.g., Bickley et al., 1995; Kaufman et al., 2012).

4 Sometimes this standard hierarchical CHC model is called a “three-stratum model” (see, e.g., Bickley et al., 1995), but it seems that what counts as a “stratum” is not understood in the same way across studies. In some studies, only levels of aggregation containing latent variables are defined as strata—for example, Floyd, Keith, Taub, and McGrew (2007) refer to models with a general, group-factor, and subtest level as “two-stratum” models, unlike Bickley et al. (1995) who call them “three-stratum” models. Carroll noted that although the word *stratum* originated with Cattell (1971) and the underlying idea resembled Thurstone’s (1947) notion of factor “order,” he used *stratum* in a different way—specifically to denote the range and heterogeneity of variables represented by a factor. As many datasets include few indicators at Stratum I, the broad abilities are still classified as Stratum II even when lower-level measures are limited (Carroll, 2005). Regarding the three-stratum framework, Carroll adopted three levels for reasons of parsimony and empirical adequacy; most CHC models follow this convention (McGrew, 2023), although Horn opposed the inclusion of a Stratum-III general factor. Notably, Bickley et al. (1995) evaluated Carroll’s three-stratum configuration prior to the later collaboration among Horn, Carroll, and Woodcock. We thank an anonymous reviewer for providing information that substantially enhanced this footnote.

5 It is necessary to understand at this point that the degree to which a subtest is “crystallized” or “fluid” is operationalized differently across studies. Typically, in CHC research, subtest crystallization/fluidity is determined with respect to whether and to what degree the G_c or G_f factor loads onto the subtest. A more precise approach involves some systematic process of determining the crystallization, knowledge dependence, or, somewhat distinctly, culture dependence of a subtest that requires thorough direct or indirect consideration of its actual content (see Kan et al., 2013, who define culture loading as the degree of adjustment made to a cognitive test to adapt it for use “from one language or culture to the next” [p. 2421], which can be operationalized, for example, as the percentage of items in a test that are adapted across versions for use in different cultural contexts; Kan et al., 2013 understand culture loading and crystallization to be very strongly related, as do the authors of the current paper). The latter procedure allows all subtests appropriately analyzed to be placed on a continuum of knowledge/culture “freedom”—knowledge/culture loading, whereas the former could suggest (misleadingly, depending on how “fluid” is defined) that every subtest on which neither G_c nor G_f loads simply is not on the crystallized-fluid spectrum. Still, those subtests that do cluster on G_c and G_f will have, respectively, higher and lower average knowledge/culture loading. One who understood test fluidity and crystallization in terms of CHC theory may reject the idea that fluidity simply is relative knowledge/culture “freedom” of a cognitive test, understanding it as more essentially defined by a specific sort of abstract reasoning that not all relatively culture-free cognitive tasks involve.

6 In reality, the statistical relationships between g and G_f that motivate this “isomorphism” claim simply show that g variance more or less fully accounts for the shared variance between subtests that compose the G_f factor in CHC models (Kan et al., 2011).

7 Among other explanations for such results, Jensen (1992) suggests that “complex cognitive learning that extends over long periods” (p. 282), precisely because of its great cumulative complexity, is highly g loaded.

et al.'s study], the G_c variables emerged as the best measures of g among the subtests Like the results for the Broad Abilities, these subtest results—namely, the highest g loadings by G_c subtests—were entirely consistent with the Wechsler literature and with empirical analyses of the KABC-II [Kaufman Assessment Battery for Children—Second Edition] and WJ III [Woodcock-Johnson III]” (pp. 133–134).

In terms of structural equation modeling, this general intelligence paradox (or g paradox) can be illustrated quite simply. One may ask: If the coefficient of the path from g (at Stratum III) $\rightarrow$ subtest (at Stratum I) tends to be larger in magnitude when estimated for subtests associated with the (Stratum-II) G_c group factor, why is it that the G_f group factor can be fixed to unity with g in many cases without losing model fit? The severity of the attenuation of the estimated $g \rightarrow$ subtest loading is negatively related to the magnitude of the Stratum-III $\rightarrow$ Stratum-II loading, as this is multiplied by the Stratum-II loading $\rightarrow$ Stratum-I (subtest-level) loading in order to recover subtest g loading. The degree of attenuation therefore should be much *lower* in the case of G_f -mediated paths, as this group factor is essentially isomorphic to Stratum-III g .

The g paradox arises in that, despite the foregoing, the G_c group factor loads to a greater degree onto its constituent subtests than does the G_f group factor onto its subtests (e.g., Caemmerer et al., 2020). Thus, the average g loading of G_c subtests is higher than the average g loading of G_f subtests, as the stronger $G_c \rightarrow$ subtest paths offset the almost perfect $g \rightarrow G_f$ path in cross-multiplication. Notably the finding of higher-magnitude g loadings in the case of the former persists even when subtest reliability is controlled (see Kan, 2011; Kan et al., 2013). Therefore, the paradox is not simply a function of the (often claimed) higher reliability of crystallized versus fluid subtests (e.g., Postlethwaite, 2011).

Attempted explanations of the paradox

Kan (2011) and Kan et al. (2013) employ investment theory (which holds that G_c is merely the result of G_f applied to the problem of knowledge acquisition over the life course) to account for what Jensen (2006) termed the *heritability paradox*, specifically “the frequent and seemingly surprising finding that h^2 increases as a function of task complexity and also as function of the degree to which tasks call for prior learned knowledge and skills” (p. 133). Jensen considered this paradoxical because performance on more complex tasks and more knowledge-based tasks are dependent to a greater degree on non-genetic factors (such as learning opportunities, cultural immersion, etc.) than are more elementary forms of cognitive processing (such as measures of reaction time) that are in turn “closer” to (what Jensen believed to be) the underlying strongly heritable mechanisms that give rise to g . A slight variant of the heritability paradox, one which results from findings that more knowledge-dependent tasks are often more g loaded than relatively culture/knowledge-free but complex tests of abstract reasoning, appears in the case of polygenic scores (PGSs) for cognitive phenotypes, which tend to correlate

more strongly with G_c than G_f measures (Loughnan et al., 2023). In relation to this paradox, Kan (2011) “suggest[s] that the explanation may lie in gene-environment correlation” (p. 17). He goes on to note that “crystallized ability tests happen to display the highest loadings on the general factor of intelligence” with “[g]eneral intelligence” (as a statistical construct) appear[ing] to be more like ‘crystallized intelligence’ than ‘fluid intelligence’” (p. 17). Kan et al. (2011) argue that:

[I]n investment theory crystallized intelligence is not a capacity but purely a statistical entity. We contend that if CHC factor G_c represents a capacity, it cannot represent crystallized intelligence, and if G_c represents crystallized intelligence, it does not represent a capacity ... from our discussion of G_c , we conclude that in investment theory the factors G_f and g represent one and the same capacity. (p. 292)

Kan et al.'s solution to the g paradox therefore is that, as G_f is g (assuming investment theory is correct), the high g loading of crystallized intelligence measures simply reflects their affinity for G_f : “Crystallized intelligence ... summarizes the crystallized abilities. It is not a substantive, causal variable. G_f accounts for the correlations between the crystallized abilities” (p. 295; see also Ashton & Lee, 2005).

The principal weakness of this solution is its reliance on Cattell's (1963) investment theory, for which direct empirical support remains limited. Perhaps the most thorough test of this theory to date did not support it: “Certain findings depart from the investment hypothesis and question the validity of Cattell's . . . theory. Is fluid ability invested in crystallized abilities? From our analyses, the answer seems to be negative” (Ferrer & McArdle, 2004, p. 949). In attempting to rescue their model from the implications of Ferrer and McArdle's (2004) work, Kan et al. (2011) suggest that factors other than G_f could co-contribute to the formation of crystallized ability, such as “exposure to information through education” (p. 300).

Another study that aims to make sense of the paradox is that of Pluck and Cerone (2021). These authors note that “[s]tatistically derived latent variables of fluid intelligence ... (almost) perfectly predict g ” (p. 3083). They further state that:

There are nevertheless some features of the positive manifold which are not so clearly explained by invoking high-level cognitive control mechanisms. In particular, the fact that crystallized intelligence, such as a person's vocabulary or their general knowledge, typically predicts more variation in g (i.e., it has overall stronger correlations with other non-verbal cognitive tasks) than any other tests, including laboratory tests of working memory or fluid intelligence. (p. 3083)

In describing what we call the g paradox, they observe that “[t]his [the high g loading of vocabulary measures] is the ‘elephant in the room’. Why should vocabulary tests, which appear rather effortless to perform, be so highly g loaded, if the explanation for the g factor is based on effortful fluid processing?” (p. 3083). As with Kan et al. (2011), Pluck and Cerone (2021) “default” to investment theory,

but they also offer novel perspective on the heritability paradox, specifically that:

[V]ocabulary ability is highly heritable, in fact the most heritable of the tests used in intelligence batteries It stands to reason that the actual lexical entries are learned, not inherited. Therefore, the quality or efficiency of the underlying conceptual systems appears to be the part under genetic transmission. In this sense, variation in lexical skill is somewhat heritable. (p. 3086)

These conceptual systems exhibit heritable variation and are more broadly part of the mind's "innate structure and processes" with "lexical entries ... fit[ted] into this conceptual system" (p. 3086). In other words, factors conditioning the integrity of the conceptual systems, adapted to language acquisition and other forms of learning, subserve knowledge development. The participation of these lexical "modules" in the generation of the positive manifold is hence claimed to reveal "a neglected but important role of lexical-conceptual knowledge in high-level, top-down domain-general cognitive processing" (p. 3082).

A difficulty for this hypothesis is that the modules underlying these conceptual systems should be neurologically and functionally well circumscribed, as per the most credible theories of modularity (see Fodor, 1983, for further discussion; see also Woodley of Menie & Sarraf, 2021). Functional magnetic resonance imaging (fMRI) data, however, suggest that G_c is associated with the quality of broad networks likely involving many distinct cognitive systems, which facilitate the acquisition and retention of knowledge (Genç et al., 2019). If G_c really were merely a product of quality variation with respect to a *specific* set of (well-circumscribed) neurological systems, this would likely have been revealed in such studies. Pluck and Cerone's (2021) proposed solution therefore does not solve the paradox.

Formative vs. reflective models: A new level of explanation

A key insight into the potential causes of the g paradox comes from research into *formative* versus *reflective* models of g . In formative models, genetic and environmental influences act on observed indicators, from which latent variables are constructed. The latent variable itself does not exert causal influence. Therefore, the latent variable is said to be *formed* from the action of causal factors that generate patterns of covariance among these indicators. The latent variable in this instance is not a *causal entity*—instead it summarizes the effects of these causal factors on its constituent components. Reflective models on the other hand reverse the pattern of mediation, such that the latent variable mediates the entirety of the effects of the causal fac-

tors on the indicators. In this instance the latent variable can be said to *reflect* the action of these causal factors and can therefore be described as a *causal entity* responsible for the covariance among its components (Bollen & Pearl, 2013; Borsboom et al., 2003; Bruins et al., 2023).

