Crystallized Intelligence: Exploring the Dark Matter of Intelligence

Ulrich Schroeders
Department of Psychology, University of Kassel
Email: schroeders@uni-kassel.de

September 6, 2025

This is a preprint version of a chapter going to be published in *A Research Agenda for Intelligence*, edited by Gilles E. Gignac and published by Edward Elgar Publishing Ltd.

Contents

Do you know the capital of Mozambique? What does the word Надежда mean? Which author voluntarily declined the Nobel Prize in Literature? Although these questions may appear unrelated, they could all serve as indicators of crystallized intelligence (gc). At the same time, they illustrate some of the challenges involved in defining, assessing, and modeling gc. Several questions arise: How can such diverse content be traced back to a single underlying construct? What role do language skills play? And does the knowledge really reflect widely shared cultural content or more specialized expertise? It also becomes evident that the likelihood of answering the above questions correctly depends on factors such as birthplace, native language, and year of birth. Someone born in Maputo, a Russian speaker, or a person who witnessed the media controversy in 1964, when Sartre famously declined the Nobel Prize on the day it was announced, will likely know the correct answers. A fair assessment of gc across national borders, language groups, and historical eras is therefore challenging. If the construct of gc is so multifaceted, difficult to measure fairly, and often vague in its operationalization, why should we study gc at all? In this chapter, we aim to address this question by first outlining prominent theoretical perspectives on gc (see Section 1), followed by a discussion about its structure and organization (see Section 2). The subsequent sections examine gc from a developmental perspective (see Section 3) and review empirical findings on its determinants and consequences (see Section 4). We then turn to the main focus of the chapter—the measurement and psychometric modeling of gc (see Section 5). Finally, we discuss how large language models can inspire future work on how to understand, measure, and model gc (see Section 6). Taken together, these perspectives demonstrate that gc remains not only a central pillar of intelligence research but also a construct whose mysteries can be re-examined in light of advances in artificial intelligence and large language models.

1 Nature and Scope of gc

Gc is a prominent ability construct in major theories of the structure of intelligence. It features prominently in the extended Gf-Gc theory, in Carroll's Three-Stratum Theory, and, accordingly, in the integrated Cattell-Horn-Carroll (CHC) theory of intelligence (McGrew, 2009; McGrew et al., 2023), where it stands alongside fluid intelligence (gf) as one of the most dominant factors. Moreover, virtually all other theories of cognitive abilities recognize the importance of the competencies associated with gc, even if they use different terminology. Essentially, tests of gc predated the development of intelligence tests in their current form. Intelligence tests, with their strong emphasis on gf, can be seen as a countermodel to traditional assessments of school readiness, as the first scientifically grounded intelligence test was specifically designed to move beyond earlier approaches that relied on evaluating acquired knowledge (i.e., the pedagogical method; Binet & Simon, 1904). However, the distinction between knowledge and reasoning ability predates psychology as a scientific discipline. Already Aristotle made the distinction between a person's innate potential (δυναμις (dynamis) – potential, capacity, disposition) and the knowledge to be acquired through experience and instruction (ἐπιστήμη (epistēmē) – structured knowledge). And this tension between fluid and crystallized abilities remains one of the defining themes in intelligence research.

Despite a long research tradition on gc, major intelligence frameworks and test batteries differ in their definition and operationalization of gc. Cattell's original definition of crystallized intelligence was broad, encompassing skills and knowledge across diverse content domains (Cattell, 1987). What unites these skill sets is that they reflect the outcomes of prior learning and its application. In line with this emphasis on acquired knowledge, Horn (2008) described gc as an ability factor that captures acculturated knowledge. In contrast, in Carroll's (1993) Three-Stratum Theory, gc is mainly defined by language-related abilities. Below the Stratum II factor, he summarizes lexical knowledge, listening and reading comprehension, spelling, and writing abilities in the native language and in foreign languages. Accordingly, in his synthesis of both major theories, McGrew (2009) characterizes gc as culturally acquired knowledge gained through acculturation, encompassing both the knowledge and application of language, information, and concepts specific to a given culture. In addition, McGrew also discusses several tentatively identified broad factors in the publication, among them grw (= reading writing) and gkn (= (general) domain-specific knowledge). Grw skills are intended to be distinct from general cultural knowledge, yet seem to reflect competencies emphasized in school-based curricula (and the respective indicators typically used in educational large-scale assessments), whereas gkn appears to represent a residual category of highly specialized knowledge (e.g., mechanical knowledge or geography achievement; McGrew, 2009). In particular, the distinction between gc and gkn is to some extent artificial, as these factors are better understood as lying on a continuum. Certainly, there are forms of knowledge so specific that they clearly fall outside of what a society values as common, shared knowledge (e.g., American Sign Language), while others are unambiguously part of it (e.g., basic physical laws). Yet for a substantial proportion of items the localization on the continuum is less clear, and judgments are subject to cultural and temporal change. For example, prior to the COVID-19 pandemic, the term replication rate was specialized knowledge largely restricted to epidemiologists or mathematicians, but the term was quickly integrated into the cultural knowledge. As a starting point for locating an item on this gc–gkn continuum, an item may be considered a meaningful indicator of gc when it can reasonably be assumed that all members of a population have had at least some, and ideally equal, exposure to its content. This is why Cattell (1987, p. 144) suggested that gc in adulthood could be assessed by a "school version". By contrast, items assessing highly specialized knowledge are more likely to reflect non-generalizable, idiosyncratic learning experiences. As a heuristic for item assignment, correlations with other constructs may be useful: whereas gc items tend to correlate strongly with g and can be assessed in the general population, gkn is typically developed through an individual's work experience, interests, passions or life circumstances and should be considered only in individuals with the relevant learning background (Flanagan & Dixon, 2014).

The strong association between verbal comprehension and gc is also taken up by Kan et al. (2011), who pointed to a seeming inconsistency: according to Cattell's investment theory, gc is not a psychological capacity per se but rather a statistical abstraction summarizing accumulated knowledge across domains. If, however, gc is interpreted as a substantive capacity within the CHC framework, this view conflicts with the investment theory, which holds that gc arises through the investment of gf. Kan and colleagues argue and substantiate using confirmatory factor analysis that gc is identical to a verbal comprehension factor. There are several critical points to raise: The authors reanalyze a dataset that exclusively relies on verbal indicators to measure gc and explicitly exclude domain-specific knowledge. Although this approach mirrors a widespread practice in the assessment of gc, it significantly narrows the scope of the construct. Moreover, they partialed out education from the indicators, significantly changing the construct, because by definition, gc reflects the accumulation of learning and knowledge. If education, as its primary developmental driver, is partialed out, what remains is a residual factor with reduced variance that is poorly defined and difficult to interpret. While such a residual may still capture elements of informal learning and everyday knowledge acquisition, it represents only a narrow facet of gc and may approximate verbal ability more closely. Finally, the sample is restricted to 483 male U.S. Naval recruits aged 17 to 22, which limits the generalizability of the results.

