Cognitive genomics: recent advances and current challenges

Authors:

Joan Fitzgerald BSc, HDipPsych, MSc.

Derek W. Morris BSc, Ph.D

Gary Donohoe BA, BD, HDip, MPsychSc, DClin Psych, PhD

Affiliation:

Neuroimaging, Cognition & Genomics (NICOG) Centre, School of Psychology and Discipline of Biochemistry, National University of Ireland Galway, Galway, Ireland

Corresponding author:

Gary Donohoe
Rm 1040 School of Psychology
National University of Ireland
University Road, Galway.
Email: gary.donohoe@nuigalway.ie

Word Count: 4,150

Abstract:

Purpose of review:

We review recent progress in uncovering the complex genetic architecture of cognition, arising primarily from genome-wide association studies (GWAS). We explore the genetic correlations between cognitive performance and neuropsychiatric disorders, the genetic and environmental factors associated with age-related cognitive decline, and speculate about the future role of genomics in the understanding of cognitive processes.

Recent Findings:

Improvements in genomic methods, and the increasing availability of large datasets via consortia cooperation has led to a greater understanding of the role played by common and rare variants in the genomics of cognition, the highly polygenic basis of cognitive function and dysfunction, and the multiple biological processes involved.

Summary:

Recent research has aided in our understanding of the complex biological nature of genomics of cognition. Further development of data banks and techniques to analyses this data hold significant promise for understanding cognitive ability, and for treating cognitively related disability.

Introduction:

Long before the development of modern genomic methods, as far back as the early 1900's, the heritability of cognitive performance was recognized through twin and adoptive studies [1]. In a study of ~10,000 monozygotic and dizygotic twins, concordance in measures of intelligence was found to be 0.86 and 0.60 respectively [2, 3]. Follow-up longitudinal twin research had further shown that heritability actually increases during childhood development; this is explained by genetic innovation in early childhood, whereby increasing numbers of genes become activate during cognitive development, thus amplifying the contribution of genetics over environment [4]. Given that the estimates of heritability of intelligence, estimated at 50% across the lifespan, it was originally assumed that it was only a matter of time until the key gene(s) involved in cognition were identified [5, 6*]. However, the complex and highly polygenic nature of cognitive phenotypes is now well established, with literally hundreds of genes statistically associated with variation in cognitive function, and implicating a wide variety of processes related to brain development and neuron to neuron communication [7**, 8*]. This article reviews recent progress in uncovering the complex genetic architecture of cognition, arising primarily from genome-wide association studies (GWAS). We discuss the genetic correlations between cognitive performance and neuropsychiatric disorders, and the genetic and environmental factors associated with age-related cognitive decline. We conclude by speculating about the future role of genomics in the understanding of brain function and cognitive processes.

Measuring cognition

'General' cognitive ability or 'intelligence' refers to our ability to reason, learn and solve problems and is measured based on performance on tests of processing speed, vocabulary size, abstract verbal and non-verbal reasoning, and visuo-spatial skills. These scores are aggregated to yield a general ability score or statistically reduced into a single factor or component referred to as Spearman's '*g*' [9, 6]. Typically, a principal components analysis of individual subtests yields a single factor that explains ~50% of variance in measures used, reflecting the strong correlation usually observed between these cognitive tasks. Combining data from multiple sources shows that *g* is a robust value, valid in both western and non-western countries [10, 11].

Notwithstanding the moderate correlations observed between many cognitive tasks, several measurement issues exist. These include low test/retest reliability for some aspects of cognition, [12, 13], a bewildering array of different measures of the same domain and even multiple versions of the same test, all of which complicates attempts to combine data from different groups to achieve the sample sizes required for genomic studies. Even where the same measures have been collected in very large population-based cohorts such as the UK Biobank, the use of shorter cognitive tests within a larger battery of health relevant tasks have led to issues of task validity [14, 15].

Yet another issue of phenotypic complexity in large-scale studies relates to the use not of cognitive tests *per se* but to the use of proxy measures of cognitive ability. Given the need to combine different datasets to increase sample size to boost power for gene discovery and the lack of comprehensive cognitive data in these datasets, some readily available proxy phenotypes have been used including years of education (YOE) and educational attainment (EA). Based on samples of >70,000 English children, the correlation between EA and *g* was observed to be 0.81 [16]. Recent analyses in very large data sets have shown this correlation to be closer to 0.7[17].