The *mutualism model* represents a major step toward incorporating these distinctions into descriptions of g as a phenomenon. It "implies a formative measurement model in which ' g ' is an index variable without a causal role" (van der Maas et al., 2014, p. 12). Specifically, g indexes the action of developmental influences that selectively reinforce associations between distinct cognitive processes, giving rise to the positive manifold (Van Der Maas et al., 2006). g cannot be therefore reduced to a single reflective entity, with a distinct neurological or genetic etiology; but van der Maas et al. (2014) "[n]ote that this line of reasoning does not apply to genetic and brain research on *components* of intelligence (for instance working memory) as these components often do have a realistic reflective interpretation. Working memory capacity may very well be based on specific and independently identifiable brain processes, even if g is not" (p. 14). Mutualism presupposes a heterarchy among these processes. We argue that this framework does not adequately capture the hierarchical structure implied by the g paradox, particularly the psychometric stratum or level dependence of the involved effects.

We argue instead that the G_c and G_f factors arise due to the action of relatively *distinct* genetic, environmental, neurological, and potentially other processes, manifesting at different *levels of psychometric aggregation*, with formative sources of covariance operating at Stratum I (or the subtest level) and being associated with G_c ,⁸ and reflective/causal entities occurring at Stratum-II G_f , which are isomorphic to and completely intersubstitutable with Stratum-III g (in the CHC model) (Kan et al., 2011).

Given the aforementioned observations, the g paradox is best explained as occurring due to the following scenario: While the G_f group factor corresponds to some cluster of entities, such as working memory, (other) executive functions, and abstract reasoning ability, which collectively yield a coherent "cognitive control" architecture (e.g., Chen et al., 2019), relevant subtests are individually unable to tap all of these simultaneously with sufficient depth because of constraints on realistic item difficulty and complexity (G_f testing demands sophisticated real-time information processing and problem solving on the part of the test-taker, and such demands can only be so great before items become unsolvable for too many people to be psychometrically useful), limiting the precision of the correspondence between the subtest and the relevant entity cluster. If, structurally, G_f is relatively circumscribed/regionalized, and if adequately testing single components of it in relative isolation can be very cognitively taxing, it is unsurprising

⁸ On the matter of G_c having a formative basis we agree with Kan et al. (2011). We disagree on the nature of the formative influences on G_c . We propose that it has a more substantively biological basis and is therefore a true capacity (not to be confused with an entity), as will be explained further in the Discussion, whereas, as already noted, Kan et al. subscribe to investment theory supplemented with posits of cultural and educational influences.

that global activation and measurement of everything that goes into G_f is challenging or impossible to achieve with any one subtest. This explains the relatively low average loading of G_f onto its constituent subtests (compared to G_c). By contrast, crystallized ability, which in its “essence” is *basically the sum total of declarative knowledge and effective use of it*, can be indexed well with a straightforward and one-dimensional testing modality, on the assumption that any test of a “broad” declarative knowledge domain will tend to reveal with high fidelity a person’s overall capacity to gain and retain such knowledge (e.g., vocabulary or information; see Weiss, Holdnack, & Saklofske, 2019 for similar arguments). Moreover, since solving items on a crystallized subtest typically places minor demands on attentional, working-memory, executive-function, reasoning, etc. capacities, often requiring merely the mental accessing and “regurgitation” of previously learned information, the same item-level constraints on fluid-ability testing do not apply in the case of crystallized-ability testing—what is needed is simply a sufficiently broad range of items to effectively proxy a person’s overall success in obtaining, retaining, and using declarative knowledge. This would explain why on average G_c loads so strongly onto its constituent subtests.

Spearman’s sandwich and two gs

As G_f subtests measure different aspects of an *entity*, the high-magnitude Stratum-III $g \rightarrow$ Stratum-II G_f path involves *reflective* factors. G_c , on the other hand, is clearly not an entity, but a broad *capacity* that can be more easily measured with items tapping proxies of overall information storage and effective use thereof. The strong indifference of G_c -salient subtests with respect to the measurement of this capacity aligns with the idea that the latter is *formative*, not reflecting any specific neurological or genetic entity,⁹ *hence when g loadings are estimated at the level of Stratum I*, crystallization (or culture loading) appears as a vector driving g loading across all subtests. We contend that these two *distinct* sources of covariance represent two effectively *distinct* gs, even though they overlap extensively in standard hierarchical models. To illustrate how these two gs can co-exist hierarchically we use the analogy of *Spearman’s sandwich* (illustrated in Figure 1).

As there is no meaningful distinction between Stratum-III g and Stratum-II G_f in CHC models (Caemmerer et al., 2020; Kan et al., 2011), this leaves an attenuated two-Stratum hierarchy of the two gs, with a formative Stratum-I g sitting immediately below a fully reflective Stratum-II g . The distinctions between the three different models discussed are illustrated in Figure 2.

Converging lines of evidence

Four major converging lines of evidence can be marshaled in support of this “two gs” model:

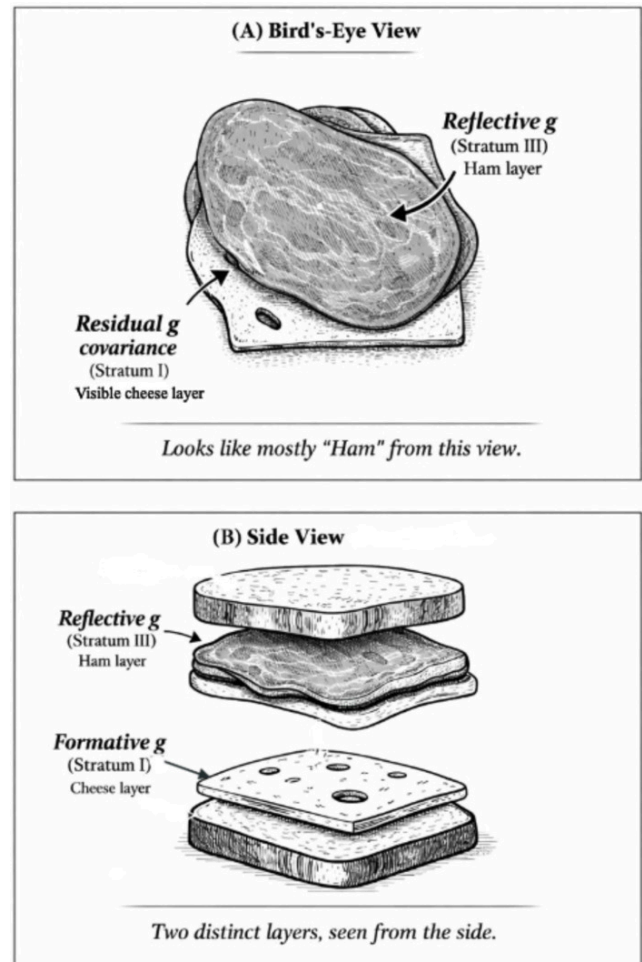


Figure 1. Spearman’s sandwich model.

Panel A: A slice of ham (analogous to reflective g at Stratum III in the CHC model) sits atop a slice of cheese (analogous to formative g at Stratum I). The former source of general covariance overlaps the latter, making only the “residual covariance” of the cheese (i.e., that which is not covered by the ham) visible when a bird’s-eye view (top-down latent variable) is taken. Panel B: When the layers of the sandwich are separated, the “fullness” of the (Stratum-I formative g) cheese layer becomes visible.

1. **Lack of evidence for G_c as an entity:** Kan et al. (2011) found, via confirmatory factor analysis, that “ G_f and g were statistically indistinguishable. G_c was effectively absent, because it was statistically equivalent to verbal comprehension. Factors G_c and g could be removed from the model without any reduction in model fit” (p. 292). They go on to argue “that in the CHC taxonomy the factors G_c and g are redundant as explanatory variables” (p. 292). This indicates that G_c , unlike G_f , is not an entity and is therefore purely a formative function of the action of factors acting via certain subtests.
2. **Strong associations between G_f and executive functioning:** G_f is intimately linked with executive functions that generate connections between distinct

⁹ It should be noted that emergent (or formative) factors *can* satisfy criteria associated with validity and so can be thought of as substantive (see Figueredo, 1999).

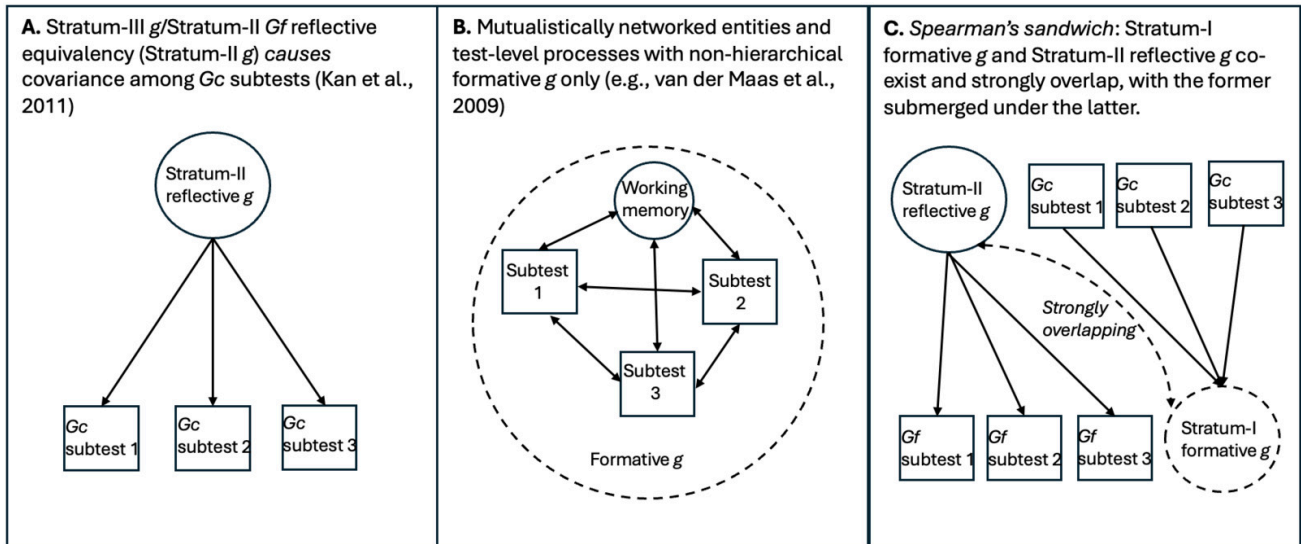


Figure 2.