Carroll (1993) repeatedly emphasized the strong association between knowledge assessments and language measures, and suggested that this close relationship may result both from similarities in how these constructs are operationalized, overlapping acquisition processes (i.e., reading), and common methods of measurement. In this context, one could envision gc tasks that rely on visual rather than verbal material—for instance, identifying countries from their silhouettes or flags (geographical knowledge), recognizing famous faces of historical figures (historical knowledge), or classifying plants and animals (biological knowledge). These could be juxtaposed with multiple-choice declarative knowledge tests (in geography, history, and biology) and even stronger language-based gc tasks still assessing domain-specific knowledge such as vocabulary or cloze tests. These different groups of indicators can be arranged along a continuum that allows for the examination of the influence of language on the measurement of gc more systematically. The assumption is that visual and language-based declarative knowledge tasks converge strongly, which would be consistent with the view that gc is essentially knowledge, with language serving only as an efficient way to acquire and to assess knowledge efficiently.

In a large-scale study of 6,701 adolescents, Schipolowski et al. (2014) empirically examined whether the differing conceptions of gc—as knowledge versus language-related abilities—ultimately capture the same construct. In a correlated factor model, a broad verbal ability factor (comprising reading, listening, writing, orthography, language usage, and cloze tests) and a knowledge factor (covering content from the natural sciences, humanities, and social sciences) were found to correlate strongly (r = .91), but not perfectly. Approximately 17% of the variance in the knowledge factor remained independent of both verbal ability and fluid intelligence, while about 12% of the variance in verbal ability was unique. In this study, vocabulary tests—arguably the most common indicators of gc (Carroll, 1993)—were also administered and showed strong correlations with both verbal ability (r = .96) and declarative knowledge (r = .91), which highlight the role of the semantic lexicon as a central nexus between culture, language, and knowledge.

In summary, the lack of definitional clarity surrounding gc reflects both its inherent complexity and the persistent conflation of the construct with its measurement. In this chapter, it is argued that distinctions among facets of gc discussed in the literature—vocabulary, semantic knowledge, and declarative knowledge—originate more out of practical convenience of its measurement than from divergent theoretical positions. Thus, we will adopt Horn's (2008, p. 197) definition of gc as "breadth and depth of knowledge of the language, concepts, and information of the dominant culture." Accordingly, language comprehension is not synonymous with gc, but language is the primary medium through which knowledge is acquired, transmitted, and applied, and both concepts are closely related. As will be shown later, recent advances in large language models have brought about a striking convergence of declarative knowledge and semantic processing (see also Section 6).

2 Structure and Organization of gc

In his extensive scholarly work, Cattell described cognitive abilities from various perspectives, often employing an artful and metaphorical style. In his classic work *Abilities: Their Structure, Growth, and Action* (Cattell, 1987, p. 187), he emphasized the developmental aspect of measurement, writing that “crystallized ability begins after school to extend into Protean forms and that no single investment such as playing bridge or skill in dentistry can be used as a manifestation by which to test all people.” Cattell’s use of the term “Protean forms” is both intriguing and illuminating: Proteus, a sea god in Greek mythology, possessed the gift of prophetic knowledge but was unwilling to share it. Whenever questioned, he changed his form to evade capture. Similarly, the question of the dimensionality of knowledge seems to resist clear scientific resolution. An empirical examination of the "shape-shifting" construct gc appears fundamentally problematic, given that knowledge is immensely diverse, rapidly expanding across countless domains, and increasingly fragmented into specialized subdomains. Building on Cattell’s (1987) suggestion to sample knowledge as broadly as possible, Steger et al. (2019) administered more than 4,000 knowledge questions across 34 subject domains (e.g., chemistry, the arts, and politics) to over 1,000 participants in a smartphone-based study. Using bass-ackwards analysis—a procedure that involves estimating a series of orthogonally rotated principal components analyses (PCAs) on domain-level data—they explored the structural organization of knowledge. Increasing the number of components revealed the unfolding of a hierarchical structure, with the first component accounting for only 42% of the total variance, whereas a five-component solution—covering knowledge in the humanities, social studies, life sciences, behavioral sciences, and natural sciences—explained 67% of the variance.

Understanding the dimensionality of gc is important, because different levels of the knowledge hierarchy have been shown to contribute differently to the prediction of external criteria. In one study, Schroeders et al. (2021) attempted to predict participants’ age based on their responses at different levels of knowledge—ranging from an overarching gc at the apex to broad knowledge areas, specific domains, and nuances. Using elastic net regression with nested resampling, they found that the overarching gc factor and age were virtually unrelated in the test samples. Broad knowledge areas accounted for 24% of the variance in age, and specific domains explained 28%. Most notably, however, item-level data predicted 54% of the variance in age, highlighting the substantial informational value carried by individual items. Of course, this hierarchical view of knowledge is a simplification. First, the proposed hierarchy levels are not clear-cut. For example, within the domain of chemistry, one can easily include mid-level layers below the domain level such as subdomains (e.g., organic, inorganic, and physical chemistry) or topics/concepts (e.g., polymers, aromatic compounds, or amines). Second, nearly every item spans multiple content areas, even across domains that appear only loosely related, such as, for example, mathematics and art history if you consider the question: “What role did the golden ratio play in the composition of Renaissance paintings?” These examples highlight the ambiguity of assigning items to a specific hierarchical level or to a single domain, given their cross-domain nature. Moreover, the results demonstrate that declarative knowledge items have substantial and meaningful item-specific variance and cannot be treated as interchangeable.