Genome-wide association studies (GWAS)

Early genetic studies of cognition focused on 'candidate' genes selected on the basis of their hypothesized biological importance to illness risk. However, a failure to replicate the findings from these studies, together with the emergence of genome-wide approaches to gene discovery in the past ten years have meant that a majority of recent discoveries in both cognitive and psychiatric genetics have come via GWAS. A major initial challenge in adopting this approach was the limited sample sizes of available cohorts, which hindered identification of genome-wide significant results in early GWAS of cognitive phenotypes [18]. In order to boost power for genetic studies, several consortia were formed to pool sample resources to yield more significant outcomes. In 2015 the Cohorts for Heart and Ageing Research in Genomic Epidemiology (CHARGE) consortium combined data from 31 cohorts (n=53,949) and performed a meta-analysis of GWAS using a general cognitive factor derived from principle component analysis of several tests [19]. This analysis identified three loci, on chromosomes 6, 14 and 19, as relevant to cognitive processes [19]. A further analysis by the Cognitive Genomics Consortium (COGENT) combined 21 cohorts

(n=35,298) and confirmed the findings of the CHARGE study as well as identifying two more significant loci on chromosomes 1 and 2 [18]. This study also compared the top SNPs from larger EA studies (n = 164) and found 31 SNPs that were significantly associated with EA in other studies that were also nominally significant in this study; all had the same direction of effect showing a robust genetic correlation between EA and cognition.

The UK Biobank project was initiated to generate a very large dataset based on the UK population where data was collected on over 500,000 people [20]. Initially, genotypic data was released for ~150,000 individuals in May 2015 and was used in combination with existing data in a number of cognition GWAS that confirmed previous findings and uncovered more associated loci [14, 17, 21, 22]. The full data set on >500,000 individuals was released in July 2017 and has a proved a "game-changer" in GWAS of cognition function by facilitating studies with samples sizes of >100,000 individuals that have identified hundreds of independent associated loci. Study of the combined CHARGE, COGENT and UK Biobank cognitive and genetic datasets (n=300,486 participants) have identified 146 genome-wide significant loci and 709 genes associated with general cognitive function [23**]. Associated genes show enriched expression in most brain regions with strongest signals in the cerebellum and cortex and *in silico* biological investigations of these genes points to processes such as *neurogenesis*, *regulation of nervous system development* and *neuron differentiation* being affected. A second study based on COGENT and UK Biobank samples plus other samples (n = 267,867 participants) and published around the same time, found a total of 205 loci (implicating 1,016 genes) to be associated with intelligence [24**]. Analysis of biological processes implicated by these associated found the pathways involving *regulation of nervous system development* and *central nervous system neuron differentiation* to be enriched for associated genes, plus *regulation of synapse structure or activity* was significantly enriched too. Beyond enriched expression of associated genes in multiple brain regions, single cell analysis identified the most enriched cell types for genes associated with intelligence to be medium spiny neurons (striatum), CA1 pyramidal neurons (hippocampus) and pyramidal neurons (somatosensory cortex).

In addition to data from publicly funded biobanks, commercial companies such as 23andMe have also collaborated in cognitive genomic research [25]. In the largest study to date on EA, Lee et al. combined data from 71 cohorts to yield a sample size of 1,131,881individuals, of which 365,538 samples were provided by 23andMe [7]. This analysis identified 1,271 lead SNPs that were

independently genome-wide significant, again demonstrating the positive correlation between sample sizes, and number of variants identified (see Figure 1). Lee et al. used multi-trait analysis of GWAS (MTAG), an approach that exploits the phenotypic and genetic correlations between different phenotypes (e.g. cognitive ones) to increase statistical power [26]. By combining GWAS results from studies of EA, cognitive performance, and mathematical ability (for a total n=1,311,438), Lee et al were able to increase their number of genome-wide significant loci to 1,624 (n=1,311,438). Biological annotation analysis suggested that genes near to these SNPs are strongly enriched for expression in the central nervous system. These genes show elevated expression in the prenatal brain, where they are involved in many developmental processes, but also have high expression in the postnatal brain where genes were involved in nearly all levels of neuron-to-neuron communication and synaptic plasticity. Of note, while neurons were strongly enriched for EA-associated genes, astrocytes and oligodendrocytes were not, leading the authors to conclude that cognitive variation was not associated with genetic differences in myelin related axonal transmission speeds [7]. This conclusion contrasted with findings from a MTAG study by Hill et al [27*] that combined GWAS of intelligence [17] with EA [14] (n=248,482) to identified 187 genetic loci associated with intelligence. Biological annotation analysis showed associated genes to be enriched in a number of processes including neurogenesis, synaptic plasticity, cell development and myelination, specifically oligodendrocyte differentiation [27]. The disagreement between these two studies suggests a need for further studies to clarify whether the genetic architecture of cognitive implicates white matter microstructure and oligodendrocytes function.