Panel A: Kan et al.'s (2011) model: Stratum-III g /Stratum-II G_f isomorphic and fully reflective of covariance among G_c subtests. Panel B: Mutualism model: Networked entities and test-level processes yield a non-hierarchical formative g indexing overall network activity. Panel C: *Spearman's sandwich*: A formative Stratum-I (dashed circle) g emerges from the G_c subtests; a Stratum-II g factor is reflected in the covariance among G_f subtests; as the formative Stratum-I g exists at a lower-level of psychometric aggregation relative to its reflective Stratum-II counterpart, these two g s will overlap considerably, but *not* perfectly, suggesting that some discrete stratum-specific residual covariance associated with the former should be detectable (i.e., should be left over when the Stratum-II g is fully controlled).

mental capacities such as working memory and attention. This is evident in both the human (Chen et al., 2019; Conway et al., 2021; Salthouse & Pink, 2008) and non-human-animal literature (Burkart et al., 2017). Consistent with this, van Aken, Kessles, Wingbermühle, van der Veld, and Egger (2016) found “[a] very high correlation between G_f and EF [Executive Functioning] ... (0.91), with working memory being the most profound indicator.” However, only “[a] moderate to high correlation between G_c and EF was present” (all quotes from p. 31). These findings strengthen the prediction that these cognitive systems might be the source of G_f 's entityity.

3. **Discriminability in functional magnetic resonance imaging (fMRI) studies:** In terms of functional analysis of brain activity patterns, G_c and G_f are discriminable (Thiele et al., 2024). Moreover, the data are also highly suggestive of a formative basis for G_c in the relevant network architecture. Genç et al. (2019) found that a key component of G_c , general knowledge, “is heavily dependent on the *quality* of a widely distributed brain network that allows for an efficient integration of information” (p. 600; emphasis added). G_f on the other hand, is associated with neurologically well-circumscribed structures. In one fMRI study of G_f that of Ebisch et al. (2012), it is noted that “tasks [induction, visualization, and spatial relationships] activate a shared frontoparietal network. Specific activations were also observed, in particular for induction and visualization” (p. 331). This apparent distinctness at the level of generalized vs. localized activity patterns also suggests that the processes that undergird the formation of G_c are more robust to age-related degeneration (being highly re-

dundant) than are relatively neurologically delimited processes, such as those associated with G_f , potentially accounting for the different observed trajectories assumed by these two group factors with advancing age (McDonough et al., 2016; Yuan et al., 2018).

4. **Genetic discriminability:** Christoforou et al. (2014) used GWAS-based path analysis to differentiate between measures of fluid and crystallized intelligence. It has also been found that neuropsychiatric disorders exhibit different genetic associations with fluid and crystallized intelligence (Londono-Correa et al., 2025). This strengthens the inference of separate sources of covariance with different organic bases, which could correspond to distinct g s.

Predictions to be tested

First, an attempt will be made to establish whether, using one especially high-quality dataset (Caemmerer et al., 2020), the g paradox can be recovered via re-analysis of path models. Specifically, the aim is to determine if G_c subtests exhibit higher average g loadings than G_f subtests, despite G_f and g being “not statistically distinguishable” (p. 9) at Stratum II.

The work of Caemmerer et al. (2020) is noted for its exceptionally broad and deep coverage of mental abilities—66 subtests total from which six group factors were extractable: G_f , fluid intelligence; G_c , crystallized intelligence; G_s , processing speed; G_{wm} , working memory; G_v , visuospatial processing; G_l , learning efficiency. These subtests were sourced from six gold-standard intelligence test batteries: The Kaufman Assessment Battery for Children, Second Edition (KABC-II); the Wechsler Intelligence Scale for Children, Third, Fourth, and Fifth Editions (WISC-III, WISC-IV, and WISC-V); the Differential Abilities Scale, Sec-

ond Edition (DAS-II); and the Woodcock-Johnson Tests of Cognitive Abilities, Third Edition (WJ III). Moreover, data for these subtests were gathered in a body of samples comprising 3,927 youths (aged 6 to 18 years; these batteries have been found to exhibit strong measurement invariance with age). These were “*standardization and linking samples collected by Pearson Assessments during norming and validity studies of their intelligence measures*” (Caemmerer et al., 2020, p. 3; emphasis added), and therefore the quality of sampling is very high.

If the g paradox is present in Caemmerer et al.’s data, we would expect the following hypotheses to obtain:

H1: The correlation between Stratum-I or subtest g loading and degree of subtest affinity for the G_c group factor should be *positive*, and of *larger magnitude*, when compared to the affinity patterns involving the other five group factors. (“Affinity” is here a shorthand term for the degree of association between a subtest and a given group factor.)

H2: The rated “culture loadings” of the group factors should moderate the strength of these affinity patterns. This is a test of Kan’s (2011) thesis that so-called “culture loading” is a major source of variation in Stratum-I or subtest g loadings, with more culture-loaded subtests being more crystallized (see also Kan et al., 2013).

The rationale for *H1* and *H2*, in light of the theory here advanced about the nature of the g paradox, is that the more crystallized a subtest is, the more closely it should tap an individual’s ability to master information acquired through culture, and such mastery should depend on the quality of the brain’s global information-storage network (the formative Stratum-I g). This should yield a vector across subtests driving higher g loading as a function of higher crystallization and culture loading. The “fluidity”/“culture-freedom” vector, by contrast, should be weaker, given that single subtest measures of fluid ability will never be particularly strong measures of the germane construct, for reasons specified earlier in this Introduction, and that, given the assumption that fluid and crystallized intelligence are the primary sources of g , factors less related to both constructs should have weak or even negative associations with g .

These findings should be robust to controlling for subtest-level reliability, as per Kan et al. (2013).

The following additional hypothesis was tested via re-analysis of the covariance matrices provided to us by Jacqueline Caemmerer (note that only these matrices and no standardization data were shared):

H3: A key prediction of our model is that the “widely distributed brain network” that forms G_c (and thus formative Stratum-I g) is easier to activate on a single-measure basis than are the entities composing G_f (reflective Stratum-II g). To test this, we determine whether network analysis of the portions of Caemmerer et al.’s (2020) covariance matrices corresponding to the 12 “pure” (non-cross-loading) G_c subtests and the 12 “pure” G_f subtests allows distinct G_c and G_f networks to be estimated and compared on density,

weighted-strength, edge-weight, and node-centrality statistics. As the network model necessarily constrains the analysis to interrelations among subtests at Stratum-I, G_c (proxying a distinct g “factor” whose interrelations are driven by *formative* processes) is predicted to perform better than the corresponding G_f network on all of these network quality measures.

Research employing a number of large genotyped cohorts has found evidence that, in the case of g , both formative and reflective processes are at work, with phenotypic g mediating between 31% and 81% of the effects of genetic g factors (confirmatory factor-analytic latent variables comprising associated intelligence PGSSs; de la Fuente et al., 2021) on the relevant subtests in the different studies (see Woodley of Menie et al., 2025; for a similar study using a broader set of traits, see de la Fuente et al., 2025). Generally weaker residual *direct* positive effects of genetic g were also found on many subtests, suggesting the presence of discrete formative influences. In light of these observations another hypothesis will be tested:

H4: In the four SEM analyses conducted by Woodley of Menie et al. (2025), the direct paths from genetic g to the subtests should significantly contrast with respect to a variable capturing whether the subtest is classed broadly as G_c or G_f . Meta-analysis of the outputs of these four SEMs should yield indications of a significant contrast parameter across the set of studies, with genetic g loading onto G_c -designated subtests to a significantly greater degree than in the case of the G_f -designated ones.

This is another very direct test of the model, specifically that G_c emerges from the action of genetic factors acting *directly* on relevant subtests, which are *also* independent of the indirect influence of reflective phenotypic latent variables on said subtests. It should also be possible to examine the extent of the residual covariance (allowing for an estimation of the span of the formative g) with respect to those studies that exhibit the broadest nomological breadth.

Methods

Analysis 1 (Tests of H1 and H2)

Two separate analyses are performed to test *H1* and *H2*. For the first, Stratum-I g loadings are estimated using Caemmerer et al.’s (2020) confirmatory factor-analytic model (p. 8). This is done via path cross-multiplication, with the resultant coefficients then corrected for unreliability using reliability coefficients obtained from the sources of Caemmerer et al. (p. 4). To test for the paradox a simple vector correlation model is used in which a subtest is assigned a value of 1 if it is uniquely associated with the G_c group factor and 0 if it is associated with non- G_c group factors. In three cases, subtests exhibited two separate paths, one of which indicated an association with G_c . In these instances, the subtests were coded 0.5. A positive correlation would indicate that Stratum-II G_c affinity moderates Stratum-I (subtest) g loading. Similar analyses are performed using the other five group factors as contrast conditions

(G_f , G_{wm} , G_v , G_s , and G_I ; one subtest, onto which G_c did not load, is associated with three different group factors, so in its case, group-factor affinities are coded as 0.333). Finally, the six group factors are independently ranked in terms of their “culture loading” (based on separate ratings provided by two of the authors, informed by descriptions of the constituent subtests; this method is similar to the binary “subjective cultural loading” assignment approach of Gonthier et al., 2021) and the average of these ratings is correlated with the vector of the correlations between subtest Stratum-I g loadings and their “affinity” for each of the six group factors. This tests the degree to which the rated “culture loadings” of the group factors moderate the pattern of affinity— g loading correlations.

Analysis 2 (Test of H3)

In testing $H3$ the covariance matrices from Caemmerer et al. (2020) corresponding to the 12 non-cross-loading G_c and 12 non-cross-loading G_f subtests are transformed into adjacency matrices using an absolute value threshold approach. In turn, retained correlations are treated as weighted, undirected edges. The corresponding networks are generated with the *igraph* package (Csardi & Nepusz, 2006). Edge weight is estimated for each network and operationalized as the proportion of all possible edges that are present. Kolmogorov-Smirnov (KS) tests are computed on the nonzero edge weights and distributions of strength centrality across networks. Significance tests are used to determine whether there are differences between each of the two networks in relation to these four parameters.

To control for possible Type-I errors the G_c and G_f networks are allowed to be iteratively mixed with 50% random items across 1000 iterations, to determine whether their presence leads to the same observed network differences as those noted in the initial step. In each iteration, the edge weights are pooled, randomly permuted, divided into two groups of equal size, and used to create a null distribution of differences in densities. Additionally, weighted node strength is estimated for each network, and a significance test is performed across the mean strengths.

A major advantage in employing Caemmerer et al.'s (2020) data for these analyses is that each network has the same number of nodes (12), and in each case these are subtests with unique affinities for their respective group factors, making this a very fair comparison. Unless they are properly constrained, network models have a tendency towards overfitting (Costantini & Perugini, 2017). The use of network analysis here is theoretically appropriate also given the predicted properties of the germane constructs, specifically that when constrained to Stratum I, the covariance associated with formative influences should exhibit the stronger networking properties—a network being an al-

most pure representation of formative effects on covariance.

