Individual Differences: The History of the Abandoned Child of Experimental Psychology

Kristof Kovács

1 Introduction

Founded in 1959 and currently having more than 4300 members, Psychonomic Society's self-defined mission is "to foster the science of cognition through the advancement and communication of basic research in experimental psychology and allied sciences" (*About, Mission, & History - Psychonomic Society*, n.d.). This goal is achieved through the Society's annual meetings as well as its seven journals, which cover "all aspects of cognitive and experimental psychology" (*Journals - Psychonomic Society*, n.d.). Eyeballing keywords in the journals' description one indeed finds the standard topics of research in the field: memory, cognitive development, learning, conceptual processes, psycholinguistics, decision making, sensory processes, perception, attention, psychophysics, motivation, emotion, and social behaviour. Additionally, some of the journals publish papers on brain processes related to cognition as well as on methodological issues and applications. Even though Psychonomic Society is obviously not the only scientific society dedicated to the study of cognition, it is one of the largest, and the topics of its journals are mostly representative of the field.

Apparently, individual differences in cognition are not covered in a list even so wide as to include development and social behaviour. Not that such a field does not exist. It does, but it appears as if it inhabited a parallel universe, where cognition is called intelligence and research results are disseminated in conferences devoted solely to individual differences. In this paper I review the history of this separation and identify reasons for why it occurred. There are several accounts of the history of intelligence research and measurement (Boeck et al., 2019; Brody, 2000; Cianciolo & Sternberg, 2008; Fancher, 1985; Mackintosh, 2011b; Sternberg, 2020). My intention

K. Kovács (✉)
Institute of Psychology, ELTE Eotvos Lorand University, Budapest, Hungary
e-mail: kovacs.kristof@ppk.elte.hu

J. Gervain et al. (eds.), *A Life in Cognition*, Language, Cognition, and Mind 11,
https://doi.org/10.1007/978-3-030-66175-5_17

is not to write another such history. On the contrary, the aim of this paper is to investigate why there is a separate history of cognitive ability research in the first place, instead of being a subchapter of the history of cognitive/experimental psychology. Nor will I provide a history of individual differences research in general, including personality or emotional intelligence, because the focus of the paper is individual differences *in cognition.*

First, I review the early history of the field of intelligence and how it diverged from mainstream experimental psychology. Then I survey attempts to unify these two disciplines and highlight related methodological and conceptual issues. I argue that the gap has been rapidly closing recently, mostly due to the concept of working memory capacity. This is followed by a discussion of measurement of cognitive abilities in school settings and how that is influenced by relevant cognitive theories, or in many cases, a lack of them. Finally, I discuss William Stern's century-old program for differential psychology and its surprising relevance for contemporary difficulties and controversies. I will conclude the paper with a few personal remarks about how my own interest in these matters evolved under the strong influence of Csaba Pléh.

2 A History of Separation

Intelligence tests have a notoriously bad reputation both within and outside academia. This is remarkable for an instrument developed with the explicit purpose to provide personalized help to children with learning difficulties, especially ones with a disadvantaged background (Sternberg, 2020). Paradoxically, IQ tests are also one of psychology's most visible success stories. Cognitive ability is regularly assessed in clinical, industrial/organizational and school settings, and IQ reliably predicts a number of important outcomes such as educational achievement (O'Connell, 2018; Roth et al., 2015), job success and income (Gensowski, 2018; Schmidt & Hunter, 2004), and even health and longevity (Calvin et al., 2011; Deary et al., 2019). At the same time, since psychometrics evolved independently of mainstream cognitive/experimental psychology for most of the last century, most research related to individual differences has been atheoretical (Mackintosh, 2011a)—or had been up to relatively recently. The "psychometric tradition that followed Binet has dominated the study of intelligence, and the analysis of intelligence test scores and their relationships to other variables have been the focus of research" (Anderson, 2005, p. 270.)

2.1 The Beginnings of Individual Differences Research

The scientific study of individual differences in cognition has been separated from mainstream experimental psychology since the birth of modern psychology. Both

James McKeen Cattell, the developer of the first mental test and Charles Spearman, the inventor of factor analysis and the discoverer of the general factor of intelligence were students of Wilhelm Wundt. At the same time, they were also greatly influenced by Galton's work on mental ability (Galton, 1869). Galton proposed that, since all information enters us through our senses, the source of intellectual differences must be differences in sensory processing, hence the measurement of mental ability should focus on sensory acuity.

Wundt's laboratory was the optimal place to study variation in sensation, both in terms of equipment and theoretical orientation. Cattell, while still a student of Wundt, intended to change the topic of his dissertation to individual differences in mental ability. The idea was opposed by Wundt, hence only after submitting his thesis did Cattell start working on his'mental tests', which were published in 1890, and consisted mostly of standard tasks of psychophysics, modelled after experimental paradigms from Wundt's laboratory, such as '*Least noticeable difference in Weight'* or *'Reaction-time for Sound'* (Fancher, 1985). Titchener, the main advocate of Wundtian psychology, also opposed the application of the methods of experimental psychology to the study of individual differences (Brody, 2000).

It is likely that one of the reasons behind Wundt's and Titchener's reluctance is the perceived challenge it would have posed to the paradigm of introspection: "for so long as the general experimental psychology remained dominated by the Leipzig model for experimentation (…) such a merger would remain impossible, for there was (and is) no logical way to conform single-subject (N = 1) experimentation to population-level (N = many) studies of individual differences" (Lamiell, 2019, p. 63).

By the 1920s the introspective paradigm was replaced and statistical methods were developed that could have been used to scientifically study variation and covariation in psychological traits. Unfortunately, by that time Cattell's tests received a devastating blow. Clark Wissler, a student of Cattell carried out what was probably one of the first studies of test validity: he administered the tests to university students and found no correlation either between the different sensory tests themselves or with an external criterion of intelligence, university grades. Hence not only did the validation of the tests fail, the tests not even seemed to be measuring the same unitary construct. Also, by the 1920s an alternative approach to the measurement of abilities dominated the scene.

2.2 *Binet and the Invention of the Modern Intelligence Test*

The first 'modern' test of intelligence was invented by Alfred Binet in 1905. Binet's motivation was first and foremost practical: as the result of the introduction of mandatory education laws, schools found themselves having to cope with children who are unable to keep up with their peers. Therefore, it was imperative for the educational authorities to identify children who would struggle with the curriculum as early as possible and provide them with adequate training. The French government created a commission, led by Binet, to develop an objective diagnostic tool.

Binet's idea was simple but revolutionary: he proposed that children should be compared to the average performance of other children of different ages. The age to which a given child's performance corresponds was termed the child's level of intelligence and was compared with their chronological age. Subsequently, William Stern (whose work is discussed in detail in Sect. 5) suggested that it is more appropriate to divide by chronological age than to simply see if there is a discrepancy. The result of the division became known as the intelligence quotient or IQ.

Binet's approach was different from Galton and Cattell's in three ways. First, he decided to focus on tasks that require attention and judgment. Second, his method of scoring relied on comparison to peers, thus the result was a relative, rather than an absolute score like the reaction time measures advocated by the Anglo-Saxon pioneers. Eventually, the relative one became the standard approach with the introduction of the 'deviation IQ' by Wechsler which, contrary to its name, is not a quotient: it expresses one's relative standing by comparing one's raw score to the raw score distribution of one's peer group. Third, the emphasis shifted from scientifically explaining the causes of individual differences to measuring them as accurately as possible.

