From Genes to Biology-Informed Cognitive Testing: Mapping the Genetic Architectures of Cognitive Functioning

Liliana G. Ciobanu
Adelaide Medical School
University of Adelaide, SA
AUSTRALIA

liliana.ciobanu@adelaide.edu.au

Andrew Heathcote
School of Psychology
University of Newcastle, NSW
AUSTRALIA

andrew.heathcote@newcastle.edu.au

Jimmy Breen
South Australian Health and medical Research Institute (SAHMRI), SA
AUSTRALIA

jimmy.breen@adelaide.edu.au

Nigel Rogash
South Australian Health and medical Research Institute (SAHMRI), SA
AUSTRALIA

nigel.rogasch@adelaide.edu.au

Lazar Stankov
School of Psychology
University of Sydney, NSW
AUSTRALIA

lazar.stankov@sydney.edu.au

Scott R. Clark
Adelaide Medical School
University of Adelaide, SA
AUSTRALIA

scott.clark@adelaide.edu.au

Eugene Aidman
Land Division, Defence, Science & Technology Group
Edinburgh, SA
AUSTRALIA

eugene.aidman@dst.defence.gov.au

ABSTRACT

Defence personnel are among a host of occupations selected and trained to perform under pressure. The premium on optimized cognitive performance is set to increase in future operating environments characterized by information overload, uncertainty, and complex decisions under time pressure. Cognition is broadly recognised as a key driver of both physical and mental performance, however little clarity exists about core elements of cognitive functioning that underpin optimal performance, their mechanisms and range of their modifiability. Specifically, there is limited understanding of the biological determinants of cognitive fitness in high stress environments.

Understanding the genomic architecture and molecular mechanisms of cognitive functioning is critical for developing a better understanding of how individuals may perform in cognitively challenging tasks. Current cognitive assessment tools have been developed to fit performance data, unrelated to underlying biology. While these standardised tests are useful for making decisions on aptitude and ability, they are not informative of the underlying biological phenotype and hence of limited help in designing biologically based interventions. Our project is building a biology-informed cognitive measurement framework that would enable more coherent system of performance-focused cognitive assessment and a more systematic and tailored approach to the development of monitoring and intervention protocols.

The genetic composition of the general cognitive ability (g), examined through the genome-wide association (GWAS) methodology, remains inconclusive, due to large variations in phenotype definition and measurement. The problem increases when considering sub-components of g *and non-*g *dimensions of cognitive functioning, such as executive functioning. To progress our understanding of the genetic architecture of cognitive functioning, a radical improvement in phenotype definition and measurement is required. To this end, we have chosen to focus on the smallest measurable components available in cognitive test data and examine their genetic associations. This solution, proposed 30 years ago (Atchley and Hall 1991), has not been feasible due to (1) lack of understanding of how genetically separable multiple cognitive abilities are, and (2) the highly polygenetic nature of cognition, which drives the requirement for massive sample size (hundreds of thousands) for cognitive genomics studies. Our previous study began addressing these limitations by demonstrating genetic separability of g from specific cognitive abilities such as executive functioning (Ciobanu et al., 2021). Our current study refines the core dimensions of cognitive functioning by examining both genetic and cognitive assessment data from the world's largest dataset (UK Biobank) with over 500,000 participants. First, we will identify the genetic variants associated with basic cognitive units by applying multivariate mixed model GWAS methodology. Second, for each of these variables, we will compare patterns of association across the genome. Third, we will examine how test-specific genetic variants combine to represent broader cognitive constructs and how these broader constructs are interrelated. Together these associations will shape our genome-informed model of cognitive functioning, which, in turn, will inform the design and application of cognitive assessment tailored to the distinctly different requirements of selection, training and operational support applications.*

1.0 INTRODUCTION

1.1 Intelligence and the assessment of cognitive processes

1.1.1 The nature of intelligence

There have been two somewhat different approaches to the study of cognitive abilities and intelligence. One has focused on individual differences in a single construct of intelligence across the population. The other on a more detailed multi-level, multi-component model applicable in clinical neuropsychology and potentially more translatable to underlying granular biological analysis.