Analysis 3 (Test of H4)

For the test of $H4$, the SEM analyses involving four studies (English Longitudinal Study of Aging [ELSA], Midlife in the US Genetic [MIDUS G], Health and Retirement Study [HRS], and Harmonized Cognitive Assessment Protocol [HCAP]) from Woodley of Menie et al. (2025) are re-run using broad G_c or G_f categorical variables to detect differences in path means (via the generation of a contrast parameter). These analyses are conducted using *lavaan* 0.6-18 (Rosseel, 2012). (Two of the current authors [MAW and JGRF] coordinated in assigning each subtest in each study to one or the other category [coded 1 and 0].) Random effects meta-analysis of the contrast parameters is conducted using the *metafor* package (Viechtbauer, 2015). Supplementary Table S1 contains the “ G_c ” vs. “ G_f ” coding for each of the 46 subtests from the four studies sourced from Woodley of Menie et al. (2025), along with the rationale for assigning each subtest to one or the other category.

Results

Analysis 1 (Test of H1 and H2)

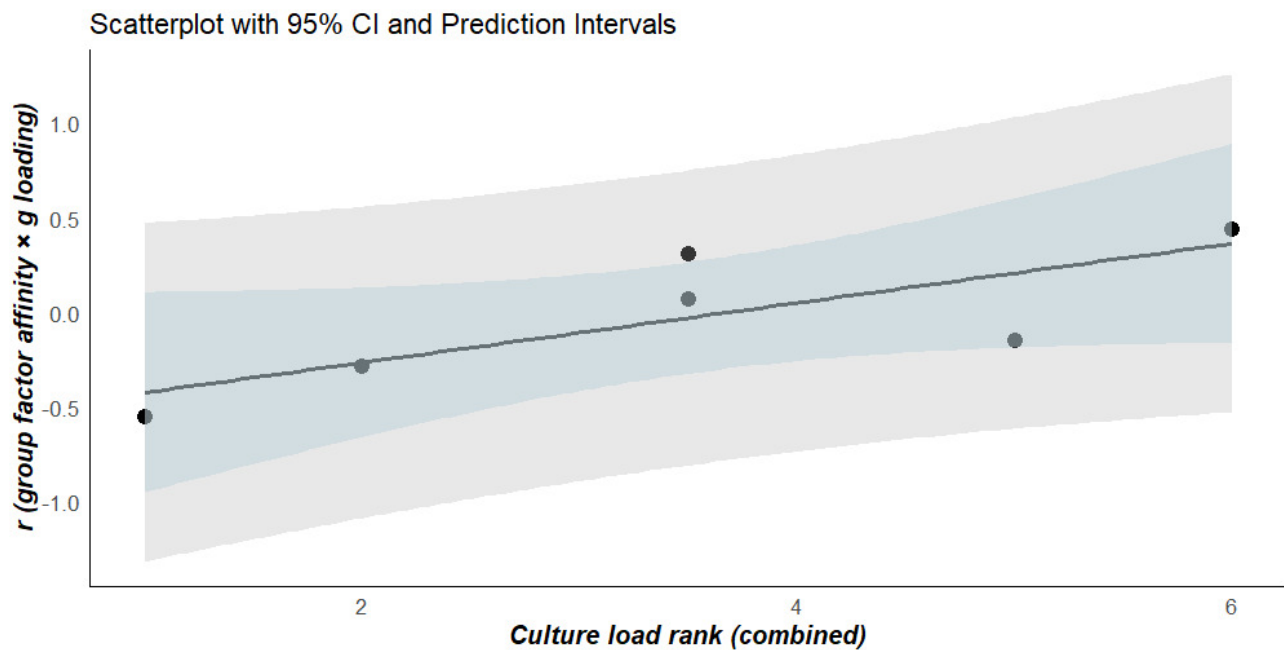
Consistent with the reality of the g paradox, (1) the average g loadings (derived via path cross-multiplication) of the 12 “pure” (non-cross-loading with any other group factor) G_c subtests is 0.71, versus 0.66 for the 12 corresponding G_f subtests (each subtest was corrected for its reliability coefficient); and (2) the loading of g onto the G_c group factor is 0.81, while the loading of g onto the G_f group factor is 0.99. The correlation between subtest (Stratum-I) g loadings and the affinities of each subtest with the (Stratum-II) group factors indicates, consistent with $H1$, that the G_c affinity pattern is the most positively associated with the (reliability-corrected) g loadings ($r_x=0.449$, 95% CI: 0.233 to 0.623). The broad factor that is clearly the most culture free, G_s , exhibits the exact opposite pattern consistent with the prediction that a dimension of crystallization is driving g loading at Stratum-I ($r_x=-0.532$, 95% CI: -0.685 to -0.333).

The “culture-loading” rankings of the group factors exhibit high inter-rater reliability ($r=0.943$, $n=6$, 95% CI: 0.561 to 0.993, $p<.05$, $K=2$ raters). The average of these “culture-loading” rankings has a directionally¹⁰ significant and positive correlation with the values estimated in the previous step ($r=.780$, $n=6$, $p_{(one-tailed)}=.034$), consistent with $H2$. Table 1 presents the two sets of “culture-loading” ranks, along with the average, and the g -loading-affinity correlations for each of the group factors. Figure 3 plots the results of testing $H2$.

10 Directionality was here assumed as the prediction of $H2$ was conditioned by the findings of Kan (2011) and Kan et al. (2013), who had already empirically demonstrated the positive correlation of culture loading, crystallization, and g loading.

Table 1. Rated “culture-loading” ranks for six group factors and (subtest-reliability-corrected) correlations (r_x) between Stratum-I average g loading and group factor affinities for six group factors.

Group factor	Culture loading rank (rater 1)	Culture loading rank (rater 2)	Culture loading rank (combined)	r_x (group factor affinity* g loading)
G_c	6	6	6	0.450*
G_f	3	4	3.5	0.303*
G_{wm}	2	2	2	-0.278*
G_v	4	3	3.5	0.060
G_s	1	1	1	-0.532*
G_l	5	5	5	-0.150

* $p < .05$, $n = 66$ subtests**Figure 3. Scatter plot graphing the relationship between “culture-loading” rank and the correlation between Stratum-I g loading and group-factor affinities.**

Analysis 2 (Test of H3)

Figures 4 and 5 depict the network structures of G_c and G_f respectively. The full names of all subtests are included in the Appendix (Table A1).

As shown in Table 2, the G_c network exhibits substantially higher density and weighted strength than the G_f network, indicating that a larger proportion of possible connections are present and that these connections are, on average, stronger. In addition, the distributions of edge weights and node centrality differ significantly between the two networks, as indicated by KS tests. Taken together, these results demonstrate that, under identical modeling constraints, G_c forms a more strongly connected and internally cohesive network than G_f without implying functional specialization or causal organization among individual nodes, consistent with H3.

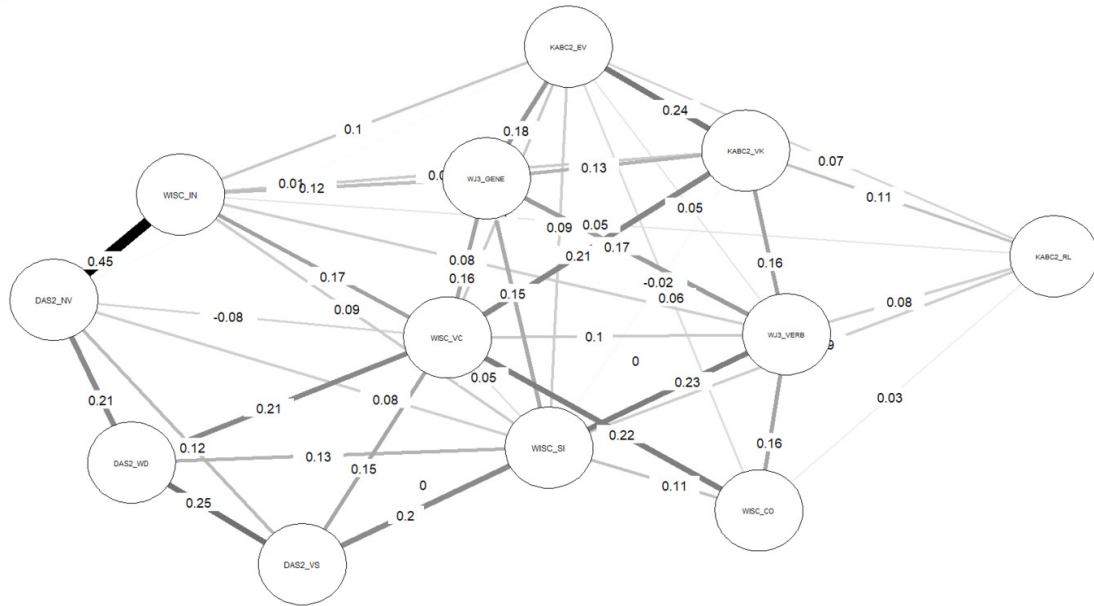
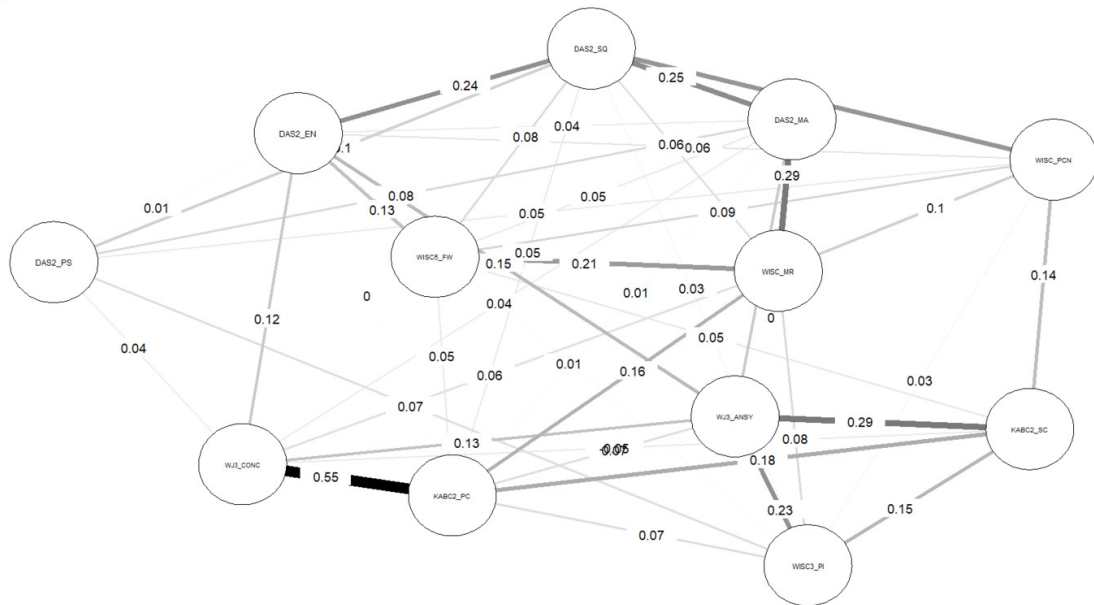
The robustness analysis indicates that the null distribution of density differences following 1000 mixed iterations was considerably smaller than the observed differences be-

tween G_c and G_f networks. As displayed in Figure 6, the observed density difference is located on the far-right tail of the null distribution, with the proportion of permuted differences that were equal to or exceeding the observed value remaining low in frequency. Consequently, these results strongly suggest that the observed density difference between G_c and G_f networks is not an artifact of random mixing or edge distribution, instead reflecting a robust structural difference between the networks.