2.3 Intelligence Testing in the US and the Politics of IQ

Binet's tests were discovered in America by H. H. Goddard, who worked in a training school and needed an objective tool to measure the severity of 'mental retardation'. An appropriate adaptation was brought about by Lewis Terman, a pioneering educational psychologist at the Stanford Graduate School of Education, who extended Binet's test and normed it on a large samples. This new test was enormously successful; IQ became one of the most well-known terms of psychology in the public.

In the US, intelligence testing developed to a largely profitable industry and was a huge success for applied psychology, especially with the advent of group tests that can be administered to a large number of people at the same time. The first group test was developed by Robert Yerkes for the selection of officers in the US Army during the First World War, and was administered to 1.7 million recruits.

The leading American and British psychologists involved in ability testing had different conceptions from Binet both about measurement itself and about the construct being measured. Much of educational assessment took place in group settings, changing the interpersonal, diagnostic testing situation developed by Binet to a mass evaluation based on a multiple-choice format.

Crucially, according to Binet, the result of an intelligence test depicts a child's actual state of development compared to their peers; he believed that intelligence is malleable. The American pioneers, on the other hand, interpreted intelligence as an innate, fixed, general cognitive ability. Therefore, they conceived of test results as limitations and believed that an hour-long assessment in a group setting at age 10–12 is capable of predicting limitations in schools and in vocations for life.

The British and American pioneers of intelligence testing also shared an enthusiasm about Galton's 'other' legacy: eugenics, the "science of improving stock" (Galton, 1883, p. 17.) Galton established the Eugenics Society in 1908 in the UK, and his ideas quickly spread in the US, too. Many of the major figures in intelligence assessment, especially in the early period, were either regular members or officials in the Eugenics Society (White, 2006).

Eugenics was popular across the ideological spectrum, with such proponents as Churchill, G. B. Shaw, or Keynes. For a eugenicist IQ-tester the mission was obvious: IQ-tests should be used to reduce the reproduction of what was called feeble-mindedness in an 'objective', 'scientific' fashion. IQ-tests were, thus, widely used in the diagnostics that led to the forced sterilization of 60 000 'feeble-minded' people in the US in the twentieth century. At the same time, the eugenic program was not conditional on testing—the introduction of IQ-tests merely supplemented already existing ideas and already established sterilization practices with psychometric tools (Mackintosh, 2011a; b).

Finally, the study of race differences, and in particular their possible genetic origins have also made the field infamous in psychology. The related debate, the so-called 'IQ controversy' gained a lot of public interest. The works that caused the most heated debates are a paper by Arthur Jensen in which he proposed that the difference between the IQ scores of Blacks and Whites is genetic in origin (Jensen, 1969), and The Bell Curve, a book that focuses on the causes of social inequality, not race per se, but race is one of the main topics discussed (Herrnstein & Murray, 1994).

Overall, research on intelligence has at times intertwined with issues in social policy or even in politics for which individual differences in cognitive abilities seem to bear relevance. For instance, The Bell Curve is subtitled "Intelligence and Class Structure in American Life", and Jensen's paper begins with the claim "Compensatory education has been tried and apparently it has failed." (Jensen, 1969, p. 1). It is hardly surprising that if intelligence researchers enter issues of policy, then it will invoke reactions from outside the field of intelligence itself. For instance, many philosophers entered the debate on both sides of the argument (e.g. Block & Dworkin, 1974a, b; Cofnas, 2016; Sesardic, 2005; Urbach, 1974). Additionally, the political aspects of this line of research have, unsurprisingly, invoked politicians, too. President Bill Clinton, for instance, voiced his opinion about The Bell Curve, telling reporters that he could not accept its conclusions because they contradict his values, even though he admittedly had not read the book itself (Clinton, n.d.).

Arguably, a part of the research on race differences verges on pseudoscience and "gives IQ testing a bad name" (Mackintosh, 2007, p. 96.) Even though it was recently proposed in an editorial in Nature that "intelligence science need not be held back by its past" (Editorial, 2017, p. 386), and even though most research on intelligence has no direct political implications whatsoever, intelligence researchers are at times still looked at with suspicion of a secret political agenda. Dubious research on race differences and its political implications have caused many psychologists to keep a distance from the field of intelligence (Anderson, 1992). It has certainly not helped the convergence of research on individual differences in cognition and mainstream psychology that the former is perceived as one of the most controversial areas of psychology.

2.4 The Puzzle of g *and the Two Disciplines of Scientific Psychology*

The success of the testing industry in educational and military settings resulted in the rapid development of the science of psychological measurement itself. The Psychometric Society was established in 1935, making it one of the oldest still functioning scientific societies in psychology. The Society's journal, Psychometrika, was established already in 1936, and it has been publishing papers on scale construction, measurement theory, test reliability, validity estimation, multivariate analysis, etc. since. Societies devoted to basic research on cognitive abilities were founded much later. The International Society for the Study of Individual Differences was founded in 1983, and it covers personality research as well as abilities. The only society devoted exclusively to the study of variation in cognitive abilities is the International Society for Intelligence Research (ISIR), founded in 2000.

One of the main areas of inquiry in this field, perhaps with the highest relevance for cognitive psychology, regards the so-called 'structure of intelligence', i.e. the dimensionality of the covariance structure of test results. The central finding in this field is the 'positive manifold': the pattern of all-positive correlations that is observed when diverse mental tests are administered to a large sample of people. Even when the tests tap on apparently different domains, such as a vocabulary test and a mental rotation test, the observed correlations are always positive.

Intelligence researchers studying the structure of abilities apply the method of factor analysis, which helps to simplify large correlation matrices of cognitive test results, assuming that the correlation between any two tests is the result of the tests' correlation with a factor (the correlation is called the tests' loading on the factor). Factors are thus *latent variables* that are not directly measurable. Since there are positive correlations between all cognitive tests, factor analysis yields a strong general factor, *g*, that accounts for 40–50% of the variance (Deary et al., 2010; Jensen, 1998).

Yet among the pattern of all-positive correlations there are clusters of correlations that are stronger than others, and through factor analysis these clusters result in factors that represent specific cognitive abilities. For example, a vocabulary test, a reading comprehension test, and a test of general knowledge might reveal relatively strong positive correlations within the positive manifold, and this cluster is thought to reflect a factor of verbal or crystallized ability. There are different models of abilities but practically all of them accommodate both *g* and specific ability factors.

There has been much debate about the interpretation of *g*. Spearman, who discovered the positive manifold, argued that different tests correlate because "all branches of intellectual activity have in common one fundamental function" (Spearman, 1904, p. 284). Spearman hence proposed the *indifference of the indicator,* which means that the actual content of the test (spatial, verbal etc.) is irrelevant, any two tests with identical *g* loading are equivalent.

What is a matter of debate is not the existence of *g*; after all, *g* is a necessary algebraic consequence of the all-positive correlations between tests (Krijnen, 2004) and the positive manifold itself is a thoroughly replicated empirical phenomenon. The controversy stems from its interpretation: is it a general cognitive ability that indeed permeates all human cognition, regardless of its domain? is it a parameter affecting all abilities like the integrity of white matter tracts in the brain? or is it just a statistical phenomenon that does not reflect a psychological or biological construct? After more than a century, this debate is still not resolved. At the same time, in the field of intelligence, the majority view is that *g* represents GCA [General Cognitive Ability]: in a survey of expert opinion this received a mean agreement score of 4.47 out of 5 (Reeve & Charles, 2008). This interpretation, however, does not sit well with a number of findings from cognitive psychology and neuroscience that point to the domain-specificity of cognition. As discussed in detail in Sect. 3.3, individual differences constructs do not necessarily allow a within-individual interpretation.