Considering the first approach, many contemporary studies can be traced back to the work of Charles Spearman (1904) and the discovery of the general intelligence factor 'g' [1], based on his observation of positive correlations between a variety of cognitive tasks. Using a statistical method of factor analysis Spearman claimed that '*g*', a single underlying intelligence factor, accounts for individual differences on a variety of observable abilities. Statistically, '*g*' is a single factor/component that explains about 40% of the variance in performance on IQ tests. Psychometrically, '*g*' captures a person's performance across any number of cognitive tasks. Psychologically, 'g' was interpreted as something like an "energy" or "power" as a function of performance across the whole cortex. More recent the work of A. Jensen (1923-2012) and H. Eysenck (1916-1997) links *g* to mental speed as captured by the measures of inspection time and choice reaction time. The existence of a single quantifiable factor for human intelligence continues to be hotly debated.

In contrast, contemporary standardized tests of intelligence followed the work of French psychologist Alfred Binet (1857–1911) who believed that intelligence was complex and could not be fully captured by a single quantitative measure. He assumed that intelligence consisted of a variety of "higher mental processes" – attention, memory, imagination, common sense, judgment, abstraction, and coping successfully with the world. In the 1930's, Thurstone (1887-1955) also questioned the role of 'g' and emphasized range of primary mental abilities such as verbal comprehension, word fluency, number facility, associative memory, and

reasoning. It is important to note that "higher mental processes" were explicit in excluding from the definition of 'g' and intelligence what was referred to as simple and elementary (e.g., sensory) sub-processes.

The American Psychological Association's defines intelligence broadly as "the ability to derive information, learn from experience, adapt to the environment, understand, and correctly utilize thought and reason." (https://dictionary.apa.org/intelligence). The implication is that processes captured by the tests of cognitive abilities can predict real-life behavior. Indeed, Intelligence Quotient (IQ) tests were developed to inform educational decisions for children and examine the aptitude of adults for certain careers. For example, the tests are used to identify children whose academic work suffers or who are "gifted" and can benefit from additional training. With adults, they have been used to select students who can enroll in higher educational institutions or for managerial jobs in business and military establishments.

Even though IQ and Spearman's '*g*' are widely treated as indices of human intelligence, there has only been limited success in understanding the underlying biological processes. It may be more profitable to focus on processes that are less general in their scope. This may be justified from both theoretical and practical points of view.

1.1.2 Process Overlap Theory (POT) and Neuropsychological/clinical Aspects of Intelligence

Shortly after the appearance of Spearman's (1904) paper [1], G. Thomson (1881-1955) proposed a sampling theory of intelligence. According to this theory, every test samples a range of the elementary human abilities. A 'g' factor emerges because each "higher-order process" captures a number of common elementary processes. More recently Kovacs and Conway (2019) have proposed Process Overlap Theory (POT) which describes overall performance on a test as a function of multiple domain-general and domain-specific abilities [2]. According to this theory 'g' is a formative construct since, as illustrated in simulation, it can appear in the absence of a general process. Broad lower-order processes proposed include fluid (Gf) and crystallized (Gc) intelligence. Similarly, based on clinical insights from patients with organic impairment in cognitive function, neuropsychology testing is designed to identify a pattern of deficits across a range of domains rather than a deficit in general measure. Such approaches are more sensitive and specific for diagnosis and management planning for patients with dementias such as Alzheimer's disease. Potentially, the chance of identifying biological substrata of less complex lower-order processes is more promising than focussing on a more complex general factor.

Neuropsychological assessment is performed using a battery of tests that sample performance across overlapping domains of thinking. Executive function (EF) is a collection of processes commonly measured in neuropsychological testing that act to monitor, analyse and control behaviour. For example, the Trail Making test requires the connection of a series on numbers and letters in consecutive order. It assesses visual attention and task switching. Another example is the Stroop test that involves naming the colour of a word when there is a mismatch between the name of a colour (e.g., "blue", "green", or "red") and the colour it is printed in (i.e., the word "red" printed in blue ink instead of red ink). When asked to name the colour it takes longer and is more prone to errors when the colour of the ink does not match the name of the colour. The test measures the ability to inhibit cognitive interference, as well as attention, processing speed, cognitive flexibility. These kinds of tasks tend to have low correlation with typical IQ tests. Two of these EF components - working memory and cognitive flexibility – are known to correlate with intelligence. The third component - inhibitory control – is clearly outside its scope and may therefore be separable at a biological level.