Analysis 3 (Test of H4)

Table 3 presents the results of estimating the contrast parameters (difference in the means of the paths estimated among the G_c - and G_f -designated subtests) for each of the four datasets analyzed in Woodley of Menie et al. (2025), along with the random effects meta-analytic contrast parameter.

Various heterogeneity statistics were also estimated for the meta-analysis. The τ^2 value (estimated amount of to-

G_c NetworkFigure 4. Network structure of G_c . $K=12$ nodes.G_f NetworkFigure 5. Network structure of G_f . $K=12$ nodes.Table 2. Model comparison between the G_c and G_f networks.

Test	G_c value	G_f value	Difference	p -value
Density	0.970	0.545	0.424	<.0001
Weighted strength	6.033	2.463	3.570	<.0001
Edge-weight KS	0.548	0.224	0.697	<.0001
Node centrality KS	6.033	2.463	0.917	<.0001

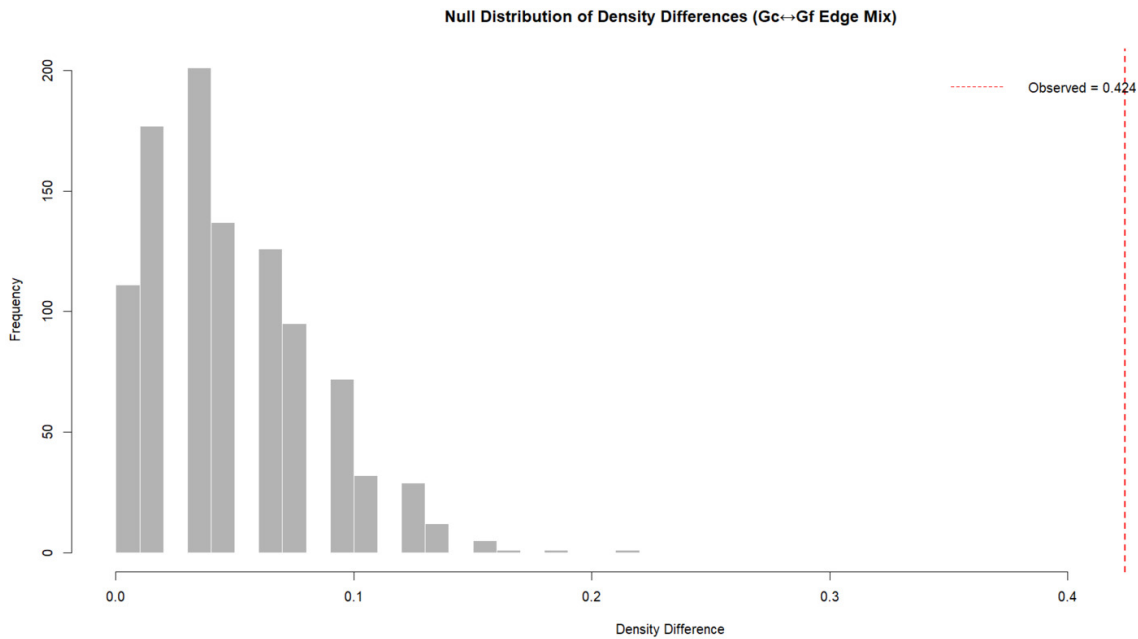


Figure 6. Null distribution of density differences following edge mixing of G_c and G_f networks.

Table 3. Contrast parameters (between “ G_c ” and “ G_f ” direct path means for four batteries, along with random effects meta-analytic estimate). SE = Standard Error.

Battery	Contrast parameter	SE	$p(> z)$
HCAP	0.035	0.008	<.0001
HRS	-0.004	0.007	.452
ELSA	0.063	0.011	<.0001
MIDUS G	0.122	0.048	<.0001
Meta-analytic value	0.04	0.019	<.05

tal heterogeneity) was 0.0012 (SE = 0.0013). The I^2 value (total heterogeneity / total variability) was 91.69%, and the H^2 value (total variability / sampling variability) was 12.04. The Q statistic (for $df = 3$) was 36.1168, $p < .0001$. These values indicate the presence of significant between-study heterogeneity.

Consistent with $H4$ the overall meta-analytic contrast parameter score is positively signed (meaning that the direct paths to G_c -assigned subtests are larger than those to G_f -assigned subtests) and is statistically significant. Moreover, the effect is significantly present in three of the four datasets.

Finally, we explore the nomological breadth of the positive manifold of the formative g factor with respect to a separate analysis of the HCAP dataset using unconstrained path estimation (this was not possible for the other datasets, as constraints were necessary for the other path models to achieve stability). Figure 7 illustrates the direct paths from the genetic g factor to the phenotypic manifest variables (estimated net of the phenotypic g factor; indirect paths through phenotypic g not shown). We argue that these genetic effects on the subtests are indexing the for-

mative Stratum-I g , hence we depict this as a latent variable *emerging* from the subtests.

70% of the subtests exhibit positively signed loadings. The likelihood of this degree of sign consistency being a chance finding is low ($p = .037$). The positive manifold of the formative g factor includes subtests that span classic crystallized (e.g., Verbal Series), classic fluid (e.g., Matrix Reasoning), working memory (e.g., Serial 7), long-term memory (e.g., Delay Story Recall), coding speed (e.g., Digit Symbol), fluency (e.g., Animal Naming), and executive functioning (e.g., Trail Making), suggesting very substantial nomological breadth, consistent with the existence of a potentially distinct g factor, and not merely a narrower factor of some kind, e.g. a group factor. The contrast parameter for this unconstrained model is statistically significant (0.029, SE=0.014, $p < .05$). See the second Supplement (S2) for the full HCAP output associated with the unconstrained model.

Discussion

The successful tests of $H1$ and $H2$ employing Caemmerer et al.'s (2020) data establish the reality of the g paradox

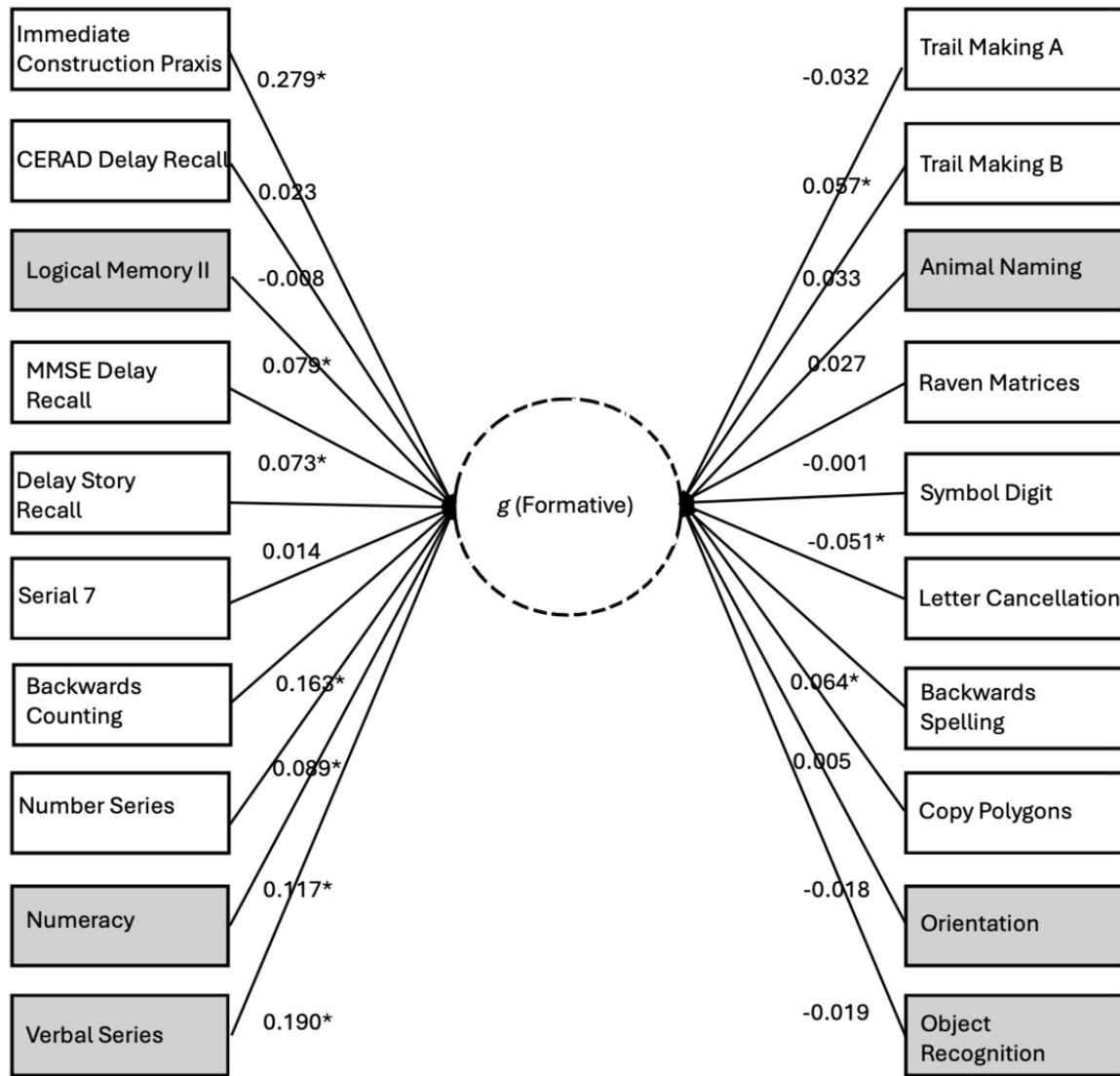


Figure 7. Direct paths from 20 subtests to a formative g factor (represented as a dashed circle, as per the convention established in Figure 2) in HCAP. Grey-shaded subtests were assigned to the G_c category, unshaded to the G_f category. * $p < .05$.

through conventional means, with the pattern of subtest-level (Stratum-I) g loadings indicating that the most g -loaded subtests are associated with the G_c group factor, which is also the most “culture-loaded” group factor. This pattern contrasts with that observed in the same data at the group factor level, with the Stratum-III g factor loading onto G_f more strongly compared to G_c (0.99 vs. 0.81). The difference in average subtest g loadings between G_c and G_f at Stratum I is numerically small (0.05). The theoretical importance of such effects, however, does not rest on their absolute magnitude, but on their direction, replicability (Kaufman et al., 2012), and structural inconsistency with standard hierarchical models. In a *strictly* reflective hierarchical model, this pattern is anomalous. Thus, even small systematic deviations are theoretically diagnostic because the null expectation is *zero difference*, not merely a small one.