The fundamental conceptual and methodological differences between experimental psychology and individual differences research led Lee Cronbach to claim in his presidential address to the American Pschological Association in 1957 that these are two distinct disciplines of scientific psychology and urged for unification: "One stream is experimental psychology; the other, correlational psychology. (…) Psychology continues to this day to be limited by the dedication of its investigators to one or the other method of inquiry rather than to scientific psychology as a whole. (…) While the experimenter is interested only in the variation he himself creates, the correlator finds his interest in the already existing variation. (…) The correlational psychologist is in love with just those variables the experimenter left home to forget. He regards individual and group variations as important effects of biological and social causes [while] individual variation is a source of embarrassment to the experimenter" (Cronbach, 1957, pp. 671–674.).

Today the term 'correlational psychology' is not in use any more, 'differential psychology' is typically used for this area of investigation. Relatedly, while researchers of individual differences were indeed restricted to correlational methods in Cronbach's time and only experimental psychologists could infer causality, advanced statistical techniques that have been developed since, such as structural equation modelling, have enabled individual differences researchers to test causal hypotheses, too.

Nevertheless, psychometrics had been drifting away from psychology for most of the twentieth century, Cronbach's call notwithstanding. As part of a recent oral history project 20 past presidents of the Psychometric Society participated in structured interviews and one of the project's most remarkable results is that "Some presidents express a certain ignorance of or lack of interest in what is going on in psychological research: They explicitly mention knowing little of psychology, or just not being interested in it. (…) it is currently possible to be a successful psychometrician and a president of The Psychometric Society, without having either a background or an active interest in psychology. (…) Having strong ties with mathematics or biostatistics is equally relevant and appropriate." (Wijsen & Borsboom, 2021, p. 333). While other presidents emphasized the contrary: their commitment to relate psychometrics to substantive issues in psychology, it seems to be an almost interdisciplinary interest rather than a natural connection between related fields within the same discipline.

From a historical perspective, it appears as if the difference between the approaches stemmed, to a large extent, from a fundamental difference in the attitude towards variation. Take, for instance, one of cognitive psychology's most cited discoveries, Miller's famous 'magical number 7', the limit of short-term memory capacity (Miller, 1956). The magical number seven was in fact seven *plus or minus two*; individual differences were acknowledged, but without attributing much significance to them.

It is also possible to interpret the results of intelligence assessment in a way similar to the magical number, that is, with an exclusive focus on the mean and ignoring the dispersion. Yet obviously no one would ever claim that the magical number of human intelligence is 100 plus or minus 30, even though approximately 96% of people do have an IQ of 100 plus or minus 30. It appears as if approaching variation as meaningful or as noise around a mean is to a large extent subjective.

At the same time, the unification Cronbach envisaged might be beneficial for both disciplines. The applied measurement of cognitive abilities is often painfully lacking in adequate theoretical background. "[I]t is the failure of individual-differences researchers to take theories seriously that has impeded scientific progress in the field". (Anderson, 2005, p. 283.)

On the other hand, experimental psychology could benefit from the statistical sophistication of psychometrics (Borsboom, 2006), for instance, complex statistical modelling developed for individual differences research can resolve methodological problems in experimentation (Maassen & Wicherts, 2019). Also, since the assessment of individual differences often takes place in high-stakes situations, this has led to high precision, i.e. a high reliability of tests. High psychometric reliability indicates that test score variance primarily reflects variance of the true score in the underlying construct and reflects measurement error to a small extent only.

Tasks in cognitive psychology are designed for within-subject experiments and as such often do not have adequate psychometric reliability. For example, classic tasks such as the Stroop, Eriksen flanker, Posner cuing, stop-signal, and go/no-go all exhibit poor test–retest reliability (Hedge et al., 2018). Similarly low or very low reliabilities were found for well-known neuropsychological tests such as the Wisconsin Card Sorting test (Bowden et al., 1998), or the Tower of Hanoi (Bishop et al., 2001; Gnys & Willis, 1991).

This means that many robust cognitive tasks are ill suited for individual differences research, which clearly is an obstacle in the way of unification. Additionally, the low reliability of measurements increases the chances of unsuccessful replications of findings.

3 Attempts at Unification

The earliest serious attempt at unification was Hunt's *cognitive correlates* approach (Hunt, 1978; Hunt et al., 1975) which sought to explain individual differences in verbal ability with the speed of lexical access. Hunt et al. used the Posner letter identification task, and concluded that the speed and accuracy of lexical access moderately, but significantly correlated with verbal ability.

The most influential integrative approach was Sternberg's *componential analysis*, which aimed at giving an account of the cognitive processes involved in solving problems similar to the ones in intelligence tests. According to Sternberg, "a component is an elementary information process that operates upon internal representations of objects or symbols." (Sternberg, 1980, p. 574). While many of his early findings are still relevant, Sternberg has since directed most of his research efforts to less 'cognitive' concepts such as *successful intelligence* (Sternberg, 2003). He once even compared the cognitive approach to intelligence to putting money in a dysfunctional parking meter, claiming that "a cognitive-experimental approach to the problem was not yielding the results I and many others had hoped for" (Sternberg, 2001, p. 190).

3.1 Conceptual and Methodological Issues

It would seem logical for the two disciplines to converge naturally, provided that one is supposed to investigate variation in the processes studied by the other. However, there are important methodological and conceptual differences. This can be demonstrated in the difference between the *structure of intelligence* and *cognitive architecture*, which are concepts from individual differences research and cognitive psychology, respectively. The structure of intelligence is a taxonomy based on the dimensionality analysis of the covariance structure of cognitive test scores. Cognitive architecture is a theory of the structure of the mind.

Broadly speaking, both are based on explorations of domain-specificity. Yet, domain-specificity bears different meanings in differential psychology and cognitive/experimental psychology. In the former it relates to the finding that individual differences in tests with characteristic content (e.g. spatial or verbal) typically correlate more strongly with one another than with tests that have different content. In the latter it means that the mind can be fractionated into processors of specific content through dissociation.

The two kinds of evidence for domain-specificity (dimensionality of covariance and dissociation) do not necessarily corroborate each other: processes can correlate and can still be dissociable (Kovacs, Plaisted, & Mackintosh, 2006). Therefore, it is a fallacy to conceptualize inter-individual structures, usually identified by factor analysis, as intra-individual constructs. In the field of intelligence, this manifests itself as the interpretation of the general factor of intelligence, *g,* an inter-individual construct, as general intelligence or general cognitive ability, an intra-individual construct. If *g* is identified as a within-individual construct, then the following statement is valid: "Bill relied on his general intelligence to correctly answer items on both the vocabulary test and the mental arithmetic test." This, however, is substantially different from saying that "If Bill performs better on the vocabulary test than most people it is very likely that he will perform better on the mental arithmetic test as well". The latter statement leaves the possibility open that Bill in fact did not use the same general ability on the two tests and there is some other reason for the results to correlate.

The positive manifold only translates to the second statement, not the first. Of course, the proposal that a general ability is involved in all cognitive performance is a sufficient explanation of the positive covariance between tests. But it is not a necessary explanation, and evidence for domain-specificity from cognitive psychology and neuropsychology questions its validity.

In cognitive psychology dissociation is demonstrated by several kinds of evidence, two of which are most dominant in adults and also most relevant for the discussion of within-individual versus between-individual methodology. The first comes from neuropsychological studies: an injury to one part of the brain results in the loss of an ability but leaves another intact, while an injury to a different part of the brain impairs the second ability but leaves the first intact. The second comes from experimental studies and is based on interference. If participants have to solve two tasks in parallel and performance on one does not deteriorate with the onset of the other, then the two tasks are considered to tap independent processes.