A polygenic score (PGS) or polygenic risk score (PRS) is a statistic measuring an individual's genetic 'loading' for variability in a trait (e.g. cognitive function) or risk of illness (e.g. schizophrenia) [45]. Using GWAS results, a PGS is a count of the number of common associated alleles carried by an individual, weighted by the strength of the alleleic associations with the disorder or trait. PGS based on the GWAS above can explain 11–13% of the variance in educational attainment and 7–10% of the variance in cognitive performance in independent samples [7]. Despite the major advances that these studies represent, this suggests that a significant gap remains between the overall heritability for cognition estimated from twin studies and SNP-based heritability for cognition (i.e. the contribution of common SNPs that can be analyzed by GWAS), reported to be 0.19 for general intelligence [24]. This missing heritability is likely due to

a variety of factors, including rare variants, gene x gene (GxG) interactions (epistasis) and gene x environment (GxE) interactions [6].

Rare variants:

Copy number variants (CNVs) are structural variants that were originally described as >1 kbp sections of DNA that can be present in a human genome at a different copy number to the expected two copies in the reference genome. These can be deletions, duplications, inversions or other complex rearrangements, and can range in frequency, but it is those that are rare that have been of most interest in the study of complex phenotypes [28-30]. Recent technological advancement of comparative genomic hybridisation and high-throughput next generation sequencing has led to an improvement in the sensitivity of detection of CNVs resulting in the redefinition of their size to >50 bp [31]. An assembly-based approach to sequencing data from two haploid genomes identified over 460,000 variants from 2bp to 28kbp. Only 10% of these variants were detected in an analysis of the 1000 Genomes Project, highlighting that structural variants have been under-called and under-studied in human genomics [32]. Structural variants contribute to genetic diversity [33] and their important contribution to the genetic variability of cognition is now recognized [29].

CNVs have been associated with disruption of cognitive development leading to intellectual disabilities and other neurodevelopmental disorders [34*]. CNVs associated with these disorders may have incomplete penetrance in a population and apparently healthy adults may carry some of the CNVs associated with these disorders without displaying symptoms [30]. A study based on the reasonably homogenous Icelandic population showed that incomplete penetrance of pathogenic CNVs for autism and schizophrenia was associated with decreased cognitive performance in the healthy population and that individual CNVs affected different cognitive domains [35*]. Examination of non-pathogenic deletions based on children from the Saguenay Youth Study (n = 1,983) and the IMAGEN consortium (n = 2,090) found that non-pathogenic deletions were associated with decreased IQ and suggested that IQ was linked to haploinsufficiency of most of the coding genome [34, 36, 37].

Thirty-three CNVs associated with risk of neurodevelopmental disorders were examined for association with cognitive performance in the UK Biobank (n = 420,247) using seven cognitive

measures. Twenty-four of the 33 CNVS were associated with reduced cognitive performance in healthy carriers and these CNVs also showed an association with reduced educational attainment and income. In addition, all 12 of the CNVs associated with schizophrenia have been associated with reduced cognitive function in healthy adults [38]. In comparison to healthy non-carriers, healthy individuals who carried at least one of the 12 copy number variants associated with schizophrenia showed reduced brain volumes in the hippocampus, nucleus accumbens and thalamus, suggesting a mediation role for hippocampal and thalamic volumes in cognitive ability [39].

Disruptive (loss-of-function) and damaging (missense) rare and ultra-rare single nucleotide variants (SNVs) in highly constrained (HC) genes, i.e. genes under negative selection, are associated with neurocognitive disorders but are also found in the healthy population where they are associated with decreased EA. In a sample of 14,133 individuals, carrying either a disruptive or damaging SNV in a HC gene was associated on average with a reduction in years of education of 2.9-3.1 months [40**]. Each additional disruptive SNV reduced the chance of going to college by on average 14%. This effect of ultra-rare disruptive and damaging SNVs on EA more than doubled when considering HC genes that are highly expressed in the brain.