The successful test of $H3$ indicates that subtests associated with G_c form a more robust network than do their G_f

counterparts. This is a much stronger test of the g paradox, as network analysis constrains comparisons to interrelations among subtests at Stratum-I (latent variables are not permitted to form). These findings are consistent with the prediction that G_c subtests tap widely distributed brain networks which contribute to the formation of the G_c (Stratum-I g) “factor,” whereas single measures of G_f by contrast are unable to tap all the germane reflective factors (associated with Stratum-II g) simultaneously with sufficient depth, leading to poorer network characteristics.

The successful test of $H4$ is consistent with the idea that genetic influences on the positive manifold that are formative, that is, having direct effects on subtests independently of latent-variable-mediated paths, should exhibit stronger associations with subtests that are categorically associated with G_c . The random effects meta-analytic contrast parameter is 0.04 ($p < .05$). The unconstrained HCAP model also yields a significant contrast parameter in addition to signals of a nomologically broad potential formative g fac-

tor. When interpreting this result, it is important to keep *Spearman's sandwich* (Figure 1) in mind. Recall that we are dealing with *residual covariance* (akin to that portion of the cheese that is not perfectly overlapped by the ham in the bird's-eye view or straightforward latent-variable model, which is analogous to the hierarchical Stratum-II g). Owing to the strong overlap between the Stratum-I and Stratum-II g s, subtracting the “ham” layer leaves only the “edges” (residual covariance) of the cheese, hence we do *not* expect the signal of a truly formative g factor to have the same *potency* as the reflective g factor. Instead, the sign suggesting the presence of a distinct source of general covariance comes from the nomological characteristics of this residual covariance structure, which, as we demonstrate in the case of the unconstrained HCAP model, are suggestive of a broadly positive manifold (70% sign concordance) weakly integrating many types of ability measure. This evinces that we are not merely dealing with a narrower group factor—the formative g is indeed a g , but the effects associated with it appear weak as an artifact of the residualization needed to statistically detect it. The remaining negatively signed effects may, at least in part, reflect instability in estimating this structure, again as anticipated given the small magnitude of residual effect sizes.

A potential difficulty for our solution to the g paradox can be addressed given the results of testing $H4$. Specifically, we have argued at length for the different natures of the two g s, such that each is more strongly associated with specific kinds of cognitive abilities (crystallized vs. fluid). A natural objection is that if these sources of covariance differ in fundamental substantive organization, it is unclear how both can qualify as general. The answer is that—with a very small number of potential exceptions for, e.g., unusually narrow executive-functioning-type tasks (which, as per our theory, would of course favor the reflective g)—the variance of almost all subtests is expected to have some formative g and some reflective g component. For example, while working memory tasks may specially depend on relatively circumscribed neurological modules, formative effects on the overall integrity of the brain will impinge on the functioning of such modules. And reflective systems, such as those underlying attention, surely are involved in the learning processes through which crystallized intelligence is developed. Both formative and reflective g , then, are broadly involved in cognition and tasks demanding abil-

ities exclusively in the domain of one or the other will be rare.

What forms Stratum-I g ? A closer look

Miller (2000b, 2000a) has postulated a key role for weakly deleterious but pleiotropic mutations in the *generation* of the positive manifold. Such mutations would act in a formative manner through their parallel effects on a wide range of cognitive systems that (in their absence) would tend to be uncorrelated. Thus, individual differences in *both* the level and strength of intercorrelations among outward manifestations of formative g should give rise to a gradient of *cognitive system integrity*. They should moreover yield a larger global *fitness factor* consisting of pleiotropically mediated phenotypic correlations involving physical and reproductive characteristics too (such as height, physical abnormalities, sperm quality, etc.) on which selection can act via (social-sorting-mediated) mate choice and even possibly at the level of sperm selection (Arden, Gottfredson, & Miller, 2009; Arden, Gottfredson, Miller, et al., 2009; Pierce et al., 2009; cf. DeLecce et al., 2020).

Miller (2000b) observes that “[t]he method of correlated vectors ... also shows that assortative mating is very focused on g -loaded mental traits. For example, spouses correlate more highly for their *vocabulary sizes* (a highly g -loaded trait) than they do for *digit span* (a modestly g -loaded trait)” (p. 264; emphasis added). Geary (2018, 2019) has identified a potentially critical mediating axis of system integrity in theorizing about a role for variation in mitochondrial numbers and efficiency in the generation of the positive manifold through their global effects on the regulation of cellular metabolism.

Furthermore, Miller (2000b) proposes that the reason that crystallized ability measures, such as vocabulary, are more (Stratum-I) g loaded and more heritable¹¹ is that they likely play a large role in conditioning human fitness in relation to social selection, since properly understanding and mentally and behaviorally managing group norms, communication systems (including language), and other culturally saturated phenomena would be key to social fitness (Miller's emphasis is on sexual selection, a kind of social selection, whereas other social-selective processes are likely of greater importance—see Peñaherrera-Aguirre et al., 2023).

11 The fact that more crystallized measures are typically more *heritable* provides strong evidence against arguments that the higher g loading of crystallized measures is merely artifactual (given the established positive association between subtest g loading and subtest heritability—see van Bloois et al., 2009 and Woodley of Menie et al., 2025; also see Kan, 2011 and Kan et al., 2013). Ashton and Lee (2005) seem to concede this (in a general way)—“[s]cores on crystallized subtests tend to be very highly correlated with external indicators of g ” (p. 440)—and so totally undermine their preceding long argument for the possibility that the association of g loading and subtest crystallization could merely reflect the bias of certain intelligence batteries toward crystallized ability measurement. They try to escape the difficulty they create for themselves by suggesting that, per investment theory, G_c subtests are, in effect, just tests of G_f applied over the life course up to the point of the intelligence test; they do not cite or discuss the work of Ferrer and McArdle (2004), which, as already noted, severely challenges this kind of investment theorizing. It must also be appreciated that Kan (2011), analyzing the extremely broad set of intelligence tests in the Minnesota Twin Study, demonstrates, in keeping with the findings of the current paper, a positive relationship between crystallization and g loading, further negating hypotheses that would attribute such an association to crystallized-biased test batteries.

Crucially, we do not deny the potential for factors associated with enculturation, such as education, to contribute to the formation of Stratum-I g (as noted, Kan et al., 2011 suggest that there may be non-investment pathways through which cultural effects can generate covariance). We argue that, ultimately, however, individual differences in condition-dependent (especially deleterious-variant-related) features of the brain's global processing networks are strongly determinative of the capacity to acquire knowledge. Compelling evidence to this end comes from Yeo, Ryman, Pommy, Thoma, and Jung (2016) who found that the g loadings among the subtests of a broad battery containing a Vocabulary measure positively moderated the magnitude of the subtest-by-total-cortical surface area (CSA) fluctuating asymmetry correlation. CSA fluctuating asymmetry is a broad index of neuro-developmental stability capturing the influence of, among other things, deleterious mutations. Thus, the most likely mechanism of general learning via Stratum-I g involves gene-environment correlation, with independent effects on covariance stemming from environmental-cultural variables likely playing a secondary role. Formal CFA analyses directly examining the role of subtests in mediating the effects of indices of mutations (or neuroanatomical proxies such as the aforementioned fluctuating asymmetry measure) on g should yield strong evidence for this claim.¹²

Finally, our solution to the g paradox has implications for the phylogenesis of the positive manifold. Stratum-I g may be primordial, and therefore highly phylogenetically conserved, being influenced by processes that globally regulate the activity of even very simple neural networks, such as those found in organisms lacking highly differentiated central nervous systems (see Peñaherrera-Aguirre et al., 2024 for relevant predictions concerning the presence of forms of primordial g in individual differences data on *Caenorhabditis elegans*). Stratum-II g , or the rudiments of it, by contrast, may have emerged in tandem with more complex and organized central nervous systems, such as those found in Arthropods, which appeared in the early Cambrian about half a billion years ago. Arthropod brains have the minimal neuro-cytoarchitectonic features necessary to accommodate more sophisticated forms of executive-functioning-type abilities or cognitive control systems—as such, Stratum-II g can be thought of as coordinating or integrating g given that such systems allow the regulation of the activity of lower-order networks and channels of inputs. The presence of mushroom bodies, which are found in the brains of certain Arthropods, for ex-

ample, have been implicated in performance on reversal-learning tasks (Peñaherrera-Aguirre et al., 2024). A greater elaboration on these more primitive structures would have given rise to the more robust suite of cognitive entities that can be readily identified as corresponding to Stratum-II g in mammalian intelligence. In light of the foregoing, we propose a new nomenclature for the two g s: *proto* g for the Stratum-I g , and *neo* g for the Stratum-II g ; these can be denoted g_p and g_n .

Limitations

While the CHC model is very widely accepted among psychometricians, it has its detractors, most notably proponents of the g -VPR (Verbal Perceptual Rotation) model (e.g., Johnson & Bouchard, 2005), who would doubtlessly object to our reliance on the G_c/G_f distinction here. It is possible that what we are claiming are two distinct g s may have an analogue in the g -VPR model context.

It might be that there are additional distinct sources of covariance at play, perhaps working through other CHC group factors—in particular, the degree to which the group factors of short-term memory and processing speed in the CHC model are reflective and/or formative entities, and how they relate to the full empirical reality of fluid ability (e.g., to what extent short-term, especially working, memory and processing speed are subsumed under fluid ability in reality, with the CHC short-term memory and processing speed factors possibly corresponding to more specialized aspects of these dimensions of cognition) are matters that require further investigation. In this vein, it is worth considering research providing evidence that “cognitive processing speed” and “cognitive processing accuracy” have discriminable neurological and genetic bases (Li et al., 2024). As it stands, the evidence presented through Analysis 1 suggests that the “fluid” and “crystallized” sources of general covariance are the primary ones, such that a two- g s model is adequate. But future lines of research could be developed to test these possibilities.