1.1.3 The Cattell-Horn-Carroll (CHC) theory

The past several years have witnessed a significant increase in the popularity of the Cattell-Horn-Carroll theory of intelligence. The latest version of this theory was described by Schneider and McGrew [3]. One of the predecessors of this model was the theory of fluid (Gf) and crystallized (Gc) intelligence. This has been

expanded from the original two factors into a ten broad factors structure. They define Gf as "…the use of deliberate and controlled procedures (often requiring focussed attention) to solve novel, "on-the-spot" problems that cannot be solved by using previously learned habits, schemas, and scripts" (p. 93). They define Gc as "the ability to comprehend and communicate culturally valued knowledge" (p. 114). In addition to these two, the well-established broad CHC abilities include: Quantitative knowledge (Gq), Reading & Writing Ability (Grw), Short-Term Memory (Gsm), Long-Term Storage and Retrieval (Glr), Visual Processing (Gv), Auditory Processing (Ga), Processing Speed (Gs), and Decision/Reaction Time/Speed (Gt). To this list they added additional tentative broad factors linked to sensory modalities – tactile (Gh), psychomotor (Gp), psychomotor speed (Gps), kinaesthetic (Gk), and olfactory (Go) - described in [4].
The CHC theory does not emphasize the role of 'g'. Instead, it points to important cognitive processes that have been missed because of the emphasis on 'g'. These processes are more specific and are likely to be better accounted for by the biological functions. In our approach we will focus on categories of cognitive tasks that vary in their specificity.

1.1.5 Disconnection between cognitive tests and the biology underlying cognitive functioning

Despite significant advances in neuroscience in understanding the neurobiology of cognitive processes, current cognitive assessment tools have been developed to fit performance data, unrelated to the underlying biology. While these standardised tests are useful for making decisions on aptitude and ability, they are not informative of the underlying biological phenotype and hence of little help in designing biological interventions. Historically, psychometrics operate at the conceptual level, while neuroscience is typically preoccupied with neuroanatomical and molecular correlates. This disconnect between cognitive tests and biology could be described using the philosophical term "genetic fallacy" – when a conclusion is based solely on a principle's origin, rather than its true context, meaning that cognitive tests available to date (conclusion), are based on historical practices of concept-based cognitive research, rather than on its true context of biological origins of individual differences in cognitive functioning. Ironically, "genetic fallacy" can be taken quite literally here – original accounts of psychometrics may be true, and they illuminate the reasons why modern cognitive tests are in its present form, but they are not conclusive in determining its merits without understanding genetics as a foundation to studying the structure of cognition and developing biology-informed measurement tools. Our project aims to address this gap by building a genetically informed cognitive measurement framework that will enable a more coherent system of performance-focused cognitive assessment and a more systematic and tailored approach to the development of monitoring and intervention protocols.

1.2 What do we know about the genetic architecture of cognition?

DNA is the inherited basis of individual differences modified by natural selection to drive evolutionary change. Twin and adoption studies have shown that genetic influence on individual differences in intelligence is substantial [5, 6]. Studying the genetic determinants of cognitive functioning is essential to understanding the underlying biology of performance. However, cognitive genomics has proved to be a challenging area of research due to a lack of consensus on the theoretical construct of cognitive functioning. Furthermore, cognition is highly polygenic, and each implicated variant only accounts a small amount of variance. Consequently, large sample sizes are required for the discovery of genetic variants associated with cognitive phenotypes at acceptable levels of statistical significance. While the problem of adequate power can be resolved by uniting the global efforts on collecting and analysing data, the definition of cognitive domains and their measurement instruments continues to contribute to the poor replicability and translation of findings. The recent success in identifying genetic correlates of general cognitive ability, *g* [7], has greatly added to appreciation of the genetic complexity of cognition. This construct is typically derived as the first unrotated principal component of multiple cognitive test metrics, under the assumption that g captures about 25 to 40% of the total variance when a battery of multiple cognitive tests is administered to a sample with a good range of cognitive ability [8, 9]. However, due to the general nature of *g*, translatability of these findings to specific performance tasks is limited (Figure 1)