It is a fallacy to conclude from evidence for dissociation that fractionated processes do not correlate (for an example of this fallacy, see Churchland, 1996, p. 253). Imagine if one measured different indicators of strength in both arms in a large sample (the strength of grip, the maximum weight one can lift, etc.). Measures of the strength of people's left arm will most probably correlate with those of the right arm, regardless of (1) people holding a weight in their left arm being able to hold a weight in the right arm at the same time (lack of interference in an experimental condition), or (2) people can lose only one of their arms in an accident with the other arm remaining intact (selective impairment due to injury).

The interpretation of (double) dissociation is a matter of debate in cognitive psychology and cognitive science. The systems approach, similar to the concept of 'modularity' (Fodor, 1983), proposes that (double) dissociation is sufficient evidence to conclude that the tasks dissociated measure independent cognitive systems or modules. Opposed to this approach is the 'components of processing framework', which proposes that "different tasks may draw differentially upon different components in a processing system. If two tasks can be dissociated (…) then there must be at least one component process that figures differently in the two tasks (…). Within this framework, dissociations are no longer used to tease apart whole systems, but only differences in reliance on components within a larger system." (Bechtel, 2001, pp. 491–492). Results on individual differences provide further support for the component process approach: if two tasks that are dissociable correlate in a normal population, then it is more likely that they measure an overlapping, but not identical set of component processes, rather than completely independent, 'encapsulated' modules.

There have been arguments against the unification called for by Cronbach, not independently of such methodological and conceptual differences. Jensen (2000) and Borsboom et al. (2009) present a similar thought experiment in which extra-terrestrial creatures study cars and, because of the reasons discussed above, eventually develop separate lines of research for (a) how individual cars work, and (b) what makes them differ in performance.

Yet, arguably, this very analogy points to the necessity of unification, at least for the field of individual differences in intelligence. A 'differential car scientist' would quickly find differences in measures of performance and would also find that these differences correlate. For instance, faster cars typically have larger engines, but also more speakers, more vivid colours, and fancier hubcaps. Vividness of colour, number of speakers, engine size and hubcap fanciness would thus load on the same factor. Which of these things has a *causal role* in explaining individual differences in acceleration? Without a general understanding of what makes cars accelerate, this is an impossible puzzle. On the other hand, knowing that losing the hubcaps will not slow the car down is enormously informative.

Therefore, from the perspective of intelligence research, unification is necessary in order for research on variation in cognitive abilities to rely on the same explanatory set as cognitive psychology and to avoid explaining differential phenomena by proposing cognitive mechanisms and processes that are contradicted by findings in cognitive psychology. From a unified perspective, the puzzle of *g* can thus be interpreted as: *Why does the variation between people in test performance appear massively domain-general if the processes they employ to solve such tests are largely domain-specific?* This 'unified' question does address the positive manifold, a phenomenon largely neglected in cognitive psychology, but is drastically different from the one emerging solely from intelligence research (What domain-general within-individual psychological construct is the equivalent of *g?*).

3.2 Working Memory Capacity and Contemporary Efforts

The long history of separation was recently followed by a short history of convergence, and this is in large part due to research on working memory. Working memory is a construct developed by cognitive psychologists to refer to the *intra-individual processes* that enable one to hold goal-relevant information in mind, often in the face of concurrent processing and/or distraction (Baddeley, 1992; Baddeley & Hitch, 1974).

Working memory was operationalized in so-called complex span tasks that require the parallel storage and processing of information. One of the first such tasks was reading span, in which subjects have to read sentences and remember the last word of each sentence. Variation in performance on this particular measure of working memory capacity was found to predict performance on reading comprehension tests (Daneman & Carpenter, 1980). Over the years, many versions of complex span tasks have been developed (for a review of complex span tasks, see Conway et al., 2005).

They are much better predictors of complex cognitive performance than simple span tasks that require storage and retrieval only, like digit span. Importantly, it is the parallel storage and processing nature of complex span that is responsible for the prediction, regardless of the actual content of the task (Turner & Engle, 1989).

In contrast to simple span, variance in complex span tests is primarily domain-general: the correlation between verbal and spatial complex span tasks, such as reading span and symmetry span, is much higher than the correlation between a verbal and a spatial simple span task, such as word span or matrix span (Kane et al., 2004). Therefore, similar to IQ tests, a general factor of working memory capacity can be extracted through factor analysis. This factor, usually called *working memory capacity* correlates strongly with fluid intelligence (the ability to solve novel problems when one cannot rely on already acquired skills or knowledge): two meta-analyses estimate that the (latent) correlation is somewhere between $r = 0.72$ to $r = 0.81$ (Kane et al., 2005; Oberauer et al., 2005). Additionally, comparing the difference between the correlation with complex span and with simple span for different abilites, it is found that the difference is highest for fluid reasoning and smallest for verbal skills and general knowledge, i.e., what working memory capacity reflects beyond pure storage and retrieval is most strongly related to fluid, and least strongly to crystallized intelligence, with *g*—unsurprisingly—in the middle (Kovacs, 2010). Overall, working memory and intelligence are strongly related but not identical constructs, and working memory capacity is differentially related to different cognitive abilities.

There are (at least) three reasons why research on working memory is turned out to be the ideal construct to abridge experimental cognitive psychology and individual differences research. First, the construction of complex span tasks is embedded in mainstream experimental or cognitive psychology (Conway et al., 2008). Second, researchers could not commit the fallacy of interpreting the general factor of working memory as a single, unitary, within-individual, domain-general working memory system that is employed in every working memory task, regardless of content, similarly to how *g* is often identified with general cognitive ability; such an interpretation would have contradicted the very findings on domain-specific storage systems that established the concept of working memory in the first place. Third, unlike many tasks in cognitive psychology, working memory tasks demonstrate appropriate reliability (Engle & Kane, 2004).

This proliferation is indeed bidirectional: most well-known test batteries of intelligence started to incorporate tasks originally designed for working memory research, such as the letter-number sequencing task (Gold et al., 1997) in the Wechsler Intelligence Scales or the nonword-repetition task (Gathercole et al., 1994) in the most recent version of the Woodcock-Johnson Tests.

Additionally, process overlap theory (POT), a recent explanation of the positive manifold also builds on working memory capacity and its relation to intelligence (Kovacs & Conway, 2016, 2019). POT explicitly aims at bridging the gap between individual differences research and cognitive psychology. The main premise of POT is that a battery of intelligence tests requires a number of domain-general processes, such as those involved in working memory and attention, as well as a number of domain-specific processes; and the patterns of overlap of general processes with specific ones result in the emergence of the positive manifold. This interaction of general and specific processes as well as this bottleneck effect is formalized in a mathematical (item response) model (Kovacs & Conway, 2016).

POT also draws on the "components of processing" framework discussed in Sect. 3.2. It interprets dissociation (including neuropsychological, experimental, and developmental bases of dissociation) as evidence for fractionating processes rather than independent, encapsulated systems. That is, dissociated tests tap processes of which at least one is different, but not necessarily sets of completely different processes. Therefore, the processes that are required for performance on different cognitive tests can overlap and can also be dissociated by brain injury etc. *at the same time*. Evidence for dissociation between domain-specific cognitive tests makes it difficult to interpret *g* as 'general intelligence', a unitary system that permeates all human cognition. But such evidence is compatible with a theory that explains the correlations between these domain-specific tests as the result of overlapping component processes.