It must be appreciated that this work has had little to say about the g paradox at the level of developmental psychology. As noted in the Discussion, the theory of the evolution of g_p and g_n entails that, while the former emerged earlier in phylogenetic time, the rudiments of the latter are likely present in Arthropods and therefore may have existed in some form in animal taxa for around half a billion years (see predictions in Peñaherrera-Aguirre et al., 2024). The g paradox thus should be a manifestation of basic structural fea-

12 PGSS primarily index Single Nucleotide Polymorphisms (SNPs) that have additive effects on cognitive phenotypes (Lewis & Vassos, 2020). While these have been used to differentiate between G_f and G_c in some studies (e.g., Christoforou et al., 2014), they do not necessarily index variation in the burdens of rare variants (mutations), which have been found to account for a substantial proportion of the heritable genetic basis of g (Hill et al., 2018), and which are theorized here to be the major *formative* basis of G_c . Furthermore, Genome Wide Association Studies (GWASs) exclude mitochondrial genomic variants, which as discussed, may also play a key role in the formation of cognitive covariance. Nevertheless, PGSSs might be reasonably expected to proxy variation in these unmeasured genetic parameters, by virtue of being correlated with traits that are thought to partially index them (such as g and health-related phenotypes). Only indirect association with the relevant genetic variance might, however, be another contributing factor to the weakness of the formative Stratum-I g depicted in Figure 7.

tures of the human brain with deep evolutionary histories, so basic that it should be detectable across the life course (rather, from the point in the life course at which both fluid and crystallized intelligence can be reliably measured and on). Indeed, we have successfully tested predictions derived from our model in both child and adult samples in the current article. Still, there is undoubtedly much research to perform in the future to elucidate how the *g* paradox appears in human ontogeny. The current authors plan in a set of analyses for a future publication to test for the presence of the paradox in child and adult samples, using covariance matrices derived from Woodcock-Johnson I-III standardization data.

.....

Acknowledgments

The authors would like to thank Jacqueline Caemmerer for kindly providing the covariance matrices used in the paper of Caemmerer et al. (2020) for reanalysis here. AJ Figueredo and three anonymous reviewers offered feedback that substantially improved this article. MIDUS is supported by multiple grants from the National Institute on Aging (5R37AG027343, 5P01AG020166, 1R03AG046312, 1U19AG051426) and also by the University of Wisconsin Institute on Aging.

Funding

None of the authors received external funding to conduct this research. ELSA is funded by the National Institute on Aging (R01AG017644), and by UK Government Departments coordinated by the National Institute for Health and Care Research (NIHR). The HRS (Health and Retirement Study) and HCAP (Harmonised Cognitive Assessment Protocol) are sponsored by the National Institute on Aging (NIA U01AG009740) and is conducted by the University of Michigan.

Competing interests

The authors declare no financial interests in the outcome of this work.

Ethics statement

All analyses were based on publicly accessible databases that did not require special application or an IRB review.

Submitted: August 17, 2025 CDT. Accepted: January 12, 2026 CDT. Published: February 04, 2026 CDT.



This is an open-access article distributed under the terms of the Creative Commons Attribution 4.0 International License (CCBY-4.0). View this license's legal deed at <http://creativecommons.org/licenses/by/4.0> and legal code at <http://creativecommons.org/licenses/by/4.0/legalcode> for more information.

References

- Arden, R., Gottfredson, L. S., & Miller, G. (2009). Does a fitness factor contribute to the association between intelligence and health outcomes? Evidence from medical abnormality counts among 3654 US Veterans. *Intelligence*, 37, 581–591. <https://doi.org/10.1016/j.intell.2009.03.008>
- Arden, R., Gottfredson, L. S., Miller, G., & Pierce, A. (2009). Intelligence and semen quality are positively correlated. *Intelligence*, 37, 277–282. <https://doi.org/10.1016/j.intell.2008.11.001>
- Armstrong, E. L., & Woodley, M. A. (2014). The rule-dependence model explains the commonalities between the Flynn effect and IQ gains via retesting. *Learning & Individual Differences*, 29, 41–49. <https://doi.org/10.1016/j.lindif.2013.10.009>
- Ashton, M. C., & Lee, K. (2005). Problems with the method of correlated vectors. *Intelligence*, 33, 431–444. <https://doi.org/10.1016/j.intell.2004.12.004>
- Bickley, P. G., Keith, T. Z., & Wolfe, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the lifespan. *Intelligence*, 20, 309–328. [https://doi.org/10.1016/0160-2896\(95\)90013-6](https://doi.org/10.1016/0160-2896(95)90013-6)
- Bollen, K. A., & Pearl, J. (2013). Eight myths about causality and structural equation models. In S. Morgan (Ed.), *Handbook of causal analysis for social research* (pp. 301–328). Springer.
- Borsboom, D., Mellenbergh, G. J., & van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review*, 110, 203–219. <https://doi.org/10.1037/0033-295X.110.2.203>
- Bruins, S., Franic, S., Borsboom, D., Dolan, C., & Boomsma, D. (2023). Structural equation modeling in genetics. In R. H. Hoyle (Ed.), *Handbook of structural equation modelling* (2nd ed., pp. 646–663). Guilford Press.
- Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral & Brain Sciences*, 40, e192. <https://doi.org/10.1017/S0140525X16000959>
- Caemmerer, J. M., Keith, T. Z., & Reynolds, M. R. (2020). Beyond individual intelligence tests: Application of Cattell-Horn-Carroll theory. *Intelligence*, 79, 101433. <https://doi.org/10.1016/j.intell.2020.101433>
- Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge University Press. <https://doi.org/10.1017/CBO9780511571312>
- Carroll, J. B. (2005). The three-stratum theory of cognitive abilities. In D. P. Flanagan & P. L. Harrison (Eds.), *Contemporary Intellectual Assessment: Theories, Tests, and Issues* (pp. 69–76). The Guilford Press.
- Cattell, R. (1963). Theory of fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology*, 54, 1–22. <https://doi.org/10.1037/h0046743>
- Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Houghton Mifflin.
- Chen, Y., Spagna, A., Wu, T., Kim, T. H., Wu, Q., Chen, C., & Fan, J. (2019). Testing a cognitive control model of human intelligence. *Scientific Reports*, 9, 1–17. <https://doi.org/10.1038/s41598-019-39685-2>
- Christoforou, A., Espeseth, T., Davies, G., Fernandes, C. P., Giddaluru, S., Mattheisen, M., ... Le Hellard, S. (2014). GWAS-based pathway analysis differentiates between fluid and crystallized intelligence. *Genes, Brain & Behavior*, 13, 663–674. <https://doi.org/10.1111/gbb.12152>
- Conway, A. R., Kovacs, K., Hao, H., Rosales, K., & Snijder, J. (2021). Individual differences in attention and intelligence: A united cognitive/Psychometric approach. *Journal of Intelligence*, 9, Article 34. <https://doi.org/10.3390/jintelligence9030034>
- Costantini, G., & Perugini, M. (2017). Network analysis for psychological situations. In J. F. Rauthmann, R. A. Sherman, & D. C. Funder (Eds.), *The Oxford Handbook of Psychological Situations* (pp. 269–286). Oxford University Press.
- Csardi, G., & Nepusz, T. (2006). The igraph software package for complex network research. *International Journal of Complex Systems*, 1695, 1–9.
- de la Fuente, J., Davies, G., Grotzinger, A. D., Tucker-Drob, E. M., & Deary, I. J. (2021). A general dimension of genetic sharing across diverse cognitive traits inferred from molecular data. *Nature Human Behaviour*, 5, 49–58. <https://doi.org/10.1038/s41562-020-00936-2>
- de la Fuente, J., Londoño-Correa, D., & Tucker-Drob, E. M. (2025). Distinguishing specific from broad genetic associations between external correlates and common factors. *Bioinformatics*, 41, btaf568.
- DeLecce, T., Fink, B., Shackelford, T., & Abed, M. G. (2020). No evidence for a relationship between intelligence and ejaculate quality. *Evolutionary Psychology*, 18, 1474704920960450. <https://doi.org/10.1177/1474704920960450>
- Ebisch, S. J., Perrucci, M. G., Mercuri, P., Romanelli, R., Mantini, D., Romani, G. L., ... Saggino, A. (2012). Common and unique neuro-functional basis of induction, visualization, and spatial relationships as cognitive components of fluid intelligence. *NeuroImage*, 62, 331–342. <https://doi.org/10.1016/j.neuroimage.2012.04.053>
- Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, 40, 935–952. <https://doi.org/10.1037/0012-1649.40.6.935>
- Figueredo, A. J. (1999). Structural parsimony as an operationalization of the criteria for the entitivity of emergent variables [Paper]. In A. J. Figueredo (Chair), *The entitivity of emergent variables II: Applicable quantitative methods*. Panel. American Evaluation Association Annual Meeting, Orlando, Florida.