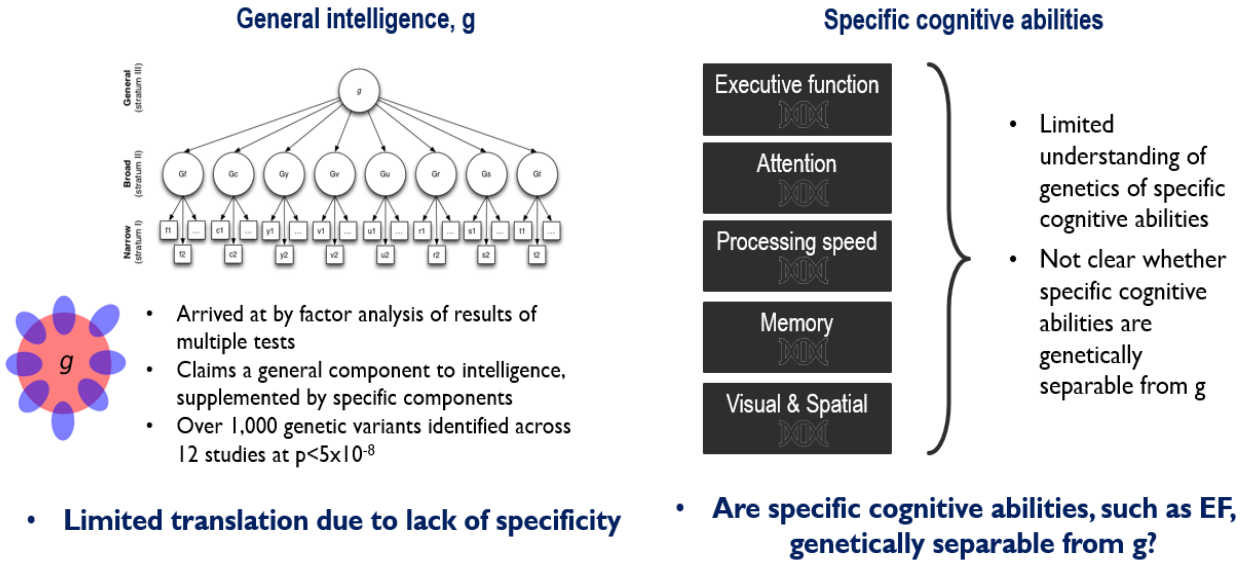

Figure 1: A current status of understanding of genetic architecture of cognition

2.0 APPROACH

2.1 Setting the stage: are specific cognitive abilities genetically separable from *g*?

The question of separability of executive function from general intelligence has been an ongoing debate for decades. While it is well known that *g* and executive function (EF) are overlapping at the phenotype as well as at the aggregate genetic levels, it is largely unknown whether these two constructs are genetically separable at the individual variant level. A deeper understanding of the relationships between individual genes and cognitive domains will enable a more detailed understanding of the molecular pathways that are shared or specific to particular cognitive tasks. In turn this knowledge will allow the development of interventions, targeted to optimise or normalise these pathways. In this study, we further develop a proof of a concept of a genetic separability of specific cognitive abilities at the individual genetic variant level (Ciobanu et al., 2021). We have previously explored whether specific cognitive abilities are separable from g, by analysing a subset of genome-wide association studies (GWAS) from the GWAS catalogue – a comprehensive database of all GWAS studies conducted to date (https://www.ebi.ac.uk/gwas/), that used measures of g and EF as outcomes in healthy participants free from psychiatric, medical or neurological conditions. The analysis identified two sets of single nucleotide polymorphisms (SNPs) associated with g (1,372 SNPs across 12 studies), and EF (300 SNPs across 5 studies) at $p<5x10^{-6}$. A comprehensive SNP-based functional annotation was conducted, followed by pathways analyses of SNP lists across the g and EF studies. We found that while some genetic variants are common for g and EF, executive functions appear to be separable from general intelligence at both structural and functional levels. Due to the limitations in sample size and in combining studies with different operationalisations of g and EF more evidence is required to better characterize these relationships. To the best of our knowledge, this study is the first to systematically compare structural and functional genetic correlates of general intelligence and executive function at an individual SNP level. It provides biologically informed evidence to inform cognitive enhancement programs focused on modifiable executive functions and can serve as a guide for future research in the field (Ciobanu et al., 2021) (Figure 2).