The most important consequence of POT is that *g* is a *formative* latent variable (Bagozzi, 2007): a summary of different, but correlated abilities, instead of being the reflection of a single, unitary ability as *g*-theory proposes. Similarly, under this framework, IQ is not conceptualized as the reflection of a single general cognitive ability, but rather as an index of mental functioning, composed of correlated specific abilities.

Therefore, contrary to the standard view, POT interprets *g* as an emergent rather than a latent property. A simulation based on the assumptions of POT demonstrated that it is indeed possible to fit a standard *g* model to data generated on the assumption of POT, i.e. without the casual role of a general mental ability (Kovacs et al., 2018). It appears as if there is a general intelligence at play even when the positive correlations are the result of a functional overlap. Importantly, this also demonstrates that being able to fit a *g*-model does not prove that the underlying structure contains anything that could be the psychological or biological equivalent of a general factor.

Overall, research on individual differences in working memory and the theoretical work that such research has proliferated are huge steps towards bridging Cronbach's two disciplines: "The point is that psychometricians and cognitive psychologists have joined forces to work together on the same problem—perhaps to the mutual benefit of both. The divorce between the two traditions of psychology (…) may be ending in a more or less happy reconciliation." (Mackintosh, 2011, p. 16.)

4 Intelligence in the Schools: Competing Approaches to Abilities

The instruments most widely used in educational and clinical practice are batteries that consist of numerous subtests with diverse content, and global ability indices are weighted sum scores calculated from the subtests or from lower-level indices. Examples of such global scores are FSIQ (Full Scale IQ) in the Wechsler scales or the GIA (General Intellectual Ability) index in the Woodcock-Johnson Tests of Cognitive Abilities.

There are two different approaches to interpreting such global indices. A 'top-down' approach builds on the concept of 'general intelligence', which, in turn, is based on the positive manifold: the finding that scores on all cognitive tests correlate positively. In the 'top-down' approach *g* represents general intelligence, a domain-general cognitive ability that, despite the superficially different content, is measured by all different subtests. IQ is a proxy for *g*, therefore, in this approach, IQ, or any global index obtained from a test battery is the most important indicator of cognitive performance. A 'bottom-up' approach puts larger emphasis on the actual tests and specific ability factors. Under this framework, global scores are conceptually identical to their technical manifestation: they are weighted averages, representing the average of a number of different cognitive abilities.

These different approaches can manifest themselves in actual batteries, too. For instance, the Wechsler scales were constructed under the assumption that while intelligence is unitary, it is expressed in multiple ways: "the subtests are different measures of intelligence, not measures of different kinds of intelligence" (Wechsler, 1958, p. 64). The WISC and the WAIS, therefore, consist of a large number of different tests, but they are not selected on the basis of any specific ability they might represent. Rather, the focus is on the sheer diversity of subtests. Only since the 4th edition do the Wechsler scales measure factors that can be meaningfully interpreted as specific abilities and provide specific ability scores in four domains beyond FSIQ.

The Woodcock-Johnson Tests of Cognitive Abilities, on the other hand, were constructed under the strong influence of CHC, a multi-dimensional model of the structure of abilities (McGrew, 2009) and they measure seven broad ability factors in the CHC framework. Even though a global ability index is provided—which is not called IQ—the test's focus is on providing a profile-type evaluation of abilities in a number of domains, thus enabling the mapping of individual strengths and weaknesses.

Crucially, these differing approaches of interpretation should be informed by theoretical frameworks, rather than only taxonomies of ability structure or predictive validity. For instance, if *g*-theory is correct and *g* is the equivalent of a general cognitive ability, then it is reasonable to focus on the global IQ score. If, however, POT is correct, and *g* is interpreted as a formative, rather than reflective variable, global indices such as IQ are reflections of *intelligence in general* rather than a proxy for *general intelligence.*

There is an additional dichotomy: the one between normative vs. ipsative assessment: relative strengths and weaknesses compared to others (normative) or compared to oneself (ipsative). It is possible that one's relative strength is below the average of a norm group, but information on intra-individual patterns might still be useful for instruction and development. Kaufman's idiographic *intelligent testing* approach emphasizes an ipsative, profile-based analysis (Kaufman, 1994; Kaufman & Lichtenberger, 2006), instead of only norm group comparisons on a global score. Arguably, profile-based assessment is much better suited for treatment or intervention purposes than a global indicator, especially if the letter is indeed a mere weighed average instead of a reflection of an overarching general ability.

5 Stern's Program(s) of Differential Psychology

Not only the concept of IQ, but the term *differential psychology* was also invented by William Stern, an influential German psychologist. Stern was also a prolific and influential researcher of personality and early child development (Pléh, 2010), and a forerunner of the positive psychology movement (Pléh, 2004, 2006). Stern declared the program of differential psychology in 1900 in the book *Über Psychologie der individuellen Differenzen (Ideen zu einer "differentiellen Psychologie)* [On the Psychology of Individual Differences: Toward a "Differential Psychology"]. In this first version he identified the three aims of differential psychology: to investigate the dimensions of inter-individual differences, to identify the causes of differences, and to explore how the consequences of differences manifest themselves in the schools or the workplace. This is an exhaustive description that covers most of the field of intelligence as it crystallized later on.

From the perspective of the current paper, the most important aspect of Stern's approach was that it was conceived as part of experimental psychology: "As Stern envisioned it, this sub discipline would not replace or in anyway compete with the general-experimental psychology formally established by Wilhelm Wundt (1832–1920) at Leipzig some two decades earlier, but would instead complement the general-experimental psychology by investigating (…) differences between individuals." (Lamiell, 2010, pp. 135–136.) The term 'differential psychology' rapidly spread and is still in use today, but, as seen in Sect. 2.4., it is conceived as another discipline of psychology, rather than an extension to experimental psychology.

In 1911, Stern published his second book on differential psychology: Die Differentielle Psychologie in ihren methodischen Grundlagen [Methodological Foundations of Differential Psychology]. This was not simply a new edition, but an extension of his program. The 1911 version of differential psychology is illustrated in Fig. 1. In Stern's theorizing, differential psychology was the collection of four methods with two different orientations for research. The first two, variation research and correlational research, focus on *attributes* through the study of many individuals. Variation research focused on individual differences in a single trait, whereas correlational research focused on the covariance patterns between at least two traits. The latter conceptually covers the factorial studies of the structure of abilities and other kinds of multivariate analysis.

Dissatisfied with a lack of focus on individuality in psychology, in 1911 Stern added two other methods that focus on individuals: psychography, the idiographic study of a single individual, and comparative studies, the comparison of two or more persons. Psychography was supposedly a non-statistical approach, but methods for ipsative testing were naturally not available at the time. It is not unreasonable to assume that Stern would have endorsed this method as a legitimate technique under