In a noteworthy and original study, Savi et al. (2019) conceptualized intelligence as an individual’s network of interrelated cognitive skills or pieces of knowledge. To model the probabilities with which skills are acquired and to account for their interrelatedness, the authors adopted the Fortuin-Kasteleyn model, originally developed in statistical physics to study phase transitions in spin systems. Using a strictly formalized and static approach, they successfully simulated well-established phenomena from intelligence research, such as the positive manifold and the hierarchical structure of abilities. In addition, they extended their work by applying idiographic network models to describe interindividual phenomena such as the rich-get-richer effect, thereby bridging intraindividual dynamics with interindividual outcomes. Their framework opens promising avenues for future developments, including the integration of retrieval dynamics, transfer of learning, forgetting processes, and motivational factors into the model. In a related line of work, Hills (2025) built evolving network models based on findings from experimental psychology. He applied basic learning principles—most notably, the Rescorla-Wagner prediction error framework—to simulate how learning updates associative strengths between concepts. Over time, this process gradually shapes a semantic network that becomes increasingly structured, dense, and competitive. Based on these model assumptions, Hills attributed the decline in gc after age 65 to enrichment processes in the mental lexicon, whereby the accumulation of associations leads to increased interference rather than neural degradation or a genuine loss. Yet some evidence suggests that this explanation may still be incomplete. For example, Tucker-Drob et al. (2022), demonstrated that changes in fluid and crystallized abilities are strongly interdependent: individuals who decline more steeply in fluid abilities also tend to show smaller gains—or even losses—in crystallized abilities.

Nonetheless, both simulation studies underscore the value of formal modeling for advancing the study of psychological phenomena (see also van Rooij & Baggio, 2021). Conceptualizing gc as evolving cognitive networks rather than static structures shifts the focus from merely measuring cognitive outcomes to understanding the underlying processes, which may clarify how we acquire, integrate, and reorganize knowledge over time. From this perspective, gc is not merely a storehouse of facts but among other influences, the cumulative and dynamic outcome of learning, reflecting the continual restructuring of semantic networks across the lifespan. Interestingly, this conceptualization of gc finds support from a very different line of work, namely neuroanatomical research. Genç et al. (2019) examined three types of neuroimaging measures to disentangle the biological foundations of fluid and crystallized abilities. The structural network efficiency (NETstruc), which reflects the connectivity of the white-matter (long fibers that connect) was most predictive of gc, whereas cortical gray matter volume and functional network efficiency (NETfunc), capturing the dynamic coordination between brain regions, were significantly associated with gf. These findings highlight a neuroanatomical dissociation between knowledge storage/retrieval (gc/glr) and information processing abilities (gf).

3 Development and Trajectories of gc

Age-related trajectories of gc depend heavily on the research design: Cross-sectional studies are prone to cohort effects, often indicating steeper rates of aging-related cognitive declines due to advantages in education and health among younger participants. In contrast, longitudinal studies may suffer from practice effects and selective attrition, often underestimating rates of cognitive decline over time. Studies employing more sophisticated designs[1] or applying corrections for these biases, tend to converge on the general pattern that many gc indicators increase steadily until approximately age 65, with declines occurring only in later adulthood (e.g., Salthouse, 2019; Tucker-Drob, 2019). Yet, these general accounts of age trajectories for gc may be misleading. On one hand, different types of knowledge—such as domain-specific or occupational knowledge—may follow distinct developmental courses depending on the learning environment (see Ackerman, 1996). On the other hand, the development of gc is highly idiosyncratic, shaped by a wide range of individual, contextual, and situational factors. A further important caveat: analyses of age-related trajectories presuppose measurement invariance of the indicators across age, a condition that is frequently assumed but rarely tested. Given the increasing differentiation and specialization of knowledge over the lifespan (e.g., during job training), it seems plausible that the factorial structure of gc would also evolve. However, empirical evidence suggested otherwise. Watrin et al. (2022), for example, tested this assumption using data from adults aged 18-70 who completed a broad declarative knowledge test and found no evidence of age-related differentiation when applying local structural equation

[1] As a sidenote, already Cattell (1987, p. 189) proposed addressing this issue by distinguishing between a "normative curve", representing typical age-related changes under common biological and cultural conditions across historical periods, and an "epogenic curve", which reflects "the special part due to the particular historical conditions in that epoch."

modeling. Similarly, a meta-analysis by Breit et al. (2022) concluded that evidence for age-related differentiation in gc is relatively sparse and inconsistent, and largely based on cross-sectional data, in which age and cohort effects are confounded.

Ackerman's PPIK theory (1996, 2000), which is mainly a theory on adult gc, posits that knowledge acquisition is shaped by a complex interplay of cognitive resources, motivational-affective dispositions, and environmental opportunities. According to this framework, gc emerges not only through the investment of gf and formal education but also through person-specific experiences. Some of these experiences are normative and culture-specific, such as those gained through the educational system (labeled as traditional gc). Others arise from occupation-specific environments, such as job-related knowledge and professional skills. Still others are highly idiosyncratic, reflecting individual engagement in specific domains (e.g., avocational pursuits like chess or astronomy). As people move into adulthood, their knowledge trajectories tend to diverge—even among those with similar cognitive baselines in early adulthood—because the learning environments become more heterogeneous. In other words, adult knowledge acquisition is increasingly shaped by stable personal interests, personality traits, and self-selected learning conditions.

The question posed by Ackerman's PPIK theory—namely, the origins of our knowledge—is highly relevant for understanding gc but has been rarely investigated in psychological research outside standardized laboratory settings. This stands in stark contrast to educational research, which is primarily concerned with questions of how and what we learn. La Belle (1982) distinguished among three modes of education: Formal education refers to the institutionalized, graded, and hierarchically structured system extending from primary school to higher education, whereas nonformal education encompasses organized learning activities outside formal institutions, often targeting specific groups (e.g., visits to the natural history museum). Finally, informal education involves the unstructured, lifelong acquisition of knowledge through everyday experiences (e.g., parental instruction). Motivational factors, perceived autonomy in learning, and the emotional salience of the content may all influence the development of gc, yet this line of research has received relatively little attention across the breadth of knowledge domains. For example, does it matter whether one reads about urban development in Paris, watches a documentary on the French capital, or visits the Eiffel Tower in person? Intuitively, one might assume so. However, one could also argue that gc is polygenetic, such that the source or modality of information may play a less decisive role in its acquisition—an idea that warrants further investigation. There is evidence that personal experience plays an important role in the encoding of certain autobiographical knowledge (not declarative knowledge), particularly when the information is emotionally charged or tied to dramatic events—such as the 9/11 attacks, the assassination of John F. Kennedy, or the Vietnam War (e.g., Rubin et al., 1998). There is also substantial research on the reminiscence bump, the phenomenon that people tend to recall memories from adolescence and early adulthood more readily and frequently, especially when personally significant experiences are involved, such as their favorite music or popular songs (e.g., Zimprich, 2020)[2]. However, the more detached or abstract the knowledge, the less pronounced this temporal effect becomes (e.g., Achaa-Amankwaa et al., 2024, for public-event knowledge). In summary, this section has reviewed key empirical findings on the development of adult gc and pointed to questions that remain unresolved. Any comprehensive theory of adult gc must address these findings, including age-related trajectories and the diverse modes of knowledge acquisition.