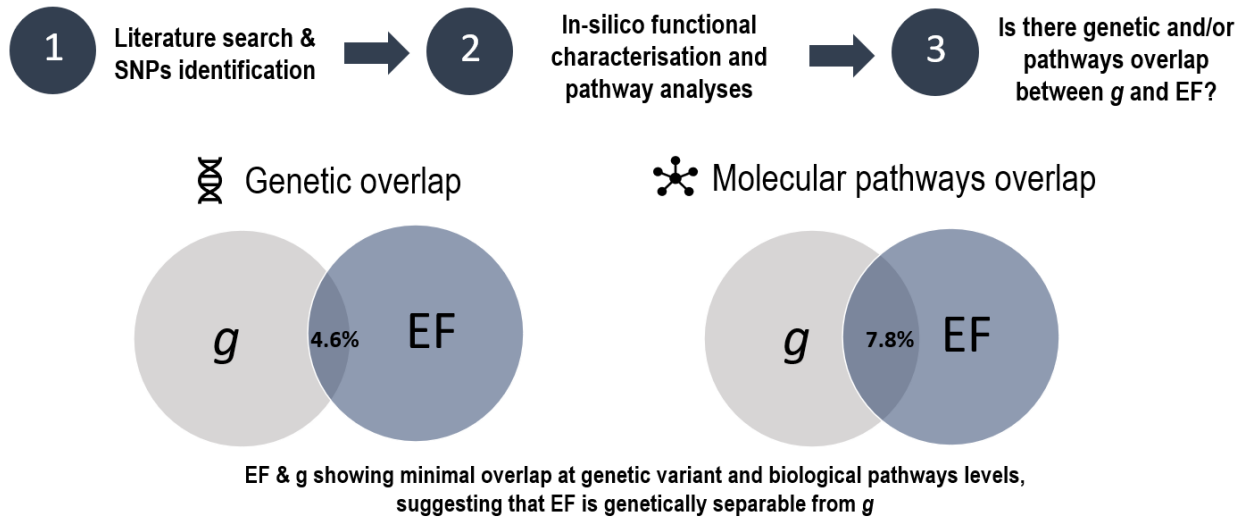

Figure 2: Analytical strategy and the main results of the study on genetic separability of EF from *g*

To further progress our understanding of the genetic architecture of cognitive functioning, a radical improvement in phenotype definition and measurement is required. To this end, to refine the core dimensions of cognitive functioning, we propose a novel framework that focuses on the smallest measurable components available in cognitive test data and examine their genetic associations. First, we propose to identify the genetic variants associated with selected cognitive test variables by applying multivariate mixed model GWAS methodology. Second, for each of these variables, we will compare patterns of association across the genome. Third, we will examine how test-specific genetic variants combine to represent broader cognitive constructs and how these broader constructs are interrelated. Together these associations will shape our genome-informed model of cognitive functioning, which, in turn, will inform the design and application of biology-informed cognitive assessment battery.

2.2 A novel framework for deriving biology-informed structure of cognition

Understanding the genomic architecture of cognition in healthy individuals is critical to identifying targets to developing effective training to enhance cognitive functioning as well as for treating impaired cognition. The lack of a clear-cut idea of what constitutes a basic unit of complex cognitive behaviour has been a major problem in understanding the genetic architecture of the hierarchical structure of cognition [10]. We argue that to shed the light on genetic underpinnings of cognitive sub-components and their correlated structure, analyses should focus on the smallest measurable units likely to display genetic variation. In an attempt to identify a basic unit of cognitive behaviour, we propose to utilise a new unified model of general intelligence based on mathematically defined network modeling [11, 12]. The Van Der Maas's mutualistic network model describes mechanisms on the level of the individual and can explain correlational structure between cognitive subtests scores – a major phenomenon in current research on cognition and intelligence that can be viewed as the empirical basis for the g and more complex factorial structures. In this model, the cognitive system is assumed to comprise of a set of basic cognitive abilities (e.g., working memory, reasoning ability, vocabulary, etc.) with a change in a process modeled as a function of autonomous change of each sub-process. The mutualistic interactions between the different processes are represented by the interaction matrix, which is assumed to be equal in all people (see [12] for more details).