In a novel approach to explaining the missing heritability in genetic studies on cognition, Hill et al, examined the high level of linkage disequilibrium found in members of the same family in the Generation Scotland family cohort (N = 20,000) [41*, 42]. This analysis using a tool based on a genome-based restricted maximum, GREML-KIN, measures both the variance explained by the genetic effects clustered in families and common SNPs and was replicated in unrelated individuals [43, 44]. Results showed that for general cognitive ability, genetic effects explained 54% of phenotypic variation, of which 31% was explained by pedigree-associated variants (which include rare variants, CNVs and structural variants) and 23% by common variants. These results are similar to heritability levels found in previous twin studies [1]. Overall these findings show that most of the pedigree variants associated with cognition were rare with allele frequencies between 0.001 and 0.01 and current genotyping platforms do not sufficiently tag these variations.

Gene by environment interactions:

Hasan and Afzal [46*] argue that to fully understand cognition, environmental effects need to be explored. Interplay between nature and nurture has been found through twin and adoption studies, with environment and genetics observed to co-vary in a manner whereby genetic make-up can determine environmental conditions. They propose that the study of candidate genes arising from next generation sequencing should include environmental parameters [46]. While PGS can explain 10% of the variation in educational attainment some of this is indirect and is explained by passive gene-environment correlation where parents and other relatives provide a rearing environment that is associated with the parental genotype [47, 48]. A recent study shows that PGS for intelligence and EA had a 60% greater predictive value when tested between families as opposed to within families. This difference disappears when socio-economic class is controlled [49]. In a further study of adopted individuals in the UK Biobank (n=6311) it was found that PGS generated from mainly non-adoptive individuals was only 50% as predictive of YOE in adoptees when compared with non-adoptive individuals and conclude that parental influences affect YOE. It was also found that individuals who have a low PGS for YOE spent longer in education if adopted supporting the gene-environment correlations theory [47]. These studies support the inclusion of environmental effects in genetic studies of cognition.

Genetic correlations between cognition and schizophrenia.

An overlap between the genetic variation associated with cognitive ability and genetic variation associated with SZ risk has long been hypothesized. Toulopoulou et al. [49], using latent factor modeling, estimated the shared genetic effects between cognition (measured in terms of memory and IQ) and risk for SZ[50]. They reported a high degree of overlap in genetic effects (the highest being for IQ) with environmental effects accounting for little phenotypic correlation. Fowler et al. however disputed the strength of this genetic overlap [51]. Using a Swedish population cohort, a much smaller overlap (~7%) in genetic variants associated with the two phenotypes was observed. Fowler et al argue that this difference related to a difference in sample selection – with studies based on patient recruitment overestimating genetic correlation between cognition and illness risk by comparison with studies based on unselected (population) cohorts. More recently, we were able to undertake a similar analysis, this time based on SNP heritability rather than general estimates

of shared genetic correlation. To assess the genetic overlap between schizophrenia liability and cognitive functions, we summarized data from the Maudsley Twin and Family Studies and the Schizophrenia Twins and Relatives Consortium (STAR Consortium), a series of studies of twins and other family members concordant or discordant for schizophrenia. Evidence of substantial genetic overlap was observed between cognitive phenotypes and schizophrenia liability (average rg = -.58; SD=0.22), although the estimates ranged widely, possibly due to the small samples involved [52].

Correlation between cognitive performance and illness liability

PGS scores for schizophrenia have been shown to be weakly associated with IQ and cognition in population samples [53-55]). For example, in a study of the ALSPAC cohort, schizophrenia PGS was associated with lower performance IQ (P= .001) and lower full IQ (P= .013) [54]. A PGS for IQ was associated with increased risk for schizophrenia (P= 3.56E-04). Bivariate genome-wide complex trait analysis (GCTA) revealed moderate genetic correlation between schizophrenia and both performance IQ (rG=-.379,P= 6.62E-05) and full IQ (rG= -.202,P= 5.00E-03), with approximately 14% of the genetic component of schizophrenia shared with that for performance IQ. Similarly, PGS for cognition was associated with severity of negative, but not positive symptoms in those with schizophrenia [56]. Finally, in a GWAS of cognition in a large sample of >3,000 patients with schizophrenia, Richards et al sought to determine whether PGS for either illness risk, educational attainment, or IQ could be used to explain to a significant degree of variation in cognitive performance in patients with SZ. PRS for both population IQ (P = 4.39 × 10–28) and EA (P = 1.27 × 10–26) were positively correlated with cognition in those with schizophrenia. In contrast, there was no association between cognition in schizophrenia cases and PRS for schizophrenia (P = .39), bipolar disorder (P = .51), or major depressive disorder (P = .49) [57**].