4 Determinants and Consequences of gc

Gc is shaped by the interplay of individual, contextual, and developmental influences. The following section highlights key determinants and consequences of gc, including cognitive factors (gf, prior gc), motivational dispositions (intellectual investment traits such as open-

[2]This is not to suggest, however, that gc is independent of age. As mentioned above, age effects at the item level are particularly pronounced, which is why such items are now even used as age-verification tasks in online surveys (Hartman et al., 2022).

ness/intellect), and contextual conditions (education and broader learning opportunities). In many models of interindividual differences, gc is treated as a determinant. However, this idea is debatable, as gc can often be understood both as a cause and a consequence in relation to other constructs. For example, education has been conceptualized both as a predictor and as an outcome of gc. To avoid the charge of arbitrariness, it is therefore crucial to clearly define what is meant by broad constructs. When education is understood as access to learning opportunities in an ability-stratified environment, including the exposure to formal instruction, it can be viewed as a predictor. Conversely, when defined as the highest degree or credential earned in the sense of educational attainment, it may be seen as a consequence. Either way, based on definition alone, a strong association between education and gc is to be expected. Accordingly, numerous studies have demonstrated robust associations between years of formal education and performance in gc tests (Ritchie & Tucker-Drob, 2018). Schooling not only provides exposure to culturally valued information, but it also cultivates domain-general learning strategies (e.g., reading skills or meta-cognitive skills) that facilitate its acquisition beyond the classroom.

According to Cattell's investment theory, gf facilitates the acquisition of gc by enabling individuals to learn and encode new information. However, empirical support for the investment hypothesis is mixed. For instance, Ferrer and McArdle (2004) found no evidence of a coupling effect between gf and gc in a longitudinal sample spanning childhood to early adulthood. In contrast, the same authors later reported strong coupling effects between cognitive abilities and academic achievement, particularly during early school years (Ferrer & McArdle, 2004). Extending this line of research, Kievit et al. (2017) proposed a mutualistic coupling account, in which gc and gf reinforce each other over time—challenging the unidirectional assumption of the traditional investment model. Moreover, the relationship between gf and gc may not remain developmentally stable. In older adulthood, the age-related decline in gf often weakens this association, as gc tends to remain relatively stable (Tucker-Drob, 2019). Hills (2025), however, argued in his work on evolving cognitive networks that these age-related changes are not independent but share a common origin—cognitive network enrichment resulting from lifelong learning.

Since gc is often conceptualized as accumulated knowledge, the best predictor of future gc is prior gc. Individuals with higher initial levels of knowledge are more likely to acquire new knowledge over time—a phenomenon commonly referred to as rich-get-richer effect or Matthew effect[3]. The knowledge-is-power hypothesis (Hambrick & Engle, 2002; Witherby & Carpenter, 2022) even further posits that prior knowledge is even more important than reasoning ability within a specific domain. In some areas of expertise, such as basketball or chess, this hypothesis may hold: Experts process new information more quickly, integrate it more deeply, and retrieve it more accurately (Hambrick, 2004). However, for broadly defined gc, such generalizations are more difficult to sustain. In a meta-analysis of 493 studies examining the relationship between prior knowledge and knowledge gains, the overall effect was close to zero, but the variability in effect sizes was substantial (Simonsmeier et al., 2022). This large heterogeneity in effects underscores the need for systematic research to identify the conditions under which the rich-get-richer effects emerge (Brod, 2021). One such condition may be the learning environment. In structured educational settings such as schools, instruction often targets the needs of lower-performing students. Such compensatory practices can reduce rather than amplify initial performance differences, thereby mitigating the rich-get-richer effect (e.g., Schroeders, Schipolowski, et al., 2016). More broadly, since learning opportunities—shaped by cultural, socioeconomic, and geographic factors—play a critical role, gc should be understood as embedded within an ecological system of opportunity structures that either enable or constrain its development.

After leaving formal schooling, personality and motivational factors gain importance and show consistent associations with gc (Ack-

[3]The name refers to the bible verse: "To all those who have, more will be given, and they will have an abundance; but from those who have nothing, even what they have will be taken away" (Matthew 13:12 and also 25:29, The New Oxford Annotated Bible, New Revised Standard Version).

erman, 1996; Ackerman & Heggestad, 1997). Among them, the two facets of the Big Five trait openness/intellect show differential associations: While openness to fantasy, aesthetics, or feelings have low or near-to-zero correlations with gc, the intellect facet exhibits substantial positive correlations (von Stumm & Ackerman, 2013). Because openness is classified within the same seek-think cluster as Typical Intellectual Engagement and Need for Cognition in the Intellect Framework (Mussel, 2013)—with seek reflecting epistemic motivation and think representing reflective processing—similar associations can be found for these intellectual investment traits. One proposed explanation is that individuals high in intellectual openness are more inclined to explore novel situations, increasing their exposure to new information that can be integrated into existing knowledge structures. Openness is therefore thought to promote self-directed learning and the active search for learning opportunities (Ackerman, 1996). An open question is whether openness/intellect primarily determines the amount of energy invested in intellectual activities—akin to a steam boiler model, with interests directing that energy—or whether it is inherently tied to specific content domains (see also Watrin et al., 2022).

Ackerman (2017) highlighted an important issue in the discussion of antecedents and consequences, which he termed the criterion problem. Rather than being treated primarily as predictors, intelligence tests have increasingly come to be used as criteria. In his review of more than a century of intelligence research, he argued that most adult intelligence tests were derived from instruments originally developed for children and adolescents. Moreover, their original measurement aim was to predict academic achievement, which has led to a shift whereby adult intelligence tests have turned "from imperfect predictor measures to idealized criterion measures of intelligence" (Ackerman, 2017, p. 994). According to Ackerman, our understanding of adult intellect remains still limited, and he underscores the need for behaviorally based validation of measures. For example, Ackerman pointed out that (specialized) domain knowledge, despite its predictive value for occupational success, is largely absent from adult intelligence tests.