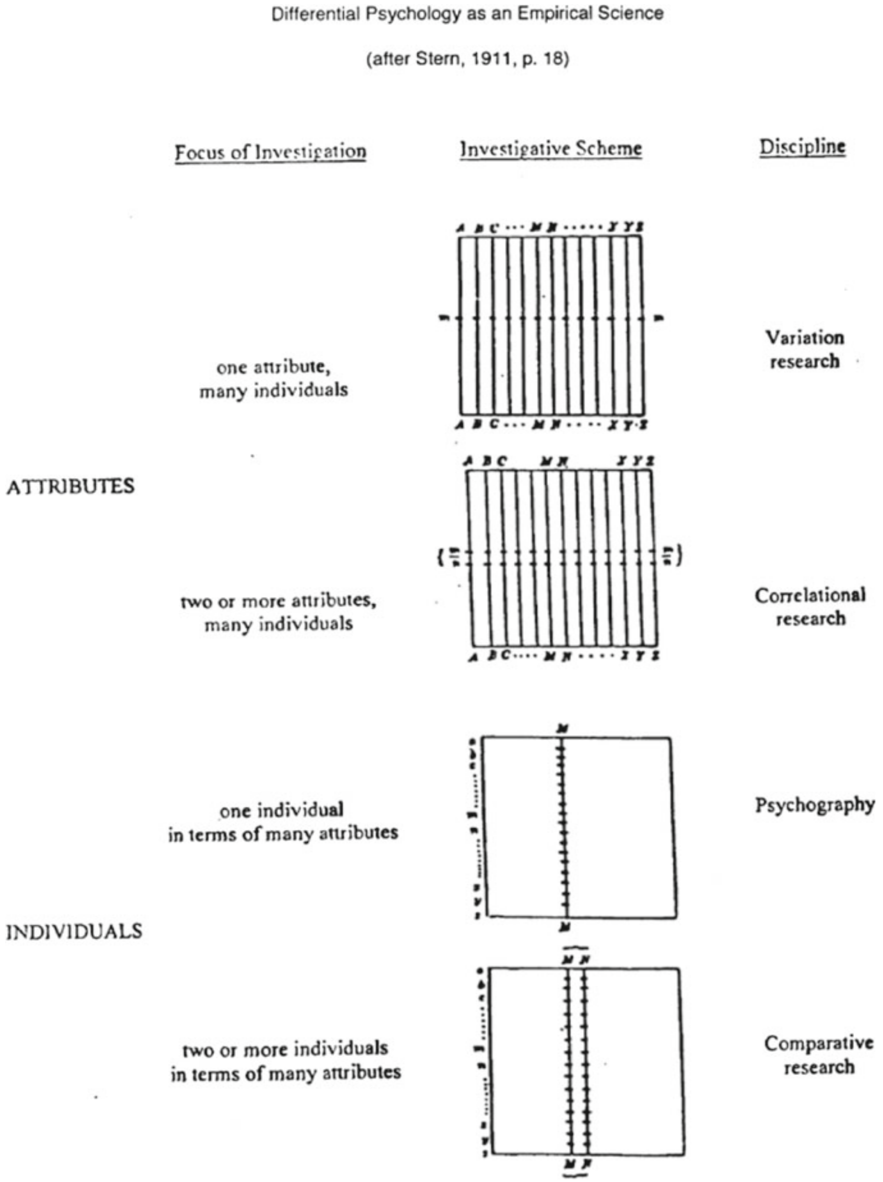

Fig. 1 The four research schemes of differential psychology in 1911: variation studies, co-variation studies, psychography, and comparison studies (from Lamiell, 2019, p. 54)

psychography. Especially since, as Pléh (2010, p. 423) points out, Stern's interests included psychological methods for career advising, which is one of the areas where ipsative assessment greatly compliments normative results.

Later on in his career, Stern was very critical of the exclusivity of normative testing (Lamiell, 1996). This included a critique of his own IQ concept: "Seventeen years ago, when I introduced the concept of the "intelligence quotient" as a measurement principle for intelligence tests, I had no idea that the "IQ" would become a kind of worldwide formula and one of the most frequently encountered expressions in American jargon. (…) But beyond that, many additional tests (…) have now been developed, standardized, and put into use, (…) always with emphasis on the objective, quantitative norm, with reference to which the single case is then compared." (Lamiell, 2012, p. 380) An ipsative approach breaks exactly this exclusivity of normative testing, hence it is quite possible that Stern would have approved of it. Unfortunately, only a handful of his works have been published in English. As a result, Stern is mostly remembered in the history of psychology as "the IQ Guy" (Lamiell, 1996).

Kaufman's classic book, *Intelligent testing with the WISC-R,* that popularised the ipsative approach, does not have Stern in its author index (Kaufman, 1979). At the same time, under the intelligent testing approach the "global IQ on any test, no matter how comprehensive, does not equal a person's total capacity for intellectual accomplishment" (Kaufman & Lichtenberger, 2006, p. 20.) Amazingly, this is almost identical to Stern's treatment from 1938: "'an intelligence quotient' may be of provisional value as a first crude approximation when the mental level of an individual is sought; but whoever imagines that in determining this quantity he has summed up 'the intelligence' of an individual once and for all, so that he may dispense with the more intensive qualitative study, leaves off where psychology should begin" (Stern, 1938, p. 60).

A telling piece of evidence of Stern's underappreciation is that, while he clearly is the originator of the term differential psychology itself (Jarl, 1958), a paper surveying the history of differential psychology only introduces Stern on the 6th page. He is mentioned as "another researcher whose work has not been as appreciated by Americans as much as it should" (Revelle et al., 2011, p. 9).

6 Personal Concluding Remarks

My own interest in individual differences emerged as an undergraduate student. My university did not offer a separate class on intelligence; most universities do not (Mackintosh, 2014). That I was admitted as a Ph.D. student at the University of Cambridge with a research proposal on a topic that I had never studied in my university curriculum was to a very large extent thanks to Csaba Pléh. Csaba was my 'tutor' for many years at Invisible College, an institution providing one-on-one tutoring to research oriented students. He accepted that I am one of his very few students with absolutely no sense to psycholinguistics, and carefully directed my interests in individual differences. In fact, Csaba was the one who gave me the textbook by Nick Mackintosh that resulted in my determination to conduct research under Nick's supervision.

Also, it is thanks to Csaba that I identify as a cognitive psychologist interested in individual differences. It has always seemed only natural. At the conferences of the Hungarian Cognitive Science Society (MAKOG), co-founded and co-organised by Csaba, I could present my undergraduate research on intelligence among talks on 'proper' cognitive topics. Little did I know back then how unusual that was.

When writing this paper, I found myself more and more absorbed by Stern's work, and since most of his writings have never been translated, for the first time I regretted not reading German at all. Csaba is—of course—familiar with Stern's work in its original. This paper would have been a better one if this book had not been meant as a surprise and I had approached him for some more mentoring.

Acknowledgements I am grateful to Judit Gervain and Bálint Forgács for their insightful comments to a previous version of this paper.

Funding The author received funding by the National Research, Development and Innovation Office of Hungary: Grant PD-17-125360 and Grant KH-18-130424.

References

About, Mission, & History—Psychonomic Society. (n.d.). Retrieved June 4, 2020, from https://www.psychonomic.org/page/about

Anderson, M. (1992). *Intelligence and development: A cognitive theory*. Blackwell.

Anderson, M. (2005). Marrying intelligence and cognition: A developmental view. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and intelligence: Identifying the mechanisms of the mind* (pp. 268–287). Cambridge University Press.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559. https://doi.org/10.1126/science.1736359

Baddeley, A. D., & Hitch, G. (1974). Working memory. *Psychology of Learning and Motivation, 8*, 47–89. https://doi.org/10.1016/S0079-7421(08)60452-1

Bagozzi, R. P. (2007). On the meaning of formative measurement and how it differs from reflective measurement: Comment on Howell, Breivik, and Wilcox (2007). *Psychological Methods*, *12*(2), 229–237; Discussion 238–245. https://doi.org/10.1037/1082-989X.12.2.229

Bechtel, W. (2001). The compatibility of complex systems and reduction: A case analysis of memory research. *Minds and Machines, 11*, 483–502.

Bishop, D. V. M., Aamodt-Leeper, G., Creswell, C., McGurk, R., & Skuse, D. H. (2001). Individual differences in cognitive planning on the tower of hanoi task: Neuropsychological maturity or measurement error? *Journal of Child Psychology and Psychiatry, 42*(4), S0021963001007247. https://doi.org/10.1017/S0021963001007247

Block, N. J., & Dworkin, G. (1974a). IQ, Heritability and Inequality, Part 1. *Philosophy & Public Affairs, 3*(4), 331–409.