Taking these polygenic studies of the relationship between variation in cognition and SZ risk, in the context of the previous heritability studies, it is clear that the genetic architecture of illness risk and normal variation in cognition overlap to at least a modest degree. This is perhaps unsurprising; given the likelihood that a distributed network of brain structures/functions contribute to both phenotypes, it would be counterintuitive if at least some of the underpinning biological processes

did not overlap. At the same time, the modest power of PGS based on GWAS of either cognition or illness risk to explain variation in the other also highlights the large degree of genetic non-overlap between these phenotypes. In terms of support for the original idea of endophenotyes – quantifiable traits whose study may help with parsing the genetic architecture of a broader (more complex) illness phenotype, this evidence is not compelling. Instead of aiding gene discovery, the real value of endophenotypes may be to identify the functional significance of already identified variants at the level of cognitive performance and other aspects of illness that predict level of disability (e.g. employment status) [58].

In terms of the direction of genetic correlation between illness susceptibility on the one hand and cognitive performance and EA on the other, some interesting differences between diagnoses have emerged. Genetic correlational analysis using GWAS data suggests a significant, negative correlation between intelligence and susceptibility to several neurodevelopmental and psychiatric disorders (e.g. schizophrenia, attention deficient hyperactivity disorder (ADHD) and major depressive disorder (MDD)), but a positive correlation between intelligence and autism spectrum disorder (ASD; see Figure 2). Correlational analysis between EA and these disorders is broadly similar, except that EA is not negatively correlated with schizophrenia, as might be expected, given the strong positive phenotypic and genetic correlation between intelligence and EA.

To understanding this puzzling discordance, Lam et al. recently teased apart the genetic findings for EA and schizophrenia to identify a subset of variants associated in the expected "concordant" direction, i.e., alleles associated with both schizophrenia risk and lower EA, but also a subset of variants that demonstrated the counterintuitive "discordant" relationship, i.e. alleles associated with schizophrenia risk but higher EA[8**]. The concordant alleles mapped to genes involved in early neurodevelopmental pathways, consistent with evidence that cognitive deficits are often present early in life before onset of schizophrenia, whereas the discordant alleles mapped to genes that functioned in adulthood synaptic pruning pathways. The authors suggest that the latter may reflect the importance of efficient synaptic pruning for academic ability on the one hand, but a liability excessive pruning on the other, contributing to schizophrenia [8].

Cognitive decline and ageing

Improved life expectancy and declining birth rates has led to an increasing percentage of the population that is greater than 60 years of age. According to the World Health Organisation the portion of the population over 60 will increase from 12% to 22% by 2050 and will reach 30% in more developed countries [59]. Cognitive decline associated with age results in increased difficulty in performing tasks that require memory or rapid information processing and can have an increasingly detrimental effect on quality of life [60]. A meta-analysis of rates of variation in changing cognitive ability showed that variation increased with age and was consistent across different cognitive domains. Measuring this variance in rates of change of cognitive decline requires strong longitudinal data and complex statistical methods [61].

The biological contribution to variation in rates of cognitive decline is evolving. In a recent opinion piece by Cabeza and colleagues [62**], three biological mechanisms are proposed to control cognitive decline in healthy ageing, and these are reserve, maintenance and compensation. Reserve is discussed in terms of brain reserve, which can be described and individual differences in structural characteristics, such as quantities of neurons and synapses, and cognitive reserve, which refers to the adaptability of cognitive processes [63]. Maintenance involves resistance to neural decline and repair of damage and compensation is when alternative neuronal pathways are recruited. The complexity of these processes infers a highly polygenic genetic contribution and suitable data sets are needed to examine the genes and biological pathways involved. A recent study used the Lothian Birth Control cohort of 1935 at four different time points between the age of 70 and 79 to measure the association of changes in g with fourteen robustly generated PGS. These PGS included EA, grip strength, schizophrenia, Alzheimer's and other health related PGS. The researchers conclude that the predictive power of PGS in not yet sensitive enough to explain the variance in cognitive decline [64*].