Developing a network is not the final word. Given the findings of previous studies on specific cognitive abilities that utilised a single cognitive test as an approximation for the smallest measurable unit of cognitive

behaviour [13] and our recent study (described above), which show that more defined clusters of cognitive abilities, such as EF, are genetically separable [14], we hypothesise that it is possible to identify intrinsic genetic variants and biological pathways associated with basic units of cognitive behaviour defined as basic sub-components of cognition in a correlational network. Using this approach we hope to bridge the worlds of cognition and biology and evolve the biology-informed structure of cognitive functioning.

2.2.1 Methodology

Due to complexity of genotype-cognitive phenotype relationships, simultaneous analyses of genomic associations with multiple cognitive traits will be more powerful and informative than a series of discrete univariate analyses. Our study refines the core dimensions of cognitive functioning by examining both genetic and cognitive assessment data from the world's largest dataset (UK Biobank: https://www.ukbiobank.ac.uk/) with over 500,000 participants. First, we apply multivariate GWAS on cognitive assessment and genotype data of healthy participants to test for genetic associations with basic sub-components (units) of cognition in a multivariate mixed model implemented in GEMMA while accounting for the relatedness of individuals and pedigree structures, as well as population substructure. Second, we assess the degree to which variation in each basic unit shares the same patterns of association across the genome by computing the linkage disequilibrium score correlation (LDSC) based on GWAS P values for each pair of sub-components of cognition. Third, we will use structural equation modelling (SEM) specifying both measurement (how basic units-specific genetic variants come together to represent broader cognitive construct) and structural (how broader cognitive constructs are related to one another) models to refine our understanding of which groups of genetic markers best explain the variance observed in each cognitive unit and map genetic correlates of a basic cognitive sub-component to broader cognitive constructs. By implementing an open-ended multivariate association method, in which the inherent phenotypical variation within each of these segments drives the association, we will describe association between SNPs and cognitive traits as well as likely biological functions of the regions surrounding these SNPs. We also highlight regions with multiple SNPs affecting different cognitive phenotypes as well as evidence for multiple SNPs working in concert to produce a single phenotype.