5 Measurement and Psychometric Modeling of gc

Gc is often studied from an individual psychological perspective, with situational and contextual influences typically framed at a specific point in time and with reference to populations assumed to have had essentially equal opportunities to acquire the relevant knowledge.[4] Nowhere are the "Protean qualities" of gc more apparent than in its dependence on culture and historical epoch. Already Cattell (1987, p.149) argued that gc has a "precarious existence", noting that "cultural change, shift of mixture of areas intellectually fashionable, or a change in the school curriculum can weaken its identity and unity as a discernible factor." Compared to the extensive body of cross-cultural research in other domains (e.g., personality research), there is surprisingly little empirical work on the comparability of cognitive abilities across cultures, especially gc. This scarcity can be attributed to several factors, including the difficulties of such group comparisons in general (see Wicherts & Dolan, 2010), the considerable logistical and personnel demands of (proctored) ability testing, and the inherent difficulty of developing gc instruments that are truly comparable across countries. For example, Watrin et al. (2023) developed a 75-item knowledge test administered in the native languages of participants from Germany, France, and the United States. Each of the 15 knowledge domains included both country-specific items (e.g., "Where in Germany is Dresden located?" or "What is the longest river in France?") and general, cross-national items (e.g., "What is amber made of?"). Not surprisingly, results showed that participants scored significantly higher on items specific to their country of residence, particularly in the social sciences and humanities, while no consistent cross-national differences emerged for items from the natural sciences. Interestingly, proxies for affinity for international exchange (e.g., travel experience, friendships, language

[4] Yet the very fact that you are reading this book chapter suggests that you belong—globally and historically speaking—to a relatively privileged group with access to gatekept knowledge. How likely would it be that you engaged with the kinds of topics that ultimately led you to this chapter if you had grown up in a rural region in a non-WEIRD country, with limited access to formal education, libraries, or digital media?

skills, cultural interests) were unrelated to performance on country-specific knowledge, suggesting that mere exposure to foreign contexts does not necessarily translate into culturally embedded knowledge.

It has repeatedly been shown that culture within a single country is not static but subject to social change (Henrich, 2020). For example, what appeared to be a negative Flynn effect in France was largely driven by gc subtests with high cultural load, where shifts in item functioning created measurement artifacts rather than reflecting true declines in crystallized ability (Gonthier et al., 2021). Any test that functions differently across countries, despite being designed to measure the same construct, indicates item or instrument bias according to van de Vijver and Tanzer (2004). Such forms of bias undermine measurement equivalence and, consequently, threaten the validity of cross-national comparisons. However, cultural, contextual, and temporal dependencies in gc measures should not be treated as methodological shortcomings, but rather as inherent features of the construct itself. The dependence on a given context reflects genuine differences in what is considered relevant knowledge within a society. According to communication theorist James Carey (2009), communication as cultural transmission is not primarily aimed at conveying information, but at expressing shared beliefs and thereby maintaining society over time. If one accepts that gc is inevitably bound to societal norms and values at a given point in time, the suitability of a cognitive test as an indicator of gc may even be evaluated by locating it along dimensions of language and cultural load (for a similar idea on the organization of subtests in a test battery, see Ortiz et al., 2018).

Considerations for the best measurement of gc are closely linked to the type of psychometric modeling approach applied. Using any statistical method, researchers implicitly adopt psychometric assumptions that strongly influence the conclusions drawn. For example, factor analysis, which is often used in the study of interindividual differences, relies on a reflective model that is arguably the most genuine model in psychology. It assumes that a latent variable, which cannot be directly observed, serves as a common cause underlying individuals' responses to observed indicators (Markus & Borsboom, 2013). This modeling perspective posits that the latent ability fully explains the correlations among individual knowledge items, whereas their unique variance reflects random error. In contrast, in a formative model, the latent construct is not the cause but rather the result of the observed indicators—it emerges from the combination of measured variables, which are seen as defining rather than reflecting the construct (for a more detailed discussion of causal-formative vs. composite-formative models, see Bollen & Diamantopoulos, 2017). According to this view, gc is fully represented by its indicators, each assigned a specific weight. Unfortunately, both reflective and the formative modeling approaches rely on assumptions that are not totally justified, such as the assumption of local stochastic independence (contradicting the reflective model) or the assumption of error-free measurement (contradicting the composite–formative model). Although such violations do not necessarily undermine conclusions from a pragmatic perspective, they show that gc can only be imperfectly represented in these statistical frameworks. Particularly problematic are the high item uniqueness and the strong influence of item sampling on test results are problematic. Beyond the age dependence of the indicators noted above, other moderators also affect gc at the item level. For example, Schroeders, Wilhelm, and Olaru (2016) used metaheuristic optimization to compile unidimensional short scales from a larger pool of knowledge items in order to maximize gender differences in favor of either women or men. Although the resulting scales all assessed the same construct at an aggregated level (e.g., knowledge in natural sciences), the observed gender differences could be shifted in virtually any direction. This underscores once again the critical role of item sampling in gc tests, which complicates both the comparability and generalizability of person scores.

These traditional psychometric approaches operationalize gc as a uni- or multidimensional latent trait. While such models are common in intelligence research, they disregard the relational and topological organization of knowledge. An alternative approach to modeling knowledge structures involves the use of psychometric network analysis (e.g., Hills, 2025; Savi et al., 2019), which conceptualizes knowledge as a system of interconnected concepts represented as nodes and their associations as edges (van der Maas et al., 2017). These models assume that gc is defined both by the nodes themselves and by the interwoven web of meaningful associations between them, much like a semantic network in which concepts are connected by semantic links. To capture this structural complexity more directly, graph embedding methods map nodes (e.g., concepts, knowledge items) and their relations (e.g., semantic similarity, co-occurrence, prerequisite structures) into a continuous vector space, where distances and directions have interpretable meaning (e.g., Grover & Leskovec, 2016). Within such spaces, clusters may correspond to knowledge domains, distances can represent semantic proximity, and trajectories may capture the developmental unfolding of knowledge structures over time. This makes it possible to simulate learning as the addition or strengthening of edges and forgetting as the weakening or removal of connections. From this perspective, the accessibility of a concept depends on the configuration of its connections within the network. Highly interconnected nodes may function as hubs that facilitate the retrieval and integration of related information, whereas sparsely connected nodes may be less readily accessible. Finally, network models allow researchers to quantify structural properties such as clustering, centrality, and path length, which can provide insights into the organization of gc and the efficiency of its long-term storage and retrieval processes (commonly referred to as glr, Goecke et al., 2025).