Genetic variation accounts for 40 to 50% of cognitive performance of older adults and 24% of the variability of cognitive change over the life span [22, 65]. Some studies show an association between genetic variants and age-related cognitive decline, yet they only explained a fraction of the phenotypic variability. In addition, many of the studies failed to replicate due to difference in cognitive measurements and other methodological issues and lack of control of participant characteristics [60]. A meta-analysis of studies on cognitive decline concluded that major

improvements were needed in research methods, in particular the use of standardized procedures across studies [66]. Interestingly, recent research has shown that neurogenesis occurs in the dental gyrus of the adult hippocampus into the 9th decade of life and healthy individuals without neurodegenerative conditions show preserved neurogenesis. The authors propose that individual resilience leads to variation in rates of neurogenesis and differing rates in cognitive decline [67].

While research into the effects of environmental factors have shown the importance of cardiovascular health, social involvement and diet on healthy ageing, our assessment of the understanding of the genomics involved in cognitive decline is hampered by the lack of strong cognitive measures coupled with large genetic data sets. As yet we do not know whether cognitive decline is genetically influenced by genes associated with general intelligence or if genes that regulate other biological processes are involved.

Cognitive genomics in the future:

The comparison of the recent release of whole exome sequencing data from the UK Biobank on ~49,000 individuals and their previously imputed genetic data identified nearly 4 four million coding SNPs and indels per individual, ~7 times higher than that observed in the imputed GWAS data. There was also a 10-fold increase in the identification of loss-of-function variants and loss-of-function variants were found in 97% of autosomal genes. Whole exome sequencing of the remainder of the UK Biobank, which is on-going, and subsequently whole genome sequencing will allow for new analysis of cognition phenotypes using rare genetic variants [68] and may give new insights into the genomics of cognition.

According to Eichler, identifying all the genetic contribution is not just a matter of increasing sample size, as variants are being missed with short read datasets that are aligned to a single reference genome, even when using whole genome sequencing [69]. He argued that more meaningful results will be obtained by diversification of genomic data. Generic research to date on cognition (and other traits) has been almost exclusively confined to samples of individuals of European ancestry. Lee et al. found that their PGS for EA was far less predictive in an African American sample [7]. Eichler proposed that the use of combinations of reference genomes from different populations, that are currently in production should in theory identify the majority of structural variants which have been untested in recent GWAS [33, 69, 70]. It also important that

reference genomes contain representation for African populations to encompass the evolutionary influences on the genome [71].

The use of whole genome sequencing, long-read and ultra-long-read sequencing technology coupled with the development of bioinformatic tools and the further extrapolation of the biological association of over 1000 lead SNPs identified by Lee et al. for EA and others should generate a great insight into cognitive processes. In addition, further development of tools and research approaches that gives us a greater insight into the interplay of the environment and genomics in healthy and psychiatric cohorts will add to our understanding of the critical biological pathways involved in neurocognition.

Conclusion

There has been considerable and rapid progress in identifying the genetic architecture of cognitive performance in recent years. This has been aided, perhaps equally, by both improvements in genomic methods, and the increasing availability of data due to data sharing and cooperation. This progress has resulted in a clearer picture of the highly polygenic basis of cognitive performance, and of the multiple biological processes involved. The contribution of common genetic variation to explaining variation in cognitive performance is clear. The contribution of rar(er) variants both to intellectual disability but also cognitive variation in the general population is also clear. The overlap, but also the discontinuity, between the polygenic variation underpinning cognitive function, illness risk, and cognitive decline is also beginning to come into view. With this clarity, the need for further development of analytical and bioinformatic approaches to understand the biology of these processes has become clearly visible. In particular, the need to more clearly identify the myriad biological pathways underpinning cognitive function is underlined (e.g. the contribution of oligodendroglial-related genetic variation to cognitive performance). Similarly, the need to model how genetic variation - both common and rare - interacts with environmental factors to predict cognitive performance is also a clear priority. Given the progress made in the past 5-10 years, furthering these objectives continues to hold significant promise for understanding cognitive ability, and for treating cognitively related disability.

References

1. Plomin R, Deary IJ. Genetics and intelligence differences: five special findings. Molecular Psychiatry. 2014;20:98. doi:10.1038/mp.2014.105.

2. Plomin R, Spinath F. Intelligence: Genetics, Genes, and Genomics. Journal of personality and social psychology. 2004;86:112-29. doi:10.1037/0022-3514.86.1.112.