- Floyd, R. G., Keith, T. Z., Taub, G. E., & McGrew, K. S. (2007). Cattell-Horn-Carroll cognitive abilities and their effects on reading decoding skills: g has indirect effects, more specific abilities have direct effects. *School Psychology Quarterly*, 22, 200–233. <https://doi.org/10.1037/1045-3830.22.2.200>
- Fodor, J. A. (1983). *The modularity of mind*. MIT Press. <https://doi.org/10.7551/mitpress/4737.001.0001>
- Geary, D. C. (2018). Efficiency of mitochondrial functioning as the fundamental biological mechanism of general intelligence (g). *Psychological Review*, 125, 1028–1050. <https://doi.org/10.1037/rev0000124>
- Geary, D. C. (2019). Mitochondria as the Linchpin of General Intelligence and the Link between g, Health, and Aging. *Journal of Intelligence*, 7, 25. <https://doi.org/10.3390/jintelligence7040025>
- Genç, E., Fraenz, C., Schlüter, C., Friedrich, P., Voelkle, M. C., Hossiep, R., & Güntürkün, O. (2019). The neural architecture of general knowledge. *European Journal of Personality*, 33, 589–605. <https://doi.org/10.1002/per.2217>
- Gonthier, C., Grégoire, J., & Besançon, M. (2021). No negative Flynn effect in France: Why variations of intelligence should not be assessed using tests based on cultural knowledge. *Intelligence*, 84, 101512. <https://doi.org/10.1016/j.intell.2020.101512>
- Hill, W. D., Arslan, R. C., Xia, C., Luciano, M., Amador, C., Navarro, P., & Penke, L. (2018). Genomic analysis of family data reveals additional genetic effects on intelligence and personality. *Molecular Psychiatry*, 23, 2347–2362. <https://doi.org/10.1038/s41380-017-0005-1>
- Jensen, A. R. (1973). *Educability and group differences*. Harper & Row.
- Jensen, A. R. (1992). Understanding g in terms of information processing. *Educational Psychology Review*, 4, 271–308. <https://doi.org/10.1007/BF01417874>
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.
- Jensen, A. R. (2001). Vocabulary and general intelligence. *Behavioral & Brain Sciences*, 24, 1109–1110. <https://doi.org/10.1017/S0140525X01280133>
- Jensen, A. R. (2006). *Clocking the Mind: Mental Chronometry and Individual Differences*. Elsevier.
- Johnson, W., & Bouchard, T. J., Jr. (2005). The structure of human intelligence: It is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393–416. <https://doi.org/10.1016/j.intell.2004.12.002>
- Kan, K.-J. (2011). *The nature of nurture: The role of gene-environment interplay in the development of intelligence* [Doctoral dissertation]. University of Amsterdam.
- Kan, K.-J., Kievit, R. A., Dolan, C., & van der Maas, H. V. (2011). On the interpretation of the CHC factor Gc. *Intelligence*, 39, 292–302. <https://doi.org/10.1016/j.intell.2011.05.003>
- Kan, K.-J., Wicherts, J. M., Dolan, C. V., & van der Maas, H. L. J. (2013). On the nature and nurture of intelligence and specific cognitive abilities: The more heritable, the more culture dependent. *Psychological Science*, 24, 2420–2428. <https://doi.org/10.1177/0956797613493292>
- Kaufman, S. B., Reynolds, M. R., Liu, X., Kaufman, A. S., & McGrew, K. S. (2012). Are cognitive g and academic achievement g one and the same g? An exploration on the Woodcock-Johnson and Kaufman tests. *Intelligence*, 40, 123–138. <https://doi.org/10.1016/j.intell.2012.01.009>
- Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools*, 47, 635–650. <https://doi.org/10.1002/pits.20496>
- Lewis, C. M., & Vassos, E. (2020). Polygenic risk scores: from research tools to clinical instruments. *Genome Medicine*, 12, 44. <https://doi.org/10.1186/s13073-020-00742-5>
- Li, M., Dang, X., Chen, Y., Chen, Z., Xu, X., Zhao, Z., & Wu, D. (2024). Cognitive processing speed and accuracy are intrinsically different in genetic architecture and brain phenotypes. *Nature Communications*, 15, 7786. <https://doi.org/10.1038/s41467-024-52222-8>
- Londono-Correa, D., de la Fuente, J., Davies, G., Cox, S., Deary, I., Harden, K., & Tucker-Drob, E. (2025). Crystallized and fluid cognitive abilities have different genetic associations with neuropsychiatric disorders. *Research Square*, rs-3. <https://doi.org/10.21203/rs.3.rs-5256724/v1>
- Loughnan, R. J., Palmer, C. E., Thompson, W. K., Dale, A. M., Jernigan, T. L., & Chieh Fan, C. (2023). Intelligence polygenic score is more predictive of crystallized measures: Evidence from the adolescent brain cognitive development (ABCD) study. *Psychological Science*, 34, 714–725. <https://doi.org/10.1177/09567976231160702>
- McDonough, I. M., Bischof, G. N., Kennedy, K. M., Rodrigue, K. M., Farrell, M. E., & Park, D. C. (2016). Discrepancies between fluid and crystallized ability in healthy adults: a behavioral marker of preclinical Alzheimer's disease. *Neurobiology of Aging*, 46, 68–75. <https://doi.org/10.1016/j.neurobiolaging.2016.06.011>
- McGrew, K. S. (2023). Carroll's Three-Stratum (3S) Cognitive ability theory at 30 years: Impact, 3S-CHC theory clarification, structural replication, and cognitive-achievement psychometric network analysis extension. *Journal of Intelligence*, 11, 32. <https://doi.org/10.3390/jintelligence11020032>
- Miller, G. F. (2000a). Mental traits as fitness indicators: Expanding evolutionary psychology's adaptationism. *Annals of the New York Academy of Sciences*, 907, 62–74. <https://doi.org/10.1111/j.1749-6632.2000.tb06616.x>
- Miller, G. F. (2000b). Sexual selection for indicators of intelligence. In G. R. Bock, J. A. Goode, & K. Webb (Eds.), *The nature of intelligence*. Novartis Foundation Symposium 233 (pp. 260–275). Wiley Ltd. <https://doi.org/10.1002/0470870850.ch16>

- Peñaherrera-Aguirre, M., Sarraf, M. A., Woodley of Menie, M. A., & Figueredo, A. J. (2024). Possible evidence for the Law of General Intelligence in honeybees (*Apis mellifera*). *Intelligence*, 106, 101856. <https://doi.org/10.1016/j.intell.2024.101856>
- Peñaherrera-Aguirre, M., Sarraf, M. A., Woodley of Menie, M. A., & Miller, G. F. (2023). The ten-million-year explosion: Paleocognitive reconstructions of domain-general cognitive ability (G) in extant primates. *Intelligence*, 101, 101795. <https://doi.org/10.1016/j.intell.2023.101795>
- Pierce, A., Miller, G., Arden, R., & Gottfredson, L. S. (2009). Why is intelligence correlated with semen quality?: Biochemical pathways common to sperm and neuron function and their vulnerability to pleiotropic mutations. *Communicative & Integrative Biology*, 2, 385–387. <https://doi.org/10.4161/cib.2.5.8716>
- Pluck, G., & Cerone, A. (2021). A demonstration of the positive manifold of cognitive test intercorrelations, and how it relates to general intelligence, modularity, and lexical knowledge. In T. Fitch, C. Lamm, H. Leder, & K. Teßmar-Raible (Eds.), *Proceedings of the 43rd Annual Conference of the Cognitive Science Society* (pp. 3082–3088). Cognitive Science Society.
- Postlethwaite, B. E. (2011). *Fluid ability, crystallized ability, and performance across multiple domains: a meta-analysis* [Doctoral Thesis]. University of Iowa.
- Reynolds, M. R., Keith, T. Z., Flanagan, D. P., & Alfonso, V. C. (2013). A cross-battery, reference variable, confirmatory factor analytic investigation of the CHC taxonomy. *Journal of School Psychology*, 51, 535–555. <https://doi.org/10.1016/j.jsp.2013.02.003>
- Rosseel, Y. (2012). Llavaan: An R package for structural equation modeling. *Journal of Statistical Software*, 48, 1–36.
- Salthouse, T. A., & Pink, J. E. (2008). Why is working memory related to fluid intelligence? *Psychonomic Bulletin & Review*, 15, 364–371. <https://doi.org/10.3758/PBR.15.2.364>
- Thiele, J. A., Faskowitz, J., Sporns, O., & Hilger, K. (2024). Choosing explanation over performance: Insights from machine learning-based prediction of human intelligence from brain connectivity. *PNAS Nexus*, 3, pgae519. <https://doi.org/10.1093/pnasnexus/pgae519>
- Thurstone, L. L. (1947). *Multiple factor analysis*. University of Chicago Press.
- van Aken, L., Kessels, R. P., Wingbermühle, E., van der Veld, W. M., & Egger, J. I. (2016). Fluid intelligence and executive functioning more alike than different? *Acta Neuropsychiatrica*, 28, 31–37. <https://doi.org/10.1017/neu.2015.46>
- van Bloois, R. M., Geutjes, L. L., te Nijenhuis, J., & de Pater, I. E. (2009, December). *g loadings and their true score correlations with heritability coefficients, giftedness, and mental retardation: Three psychometric meta-analyses* [Paper]. 10th Annual Meeting of the International Society for Intelligence Research, Madrid, Spain.
- Van Der Maas, H. L. J., Dolan, C. V., Grasman, R. 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. <https://doi.org/10.1037/0033-295X.113.4.842>
- van der Maas, H. L. J., Kan, K.-J., & Borsboom, D. (2014). Intelligence is what the intelligence test measures. Seriously. *Journal of Intelligence*, 2, 12–15. <https://doi.org/10.3390/jintelligence2010012>
- Viechtbauer, W. (2015). Package “metafor.” The comprehensive R archive network.
- Weiss, L. G., Keith, T. Z., Zhu, J., & Chen, H. (2013). WISC-IV and clinical validation of the four- and five-factors interpretative approaches. *Journal of Psychoeducational Assessment*, 31, 114–131. <https://doi.org/10.1177/0734282913478032>
- Woodley of Menie, M. A., Peñaherrera-Aguirre, M., & Fuerst, J. G. R. (2025). Evidence that phenotypic g is both formative and reflective from four, large genetically-informed samples. *Twin Research & Human Genetics*, 5, 1–15.
- Woodley of Menie, M. A., & Sarraf, M. A. (2021). Controversies in evolutionary psychology. In T. K. Shackelford & V. A. Weekes-Shackelford (Eds.), *Encyclopedia of Evolutionary Psychological Science* (pp. 1399–1420). Springer.
- Yeo, R. A., Ryman, S. G., Pommy, J., Thoma, R. J., & Jung, R. E. (2016). General cognitive ability and fluctuating asymmetry of brain surface area. *Intelligence*, 56, 93–98. <https://doi.org/10.1016/j.intell.2016.03.002>
- Yuan, P., Voelkle, M. C., & Raz, N. (2018). Fluid intelligence and gross structural properties of the cerebral cortex in middle-aged and older adults: A multi-occasion longitudinal study. *NeuroImage*, 172, 21–30. <https://doi.org/10.1016/j.neuroimage.2018.01.032>

Appendix

Table A1. Glossary of terms corresponding to the 24 subtests examined in the network analysis along with their associated group factors.

<i>Code</i>	<i>Dimension</i>	<i>Description</i>
DAS2 NV	G _c	Naming vocabulary
DAS2 VS	G _c	Verbal similarities
DAS2 WD	G _c	Word definitions
KABC2 EV	G _c	Expressive vocabulary
KABC2 RL	G _c	Riddles
KABC2 VK	G _c	Verbal knowledge
WISC CO	G _c	Comprehension
WISC IN	G _c	Information
WISC SI	G _c	Similarities
WISC VC	G _c	Vocabulary
WJ3 GENE	G _c	General information
WJ3 VERB	G _c	Verbal comprehension
DAS2 EN	G _f	Early number concepts
DAS2 MA	G _f	Matrices
DAS2 PS	G _f	Picture similarities
DAS2 SQ	G _f	Sequential and quantitative reasoning
KABC2 PC	G _f	Pattern reasoning
KABC2 SC	G _f	Story completion
WISC MR	G _f	Matrix reasoning
WISC PCN	G _f	Picture concepts
WISC3 PI	G _f	Picture arrangement
WISC5 FW	G _f	Figure weights
WJ3 ANSY	G _f	Analysis-synthesis
WJ3 CON	G _f	Concept formation

Supplementary Materials

Supplement 1 (Table)

Download: <https://icajournal.scholasticahq.com/article/155379-more-than-just-1-g-the-general-intelligence-paradox-and-its-solution/attachment/325173.docx>

HCAP output (Supplement 2)

Download: <https://icajournal.scholasticahq.com/article/155379-more-than-just-1-g-the-general-intelligence-paradox-and-its-solution/attachment/325174.txt>