6 Large-Language Models and gc

Just as psychometric network models suggest that gc emerges from the interplay of interconnected pieces of information within domains, large language models (LLMs) show how complex abilities can arise from networks of relatively simple computational elements. LLMs have rapidly transformed education, communication, and research, but they also have a far-reaching impact on the research of cognitive abilities and provide an opportunity to reconsider gc both as a theoretical construct and as a measurable ability. LLMs process massive amounts of text, extracting statistical patterns and making predictions about which words, phrases, and structures most likely co-occur. In this sense, large-scale text analysis enables them not only to process language, but also to generate outputs that resemble factual knowledge acquisition in humans, although these systems differ in crucial ways (Lake et al., 2017). Nonetheless, the convergence of language processing and knowledge representation in LLMs is remarkable and parallels the dual conceptualization of gc as acquired knowledge (Cattell) and language comprehension (Carroll), two perspectives that may appear distinct but are deeply intertwined. Traditionally, vocabulary, reading comprehension, and factual knowledge have been treated as separate but strongly correlated facets of gc, but the functioning of LLMs demonstrates how these components are not only empirically related but also computationally inseparable. Language structures provide the scaffolding for knowledge representation, while knowledge provides the semantic depth that gives meaning to language. In other words, the statistical learning of linguistic regularities and the accumulation of world knowledge are two sides of the same coin. In future research, a more detailed exploration of this relationship from an assessment and a psychometric perspective constitutes a promising and stimulating area of investigation. Two ideas for future research are sketched in the following.

Gc was introduced as the breadth and depth of one's knowledge, with an emphasis on covering a wide range of culturally shared and valued content areas. Since LLMs map knowledge-related content in a high-dimensional vector space, it is, in principle, possible to measure gc by examining whether central information, ideas, and concepts are encoded in this space. From this perspective, a person's gc might be approximated as a region or distribution within a semantic space. In LLMs, embeddings—numerical vector representations of words, sentences, or concepts—constitute the central representational units and capture semantic meaning and relationships in a high-dimensional space. For example, in a BERT model (Bidirectional Encoder Representations from Transformers, Devlin et al., 2019), each word or token is encoded as a numeric vector in a high-dimensional space (typically with 768 or 1024 dimensions in BERT-based models). Instead of reducing a person's ability estimate to a single value, assessing the coverage of central regions in the n-dimensional space could convey a more realistic picture of ability. In this sense, gc—defined as the knowledge that is broadly shared, culturally valued, and commonly encountered within a society—could be measured through the overlapping regions of knowledge that are shared across many individuals. However, this perspective should not be regarded as static. Consistent with the idea of gc as an evolving network, learning could then be modeled as trajectories through this space, with time serving as an additional dimension that captures the reorganization of semantic networks.

Second, the assessment of gc can be rendered more efficient through the use of LLMs and their embeddings in rational test construction. For decontextualized abilities, such as gf test items, predictions of item difficulty using item design features work well, so that this information can be used for item construction and assembly without extensive pretesting (e.g., computerized adaptive testing). For example, based on construction rules (e.g., number of elements, type of rules implemented, amount of visual complexity), 60–80% of the variance in item difficulty for matrix problems can regularly be explained (Primi, 2014). For contextualized abilities, however, the situation is far more complex. Here, LLMs may offer a promising opportunity, serving as a logical extension of earlier corpus-based approaches. For example, Schroeders and Achaa-Amankwaa (2025) constructed a vocabulary test for adult native speakers of German and demonstrated that two linguistic features—word frequency and word length—accounted for 61% of the variance. However, the prediction accuracy could be further improved up to 80% when embeddings reduced through Principal Component Analysis are used. It must be noted that these results are preliminary and based on a comparatively small item set ($n = 110$). Similar considerations regarding the prediction of item difficulty prior to testing and rational item construction with the aid of LLMs can also be applied to other gc indicators, such as cloze tests or classical declarative knowledge tasks.

7 Conclusion

Hopefully, this chapter has conveyed that, despite—or precisely because of—the many open questions, gc remains a fascinating object of inquiry. The appeal is to engage more deeply with gc, because unlike other domains of cognitive abilities, the study of gc offers the opportunity to generate insights that are also relevant beyond the boundaries of classical intelligence research. Gc is so multifaceted that it should not be examined from the vantage point of a single discipline, but rather through interdisciplinary collaboration. The perspective presented here had a strong focus on assessment and psychometric modeling, but it represents only one way of approaching gc; other perspectives from cognitive psychology, education, communication studies, and computer science are equally essential to a more comprehensive understanding.

Across the subsections, three central propositions have been put forward and elaborated in detail. First, gc corresponds to declarative knowledge, with spoken and written language serving as its primary, though not exclusive, sign system. In mathematical terms: gc = knowledge $\gtrsim$ language. This conception of gc helps to resolve a long-standing debate over whether gc should be understood primarily as knowledge in Cattell's sense or as language comprehension in Carroll's sense. Moreover, the distinction between gc and gkn (= general domain-specific knowledge) is gradual rather than categorical, with differences largely reflecting the extent to which knowledge is broadly shared within society and accorded cultural value.

Second, neither reflective nor formative models are suitable for capturing high item uniqueness, within-item dimensionality, or the aggregation of knowledge without informational loss. Short- and long-term developments of gc could instead be conceptualized as evolving (cognitive) networks. In these models, gc may function as both cause and consequence. Simulations of intraindividual mechanisms that are eligible to explain interindividual phenomena provide initial supporting evidence for such a conceptualization (Savi et al., 2019). Taken together, this network-based perspective opens promising avenues for modeling the dynamic nature of gc.

Finally, in contrast to reasoning abilities, the scientific investigation of knowledge remains in its infancy, which is why Ackerman (2000) coined the phrase 'dark matter' of adult intelligence to describe gc. With their capacity to represent concepts and relationships in a semantic space, LLMs could be leveraged more systematically in future research to gain deeper insights into knowledge structures and performance. Building on this, some ideas have been outlined for developing new measurement instruments of gc, for improving existing ones (e.g., through rational item construction), and for capturing a person's ability in a more comprehensive manner.

Acknowledgement

The author is grateful to Priscilla Achaa-Amankwaa, Gilles Gignac, Kristin Jankowsky, Tim Trautwein, and Jonas Walter for their helpful comments and valuable feedback on earlier versions of this book chapter.