3. Plomin R, DeFries, J. C., McClearn, G. E., & McGuffin, P. . Behavioral genetics. 4th ed. New York: Worth Publishers; 2001.

4. Briley DA, Tucker-Drob EM. Explaining the increasing heritability of cognitive ability across development: a meta-analysis of longitudinal twin and adoption studies. Psychol Sci. 2013;24(9):1704-13. doi:10.1177/0956797613478618.

5. Ramus F. Genes, brain, and cognition: A roadmap for the cognitive scientist. Cognition. 2006;101(2):247-69. doi:https://doi.org/10.1016/j.cognition.2006.04.003.

6. *Plomin R, von Stumm S. The new genetics of intelligence. Nature Reviews Genetics. 2018;19:148. doi:10.1038/nrg.2017.104. **This recent review article discusses the benefits of polygenic scores in intelligence research**

7. **Lee JJ, Wedow R, Okbay A, Kong E, Maghzian O, Zacher M et al. Gene discovery and polygenic prediction from a genome-wide association study of educational attainment in 1.1 million individuals. Nature Genetics. 2018;50(8):1112-21. doi:10.1038/s41588-018-0147-3. **This study is the largest GWAS study to date on educational attainment using both public and commerically available data and highlights the role of genes involved in the prenatal brain as well post natal development.**

8. *Lam M, Hill WD, Trampush JW, Yu J, Knowles E, Davies G et al. Pleiotropic Meta-Analysis of Cognition, Education, and Schizophrenia Differentiates Roles of Early Neurodevelopmental and Adult Synaptic Pathways. bioRxiv. 2019:519967. doi:10.1101/519967. **This paper examines the pleiotropic nature of GWAS findings on intelligence and psychiatric disorders.**

9. Deary IJ, Penke L, Johnson W. The neuroscience of human intelligence differences. Nature Reviews Neuroscience. 2010;11:201. doi:10.1038/nrn2793.

10. Johnson W, Nijenhuis Jt, Bouchard TJ. Still just 1 g: Consistent results from five test batteries. Intelligence. 2008;36(1):81-95. doi:https://doi.org/10.1016/j.intell.2007.06.001.

11. Warne RT, Burningham C. Spearman's g found in 31 non-Western nations: Strong evidence that g is a universal phenomenon. Psychological Bulletin. 2019;145(3):237-72. doi:10.1037/bul0000184.

12. Hedge C, Powell G, Sumner P. The reliability paradox: Why robust cognitive tasks do not produce reliable individual differences. Behavior Research Methods. 2018;50(3):1166-86. doi:10.3758/s13428-017-0935-1.

13. De Schryver M, Hughes S, Rosseel Y, De Houwer J. Unreliable Yet Still Replicable: A Comment on LeBel and Paunonen (2011). Frontiers in Psychology. 2016;6(2039). doi:10.3389/fpsyg.2015.02039.

14. Okbay A, Beauchamp JP, Fontana MA, Lee JJ, Pers TH, Rietveld CA et al. Genome-wide association study identifies 74 loci associated with educational attainment. Nature. 2016;533(7604):539-42. doi:10.1038/nature17671.

15. Fawns-Ritchie C, Deary IJ. Reliability and validity of the UK Biobank cognitive tests. medRxiv. 2019:19002204. doi:10.1101/19002204.

16. Deary IJ, Strand S, Smith P, Fernandes C. Intelligence and educational achievement. Intelligence. 2007;35(1):13-21. doi:https://doi.org/10.1016/j.intell.2006.02.001.

17. Sniekers S, Stringer S, Watanabe K, Jansen PR, Coleman JRI, Krapohl E et al. Genome-wide association meta-analysis of 78,308 individuals identifies new loci and genes influencing human intelligence. Nat Genet. 2017;49(7):1107-12. doi:10.1038/ng.3869.

18. Trampush JW, Yang ML, Yu J, Knowles E, Davies G, Liewald DC et al. GWAS meta-analysis reveals novel loci and genetic correlates for general cognitive function: a report from the COGENT consortium. Mol Psychiatry. 2017;22(3):336-45. doi:10.1038/mp.2016.244.

19. Davies G, Armstrong N, Bis JC, Bressler J, Chouraki V, Giddaluru S et al. Genetic contributions to variation in general cognitive function: a meta-analysis of genome-wide association studies in the CHARGE consortium (N=53949). Molecular psychiatry. 2015;20(2):183-92. doi:10.1038/mp.2014.188.