Block, N. J., & Dworkin, G. (1974b). IQ, Heritability and Inequality, Part 2. *Philosophy & Public Affairs, 4*(1), 40–99. https://doi.org/10.2307/2264953

Boeck, P. De, Gore, L. R., González, T., & Martín, E. S. (2019). An alternative view on the measurement of intelligence and its history. In R. J. Sternberg (Ed.), *The Cambridge handbook of intelligence* (2nd ed., pp. 47–74). Cambridge University Press. https://doi.org/10.1017/9781108770422.005

Borsboom, D. (2006). The attack of the psychometricians. *Psychometrika, 71*(3), 425–440.

Borsboom, D., Kievit, R. A., Cervone, D., & Hood, S. B. (2009). The two disciplines of scientific psychology, or: The disunity of psychology as a working hypothesis. In *Dynamic process methodology in the social and developmental sciences* (pp. 67–97). Springer US. https://doi.org/10.1007/978-0-387-95922-1_4

Bowden, S. C., Fowler, K. S., Bell, R. C., Whelan, G., Clifford, C. C., Ritter, A. J., & Long, C. M. (1998). The reliability and internal validity of the wisconsin card sorting test. *Neuropsychological Rehabilitation, 8*(3), 243–254. https://doi.org/10.1080/713755573

Brody, N. (2000). History of theories and measurements of intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 16–33). Cambridge University Press.

Calvin, C. M., Deary, I. J., Fenton, C., Roberts, B. A., Der, G., Leckenby, N., & Batty, G. D. (2011). Intelligence in youth and all-cause-mortality: Systematic review with meta-analysis. *International Journal of Epidemiology, 40*(3), 626–644. https://doi.org/10.1093/ije/dyq190

Churchland, P. M. (1996). *The engine of reason, the seat of the soul: A philosophical journey into the brain.* MIT Press.

Cianciolo, A. T., & Sternberg, R. J. (2008). *Intelligence: A brief history*. Wiley-Blackwell.

Clinton, W. J. (n.d.). *Public papers of the presidents of the United States (1994, Book II).* U.S. Government Publishing Office. https://www.govinfo.gov/content/pkg/PPP-1994-book2/html/PPP-1994-book2-doc-pg1818-2.htm

Cofnas, N. (2016). Science is not always "self-correcting": Fact-value conflation and the study of intelligence. *Foundations of Science, 21*(3), 477–492. https://doi.org/10.1007/s10699-015-9421-3

Conway, A. R. A., Jarrold, C., Kane, M., Miyake, A., & Towse, J. (2008). *Variation in working memory*. Oxford University Press.

Conway, A. R. A., Kane, M. J., Bunting, M. F., Hambrick, D. Z., Wilhelm, O., & Engle, R. W. (2005). Working memory span tasks: A methodological review and user's guide. *Psychonomic Bulletin & Review, 12*(5), 769–786. https://doi.org/10.3758/BF03196772

Cronbach, L. J. (1957). The two disciplines of scientific psychology. *American Psychologist, 12*(11), 671–684. https://doi.org/10.1037/h0043943

Daneman, M., & Carpenter, P. A. (1980). Individual differences in working memory and reading. *Journal of Verbal Learning and Verbal Behavior, 19*(4), 450–466. https://doi.org/10.1016/S0022-5371(80)90312-6

Deary, I. J., Harris, S. E., & Hill, W. D. (2019). What genome-wide association studies reveal about the association between intelligence and physical health, illness, and mortality. *Current Opinion in Psychology, 27*, 6–12. https://doi.org/10.1016/j.copsyc.2018.07.005

Deary, I. J., Penke, L., & Johnson, W. (2010). The neuroscience of human intelligence differences. *Nature Reviews. Neuroscience, 11*(3), 201–211. https://doi.org/10.1038/nrn2793

Editorial. (2017). Intelligence test: Modern genetics can rescue the study of intelligence from a history marred by racism. *Nature*, *545*, 385–386

Engle, R. W., & Kane, M. J. (2004). Executive attention, working memory capacity, and a two-factor theory of cognitive control. *The Psychology of Learning and Motivation, 44*, 145–199.

Fancher, R. E. (1985). *The intelligence men: Makers of the IQ controversy.* Norton.

Fodor, J. A. (1983). *The modularity of mind*. MIT Press.

Galton, F. (1869). *Hereditary genius: An inquiry into its laws and consequences.* Macmillan.

Galton, F. (1883). *Inquiries into human faculty and its development*. Macmillan.

Gathercole, S. E., Willis, C. S., Baddeley, A. D., & Emslie, H. (1994). The children's test of nonword repetition: A test of phonological working memory. *Memory, 2*(2), 103–127. https://doi.org/10.1080/09658219408258940

Gensowski, M. (2018). Personality, IQ, and lifetime earnings. *Labour Economics*, *51*(2016), 170–183. https://doi.org/10.1016/j.labeco.2017.12.004
Gnys, J. A., & Willis, W. G. (1991). Validation of executive function tasks with young children. *Developmental Neuropsychology, 7*(4), 487–501. https://doi.org/10.1080/87565649109540507
Gold, J. M., Carpenter, C., Randolph, C., Goldberg, T. E., & Weinberger, D. R. (1997). Auditory Working Memory and wisconsin card sorting test performance in schizophrenia. *Archives of General Psychiatry, 54*, 159–165. https://doi.org/10.1001/archpsyc.1997.01830140071013
Hedge, C., Powell, G., & Sumner, P. (2018). The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. *Behavior Research Methods, 50*(3), 1166–1186. https://doi.org/10.3758/s13428-017-0935-1
Herrnstein, R. J., & Murray, C. (1994). *Bell curve: Intelligence and class structure in american life*. Simon and Schuster. http://books.google.com/books?hl=en&lr=&id=s4CKqxi6yWIC&pgis=1
Hunt, E. (1978). Mechanics of verbal ability. *Psychological Review, 85*(2), 109–130. https://doi.org/10.1037/0033-295X.85.2.109
Hunt, E., Lunneborg, C., & Lewis, J. (1975). What does it mean to be high verbal? *Cognitive Psychology, 7*(2), 194–227. https://doi.org/10.1016/0010-0285(75)90010-9
Jarl, V. C. (1958). Historical note on the term differential psychology. *Nordisk Psykologi, 10*(2), 114–116. https://doi.org/10.1080/00291463.1958.10780375
Jensen, A. (1969). How much can we boost IQ and scholastic achievement. *Harvard Educational Review, 39*(1), 1–123. https://doi.org/10.17763/haer.39.1.l3u15956627424k7
Jensen, A. R. (1998). *The g factor: The science of mental ability*. Praeger.
Jensen, A. R. (2000). "The g factor" is about variance in mental abilities, not a cognitive theory of mental structure: Reply to Anderson on Jensen on intelligence-g-factor. *PSYCOLOQUY*, *11*(041).
Journals—Psychonomic Society. (n.d.). Retrieved June 7, 2020, from https://www.psychonomic.org/page/journals
Kane, M. J., Hambrick, D. Z., & Conway, A. R. A. (2005). Working memory capacity and fluid intelligence are strongly related constructs: Comment on Ackerman, Beier, and Boyle (2005). *Psychological Bulletin*, *131*, 66–71; Author reply 72–75. https://doi.org/10.1037/0033-2909.131.1.66
Kane, M. J., Hambrick, D. Z., Tuholski, S. W., Wilhelm, O., Payne, T. W., & Engle, R. W. (2004). The generality of working memory capacity: A latent-variable approach to verbal and visuospatial memory span and reasoning. *Journal of Experimental Psychology. General, 133*(2), 189–217. https://doi.org/10.1037/0096-3445.133.2.189
Kaufman, A. S. (1979). *Intelligent testing with the WISC-R*. Wiley.
Kaufman, A. S. (1994). *Intelligent testing with the WISC-III*. Wiley.
Kaufman, A. S., & Lichtenberger, E. O. (2006). *Assessing adolescent and adult intelligence* (3rd ed.). Wiley.
Kovacs, K. (2010). *A component process account of the general factor of intelligence*. Unpublished Ph.D. thesis. University of Cambridge.
Kovacs, K., & Conway, A. R. A. (2016). Process overlap theory: A unified account of the general factor of intelligence. *Psychological Inquiry, 27*(3), 151–177. https://doi.org/10.1080/1047840X.2016.1153946
Kovacs, K., & Conway, A. R. A. (2019). What is IQ? Life beyond "general intelligence." *Current Directions in Psychological Science, 28*(2), 189–194. https://doi.org/10.1177/0963721419827275
Kovacs, K., Conway, A. R. A., Snijder, J., & Hao, H. (2018). *General intelligence explained (Away): Poster presented at: The 59th annual meeting of the psychonomic society; New orleans, 2018.11.15–18.*
Kovacs, K., Plaisted, K. C., & Mackintosh, N. J. (2006). Difficulties differentiating dissociations. In *Behavioral and brain sciences* (Vol. 29, Issue 2, pp. 138–139). https://doi.org/10.1017/S0140525X06349035
Krijnen, W. P. (2004). Positive loadings and factor correlations from positive covariance matrices. *Psychometrika, 69*(4), 655–660. https://doi.org/10.1007/BF02289861