Declaration of Generative AI and AI-Assisted Technologies in the Writing Process.

During the preparation of this chapter, the author used ChatGPT-5 to assist with translation and to improve the readability of some passages. The author reviewed and edited all content as necessary and takes full responsibility for the final text.

References

Achaa-Amankwaa, P., Steger, D., Wilhelm, O., & Schroeders, U. (2024). Public events knowledge in an age-heterogeneous sample: Reminiscence bump or bummer? *Psychology and Aging*, *39*(1), 72–87. https://doi.org/10.1037/pag0000786

Ackerman, P. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227–257. https://doi.org/10.1016/S0160-2896(96)90016-1

Ackerman, P. (2000). Domain-specific knowledge as the "dark matter" of adult intelligence. *The Journals of Gerontology Series B: Psychological Sciences and Social Sciences*, *55*(2), 69–84. https://doi.org/10.1093/geronb/55.2.P69

Ackerman, P. (2017). Adult intelligence: The construct and the criterion problem. *Perspectives on Psychological Science*, *12*(6), 987–998. https://doi.org/10.1177/1745691617703437

Ackerman, P., & Heggestad, E. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin*, *121*, 219–245. https://doi.org/10.1177/1069072703011002006

Binet, A., & Simon, T. (1904). Méthodes nouvelles pour le diagnostic du niveau intellectuel des anormaux [new methods for the diagnosis of the intellectual level of subnormals. *L'année Psychologique*, *11*, 191–244. https://doi.org/10.3406/psy.1904.3675

Bollen, K., & Diamantopoulos, A. (2017). In defense of causal-formative indicators: A minority report. *Psychological Methods*, *22*(3), 581–596. https://doi.org/10.1037/met0000056

Breit, M., Brunner, M., Molenaar, D., & Preckel, F. (2022). Differentiation hypotheses of intelligence: A systematic review of the empirical evidence and an agenda for future research. *Psychological Bulletin*, *148*(7–8), 518–554. https://doi.org/10.1037/bul0000379

Brod, G. (2021). Toward an understanding of when prior knowledge helps or hinders learning. *Npj Science of Learning*, *6*(1), 24. https://doi.org/10.1038/s41539-021-00103-w

Carey, J. (2009). *Communication as culture: Essays on media and society*. Routledge. https://doi.org/10.4324/9780203928912

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

Cattell, R. (1987). *Intelligence: Its structure, growth and action*. North-Holland.

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technolo-*

gies, *1*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Ferrer, E., & McArdle, J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, *40*(6), 935–952. https://doi.org/10.1037/0012-1649.40.6.935

Flanagan, D., & Dixon, S. (2014). Cattell-horn-carroll theory of cognitive abilities. In C. Reynolds, K. Vannest, & E. Fletcher-Janzen (Eds.), *Encyclopedia of special education* (4th). John Wiley & Sons. https://doi.org/10.1002/9781118660584.ese0431

Genç, E., Fraenz, C., Schlüter, C., Friedrich, P., Voelkle, M., Hossiep, R., & Güntürkün, O. (2019). The neural architecture of general knowledge. *European Journal of Personality*, *33*(5), 589–605. https://doi.org/10.1002/per.2217

Goecke, B., Weiss, S., & Wilhelm, O. (2025). Driving factors of individual differences in broad retrieval ability: Gr is more than the sum of its parts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *51*(2), 301–319. https://doi.org/10.1037/xlm0001336

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 [84, Article 101512.]. *Intelligence*. https://doi.org/10.1016/j.intell.2020.101512

Grover, A., & Leskovec, J. (2016). Node2vec: Scalable feature learning for networks [No. arXiv:1607.00653). arXiv.]. https://doi.org/10.48550/arXiv.1607.00653

Hambrick, D. (2004). The role of domain knowledge in higher-level cognition. In O. Wilhelm & R. Engle (Eds.), *Handbook of understanding and measuring intelligence* (pp. 361–372). Sage Publications.

Hambrick, D., & Engle, R. (2002). Effects of domain knowledge, working memory capacity, and age on cognitive performance: An investigation of the knowledge-is-power hypothesis. *Cognitive Psychology*, *44*, 339–387. https://doi.org/10.1006/cogp.2001.0769

Hartman, R., Moss, A., Rabinowitz, I., Bahn, N., Rosenzweig, C., Robinson, J., & Litman, L. (2022). Do you know the wooly bully? testing era-based knowledge to verify participant age online. *Behavior Research Methods*, *55*(7), 3313–3325. https://doi.org/10.3758/s13428-022-01944-y

Henrich, J. (2020). *The weirdest people in the world: How the west became psychologically peculiar and particularly prosperous*. Farrar, Straus; Giroux.

Hills, T. (2025). Cognitive network enrichment, not degradation, explains the aging mental lexicon and links fluid and crystallized intelligence [Advance online publication.]. *Psychological Review*. https://doi.org/10.1037/rev0000557

Horn, J. (2008). Spearman, g, expertise, and the nature of human cognitive capability. In P. Kyllonen, R. Roberts, & L. Stankov (Eds.), *Extending intelligence: Enhancement and new constructs* (pp. 185–230). LEA.

Kan, K.-J., Kievit, R., Dolan, C., & Maas, H. (2011). On the interpretation of the chc factor gc. *Intelligence*, *39*, 292–302. https://doi.org/10.1016/j.intell.2011.05.003

Kievit, R., Lindenberger, U., Goodyer, I., Jones, P., Fonagy, P., Bullmore, E., Network, P., & Dolan, R. (2017). Mutualistic coupling between vocabulary and reasoning supports cognitive development during late adolescence and early adulthood. *Psychological Science*, *28*(10), 1419–1431. https://doi.org/10.1177/0956797617710785

La Belle, T. (1982). Formal, nonformal and informal education: A holistic perspective on lifelong learning. *International Review of Education*, *28*(2), 159–175. https://doi.org/10.1007/BF00598444

Lake, B., Ullman, T., Tenenbaum, J., & Gershman, S. (2017). Building machines that learn and think like people. *Behavioral and Brain Sciences*, *40*, 253. https://doi.org/10.1017/S0140525X16001837