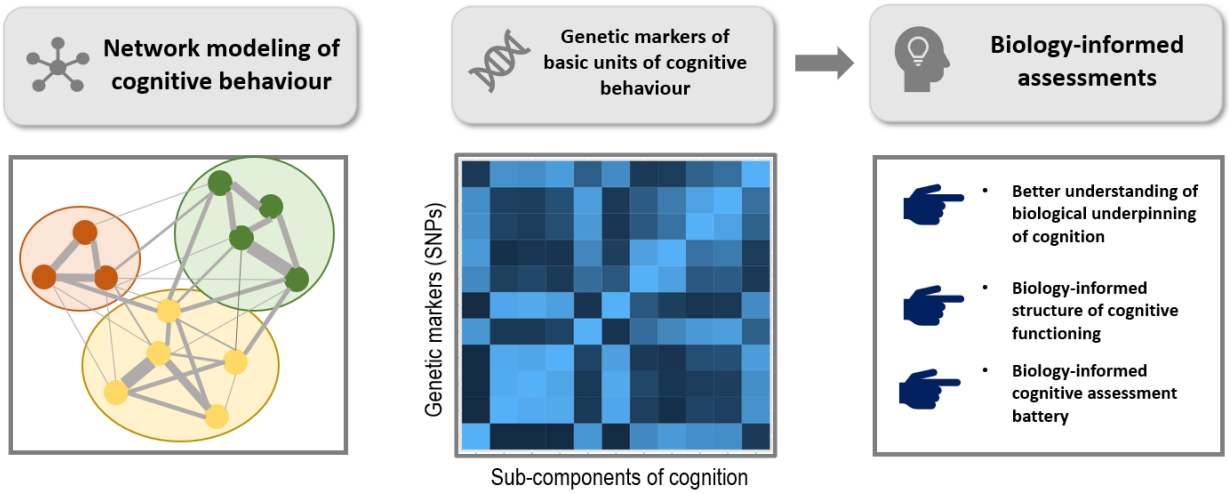

Figure 3: A summary of a novel network-based framework on biology-informed structure of cognition

3.0 SUMMARY AND FUTURE DIRECTIONS

In summary, our framework will explore the genetic basis of phenotypic relationships between different

measures of cognitive functioning and provide insights into the understanding of how complex traits of cognitive functioning are shaped by both individual and coordinated genetic actions, to enable a biologically informed methodology for measuring core elements of cognition. This research will enable a better understanding of genetic architecture and complex biological processes underlying cognitive functioning. It will produce new biological and genetic evidence to inform a review of the structure of core elements of cognitive functioning, and the development of common standards for biologically informed cognitive assessment. Together, these will provide a common platform for future research into biology-based interventions to improve and enhance cognitive functioning in high-functioning individuals.

4.0 REFERENCES

[1] Spearman, C., *"General Intelligence," Objectively Determined and Measured.* The American Journal of Psychology, 1904. **15**(2): p. 201-292.

[2] Kovacs, K. and A.R.A. Conway, *A Unified Cognitive/Differential Approach to Human Intelligence: Implications for IQ Testing.* Journal of Applied Research in Memory and Cognition, 2019. **8**(3): p. 255-272.

[3] Schneider, W.J. and K.S. McGrew, *The Cattell–Horn–Carroll theory of cognitive abilities*, in *Contemporary intellectual assessment: Theories, tests, and issues, 4th ed.* 2018, The Guilford Press: New York, NY, US. p. 73-163.

[4] Stankov, L., *Diminished 'g': Fluid and crystallized intelligence and cognitive abilities linked to sensory modalities*, in *General and Specific Mental Abilities*, D.J. McFarland, Editor. 2019, Cambridge Scholars Publishing: Cambridge, UK.

[5] Deary, I. and L. Lopez, *Genetic Foundations of Human Intelligence.* Human genetics, 2009. **126**: p. 215-32.

[6] Wright, M., et al., *Genetics of Cognition: Outline of a Collaborative Twin Study*. Twin Research, 2001. **4**(1): p. 48-56.

[7] Spearman, C., '*General Intelligence', Objectively Determined and Measured.* American Journal of Psychology, 1904. **15**: p. 201-93.

[8] Plomin, R. and I.J. Deary, *Genetics and intelligence differences: five special findings.* Molecular Psychiatry, 2015. **20**(1): p. 98-108.

[9] Stankov, L., *Overemphasized "g"*. Journal of Intelligence, 2017. **5**: p. 33.

[10] Atchley, W.R. and B.K. Hall, *A model for development and evolution of complex morphological structures* 1991. **66**(2): p. 101-157.

[11] van der Maas, H.L.J.D., Conor V.Grasman, Raoul P. P. P.Wicherts, Jelte M.Huizenga, Hilde M.Raijmakers, Maartje E. J., *A Dynamical Model of General Intelligence: The Positive Manifold of Intelligence by Mutualism.* Psychological Review, 2006. **113**(4): p. 842–861.

[12] Van Der Maas, H.L.J., et al., *Network Models for Cognitive Development and Intelligence.* 2017. **5**(2): p. 16.

[13] Ibrahim-Verbaas, C.A., et al., *GWAS for executive function and processing speed suggests involvement of the CADM2 gene.* Molecular psychiatry, 2016. **21**(2): p. 189-197.

[14] Hatoum, A.S., et al., *GWAS of Over 427,000 Individuals Establishes GABAergic and Synaptic Molecular Pathways as Key for Cognitive Executive Functions.* 2019: p. 674515.