Lamiell, J. T. (1996). William stern: more than "The IQ guy." In G. A. Kimble, M. Wertheimer, & C. White (Eds.), *Portraits of pioneers in psychology* (pp. 73–84). American Psychological Association.
Lamiell, J. T. (2010). Why was there no place for personalistic thinking in 20th century psychology? *New Ideas in Psychology, 28*(2), 135–142. https://doi.org/10.1016/j.newideapsych.2009.02.002
Lamiell, J. T. (2012). Introducing william stern (1871–1938). *History of Psychology, 15*(4), 379–384. https://doi.org/10.1037/a0027439
Lamiell, J. T. (2019). *Psychology's misuse of statistics and persistent dismissal of its critics.* Pelgrave.
Maassen, E., & Wicherts, J. M. (2019). Distinguishing specific from general effects in cognition research. *Journal of Applied Research in Memory and Cognition, 8*(3), 288–292. https://doi.org/10.1016/j.jarmac.2019.06.007
Mackintosh, N. J. (2007). Review of "Race differences in intelligence: An evolutionary hypothesis" by Richard Lynn. *Intelligence, 35*(1), 94–96. https://doi.org/10.1016/j.intell.2006.05.001
Mackintosh, N. J. (2011). *IQ and human intelligence.* Oxford University Press.
Mackintosh, N. J. (2011b). History of theories and measurement of intelligence. In R. J. Sternberg (Ed.), *The Cambridge handbook of intelligence* (pp. 3–19). Cambridge University Press. https://doi.org/10.1017/cbo9780511977244.002
Mackintosh, N. J. (2014). Why teach intelligence? *Intelligence, 42*, 166–170. https://doi.org/10.1016/j.intell.2013.08.001
McGrew, K. S. (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
Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review, 63*(2), 81–97. https://doi.org/10.1037/h0043158
O'Connell, M. (2018). The power of cognitive ability in explaining educational test performance, relative to other ostensible contenders. *Intelligence*, *66*(2017), 122–127. https://doi.org/10.1016/j.intell.2017.11.011
Oberauer, K., Schulze, R., Wilhelm, O., & Süss, H.-M. (2005). Working memory and intelligence--their correlation and their relation: Comment on Ackerman, Beier, and Boyle (2005). *Psychological bulletin*, *131*, 61–65; Author reply 72–75. https://doi.org/10.1037/0033-2909.131.1.61
Pléh, C. (2004). The tradition of positive psychology in Europe. *Ricerche Di Psicologia, 27*(1), 13–22.
Pléh, C. (2006). Positive psychology traditions in classical European psychology. In M. Csikszentmihalyi & I. S. Csikszentmihalyi (Eds.), *A life worth living: Contributions to positive psychology* (pp. 19–28). Oxford University Press.
Pléh, C. (2010). *A lélektan története*. Osiris.
Reeve, C. L., & Charles, J. E. (2008). Survey of opinions on the primacy of g and social consequences of ability testing: A comparison of expert and non-expert views. *Intelligence, 36*(6), 681–688. https://doi.org/10.1016/j.intell.2008.03.007
Revelle, W., Wilt, J., & Condon, D. M. (2011). Individual differences and differential psychology: A brief history and prospect. In *The Wiley-Blackwell handbook of individual differences* (pp. 1–38). Wiley-Blackwell. https://doi.org/10.1002/9781444343120.ch1
Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence, 53*, 118–137. https://doi.org/10.1016/j.intell.2015.09.002
Schmidt, F. L., & Hunter, J. (2004). General mental ability in the world of work: Occupational attainment and job performance. *Journal of Personality and Social Psychology, 86*(1), 162–173. https://doi.org/10.1037/0022-3514.86.1.162
Sesardic, N. (2005). *Making sense of heritability*. Cambridge University Press. https://doi.org/10.1017/CBO9780511487378
Spearman, C. (1904). "General Intelligence", objectively determined and measured. *The American Journal of Psychology, 15*(2), 201–292. https://doi.org/10.2307/1412107

Stern, W. (1938). *General psychology from the personalistic standpoint*. Macmillan.
Sternberg, R. J. (1980). Sketch of a componential subtheory of human intelligence. *Behavioral and Brain Sciences, 3*(4), 573–584. https://doi.org/10.1017/S0140525X00006932
Sternberg, R. J. (2001). How much money should one put into the cognitive parking meter? *Trends in Cognitive Sciences, 5*(5), 190. https://doi.org/10.1016/S1364-6613(00)01644-2
Sternberg, R. J. (2003). A broad view of intelligence: The theory of successful intelligence. *Consulting Psychology Journal, 55*(3), 139–154. https://doi.org/10.1037/1061-4087.55.3.139
Sternberg, R. J. (2020). A history of research on intelligence part II: Psychological theory, research, and practice in the nineteenth and twentieth centuries. In *Cambridge handbook of intelligence* (2nd ed., pp. 31–46). Cambridge University Press.
Turner, M. L., & Engle, R. W. (1989). Is working memory capacity task dependent? *Journal of Memory and Language, 28*(2), 127–154. https://doi.org/10.1016/0749-596X(89)90040-5
Urbach, P. (1974). Progress and degeneration in the "IQ debate" (I). *The British Journal for the Philosophy of Science, 25*(2), 99–135.
Wechsler, D. (1958). *The measurement and appraisal of adult intelligence*. Williams & Wilkins.
White, J. (2006). *Intelligence, destiny and education: The ideological roots of intelligence testing*. Routledge.
Wijsen, L. D., & Borsboom, D. (2021). Perspectives on psychometrics Interviews with 20 past psychometric society presidents. *Psychometrika, 86*(1), 327–343. https://doi.org/10.1007/s11336-021-09752-7