Markus, K., & Borsboom, D. (2013). Reflective measurement models, behavior domains, and common causes. *New Ideas in Psy-*

chology, *31*(1), 54–64. https://doi.org/10.1016/j.newideapsych.2011.02.008
McGrew, K. (2009). Chc theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*(1), 1–10. https://doi.org/10.1016/j.intell.2008.08.004
McGrew, K., Schneider, W., Decker, S., & Bulut, O. (2023). A psychometric network analysis of chc intelligence measures: Implications for research, theory, and interpretation of broad chc scores "beyond g. *Journal of Intelligence*, *11*(1), 19. https://doi.org/10.3390/jintelligence11010019
Mussel, P. (2013). Intellect: A theoretical framework for personality traits related to intellectual achievements. *Journal of Personality and Social Psychology*, *104*(5), 885–906. https://doi.org/10.1037/a0031918
Ortiz, S., Piazza, N., Ochoa, S., & Dynda, A. (2018). Testing with culturally and linguistically diverse populations: Moving beyond the verbal–performance dichotomy into evidence-based practice. In D. Flanagan & E. McDonough (Eds.), *Contemporary intellectual assessment* (4th, pp. 684–712). Guilford Press.
Primi, R. (2014). Developing a fluid intelligence scale through a combination of rasch modeling and cognitive psychology. *Psychological Assessment*, *26*(3), 774–788. https://doi.org/10.1037/a0036712
Ritchie, S., & Tucker-Drob, E. (2018). How much does education improve intelligence? a meta-analysis. *Psychological Science*, *29*(8), 1358–1369. https://doi.org/10.1177/0956797618774253
Rubin, D., Rahhal, T., & Poon, L. (1998). Things learned in early adulthood are remembered best. *Memory & Cognition*, *26*(1), 3–19. https://doi.org/10.3758/BF03211366
Salthouse, T. (2019). Trajectories of normal cognitive aging. *Psychology and Aging*, *34*(1), 17–24. https://doi.org/10.1037/pag0000288
Savi, A., Marsman, M., Maas, H., & Maris, G. (2019). The wiring of intelligence. *Perspectives on Psychological Science*, *14*, 1034–1061. https://doi.org/10.1177/1745691619866447
Schipolowski, S., Wilhelm, O., & Schroeders, U. (2014). On the nature of crystallized intelligence: The relationship between verbal ability and factual knowledge. *Intelligence*, *46*, 156–168. https://doi.org/10.1016/j.intell.2014.05.014
Schroeders, U., & Achaa-Amankwaa, P. (2025). Developing nova: A next-generation open vocabulary assessment. *PsyArXiv*. https://doi.org/https://osf.io/preprints/psyarxiv/vhakw_v2
Schroeders, U., Schipolowski, S., Zettler, I., Golle, J., & Wilhelm, O. (2016). Do the smart get smarter? development of fluid and crystallized intelligence in 3rd grade. *Intelligence*, *59*, 84–95. https://doi.org/10.1016/j.intell.2016.08.003
Schroeders, U., Watrin, L., & Wilhelm, O. (2021). Age-related nuances in knowledge assessment [85, Article 101526.]. *Intelligence*. https://doi.org/10.1016/j.intell.2021.101526
Schroeders, U., Wilhelm, O., & Olaru, G. (2016). The influence of item sampling on sex differences in knowledge tests. *Intelligence*, *58*(3), 22–32. https://doi.org/10.1016/j.intell.2016.06.003
Simonsmeier, B., Flaig, M., Deiglmayr, A., Schalk, L., & Schneider, M. (2022). Domain-specific prior knowledge and learning: A meta-analysis. *Educational Psychologist*, *57*(1), 31–54. https://doi.org/10.1080/00461520.2021.1939700
Steger, D., Schroeders, U., & Wilhelm, O. (2019). On the dimensionality of crystallized intelligence: A smartphone-based assessment. *Intelligence*, *72*, 76–85. https://doi.org/10.1016/j.intell.2018.12.002
Tucker-Drob, E. (2019). Cognitive aging and dementia: A life-span perspective. *Annual Review of Developmental Psychology*, *1*(1), 177–196. https://doi.org/10.1146/annurev-devpsych-121318-085204
Tucker-Drob, E., Fuente, J., Köhncke, Y., Brandmaier, A., Nyberg, L., & Lindenberger, U. (2022). A strong dependency between changes in fluid and crystallized abilities in human cognitive aging [Article eabj2422.]. *Science Advances*,

8(5). https://doi.org/10.1126/sciadv.abj2422

van Rooij, I., & Baggio, G. (2021). Theory before the test: How to build high-verisimilitude explanatory theories in psychological science. *Perspectives on Psychological Science*, *16*(4), 682–697. https://doi.org/10.1177/1745691620970604

van de Vijver, F., & Tanzer, N. (2004). Bias and equivalence in cross-cultural assessment: An overview. *European Review of Applied Psychology*, *54*(2), 119–135. https://doi.org/10.1016/j.erap.2003.12.004

van der Maas, H., Kan, K.-J., Marsman, M., & Stevenson, C. (2017). Network models for cognitive development and intelligence. *Journal of Intelligence*, *5*(2), 16. https://doi.org/10.3390/jintelligence5020016

von Stumm, S., & Ackerman, P. (2013). Investment and intellect: A review and meta-analysis. *Psychological Bulletin*, *139*(4), 841–869. https://doi.org/10.1037/a0030746

Watrin, L., Schroeders, U., & Wilhelm, O. (2022). Structural invariance of declarative knowledge across the adult lifespan. *Psychology and Aging*, *37*(3), 283–297. https://doi.org/10.1037/pag0000660

Watrin, L., Schroeders, U., & Wilhelm, O. (2023). Gc at its boundaries: A cross-national investigation of declarative knowledge. *Learning and Individual Differences*, *102*, 102267. https://doi.org/10.1016/j.lindif.2023.102267

Wicherts, J., & Dolan, C. (2010). Measurement invariance in confirmatory factor analysis: An illustration using iq test performance of minorities. *Educational Measurement: Issues and Practice*, *29*(3), 39–47. https://doi.org/10.1111/j.1745-3992.2010.00182.x

Witherby, A., & Carpenter, S. (2022). The rich-get-richer effect: Prior knowledge predicts new learning of domain-relevant information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *48*(4), 483–498. https://doi.org/10.1037/xlm0000996

Zimprich, D. (2020). Individual differences in the reminiscence bump of very long-term memory for popular songs in old age: A non-linear mixed model approach. *Psychology of Music*, *48*(4), 547–563. https://doi.org/10.1177/0305735618812199