20. Sudlow C, Gallacher J, Allen N, Beral V, Burton P, Danesh J et al. UK biobank: an open access resource for identifying the causes of a wide range of complex diseases of middle and old age. PLoS Med. 2015;12(3):e1001779. doi:10.1371/journal.pmed.1001779.

21. Hill WD, Davies G, Harris SE, Hagenaars SP, Liewald DC, Penke L et al. Molecular genetic aetiology of general cognitive function is enriched in evolutionarily conserved regions. Translational psychiatry. 2016;6(12):e980. doi:10.1038/tp.2016.246.

22. Davies G, Marioni RE, Liewald DC, Hill WD, Hagenaars SP, Harris SE et al. Genome-wide association study of cognitive functions and educational attainment in UK Biobank (N=112 151). Mol Psychiatry. 2016;21(6):758-67. doi:10.1038/mp.2016.45.

23. **Davies G, Lam M, Harris SE, Trampush JW, Luciano M, Hill WD et al. Study of 300,486 individuals identifies 148 independent genetic loci influencing general cognitive function. Nature Communications. 2018;9(1):2098. doi:10.1038/s41467-018-04362-x. **This study demonstrates the power of sample size to GWAS findings by the addition of the UK biobank data. It confirms previous findings and identifies new loci associated with neuronial communication.**

24. **Savage JE, Jansen PR, Stringer S, Watanabe K, Bryois J, de Leeuw CA et al. Genome-wide association meta-analysis in 269,867 individuals identifies new genetic and functional links to intelligence. Nature Genetics. 2018;50(7):912-9. doi:10.1038/s41588-018-0152-6. **Again using the UK biobank and other data, this study had identified the largest number of association snps for IQ to date, and reports a bio-informatic analysis of these findings.**

25. Eriksson N, Macpherson JM, Tung JY, Hon LS, Naughton B, Saxonov S et al. Web-Based, Participant-Driven Studies Yield Novel Genetic Associations for Common Traits. PLoS Genetics. 2010;6(6):1-20. doi:10.1371/journal.pgen.1000993.

26. Turley P, Walters RK, Maghzian O, Okbay A, Lee JJ, Fontana MA et al. Multi-trait analysis of genome-wide association summary statistics using MTAG. Nature Genetics. 2018;50(2):229-37. doi:10.1038/s41588-017-0009-4.

27. *Hill WD, Marioni RE, Maghzian O, Ritchie SJ, Hagenaars SP, McIntosh AM et al. A combined analysis of genetically correlated traits identifies 187 loci and a role for neurogenesis and myelination in intelligence. Mol Psychiatry. 2018. doi:10.1038/s41380-017-0001-5. **Hill et al demonstrate the additive effects of multi-trait anlaysis and its utility to generate further findings.**

28. Lee C, Scherer SW. The clinical context of copy number variation in the human genome. Expert Rev Mol Med. 2010;12:e8. doi:10.1017/S1462399410001390.

29. Feuk L, Carson AR, Scherer SW. Structural variation in the human genome. Nature Reviews Genetics. 2006;7(2):85-97. doi:10.1038/nrg1767.

30. Kendall KM, Rees E, Escott-Price V, Einon M, Thomas R, Hewitt J et al. Cognitive Performance Among Carriers of Pathogenic Copy Number Variants: Analysis of 152,000 UK Biobank Subjects. Biological Psychiatry. 2017;82(2):103-10. doi:https://doi.org/10.1016/j.biopsych.2016.08.014.

31. Nowakowska B. Clinical interpretation of copy number variants in the human genome. J Appl Genet. 2017;58(4):449-57. doi:10.1007/s13353-017-0407-4.

32. Huddleston J, Chaisson MJP, Steinberg KM, Warren W, Hoekzema K, Gordon D et al. Discovery and genotyping of structural variation from long-read haploid genome sequence data. Genome Research. 2017;27(5):677-85.

33. Chiang C, Scott AJ, Davis JR, Tsang EK, Li X, Kim Y et al. The impact of structural variation on human gene expression. Nature Genetics. 2017;49:692. doi:10.1038/ng.3834

34. *Huguet G, Schramm C, Douard E, Jiang L, Labbe A, Tihy F et al. Measuring and Estimating the Effect Sizes of Copy Number Variants on General Intelligence in Community-Based Samples. JAMA Psychiatry. 2018;75(5):447-57. doi:10.1001/jamapsychiatry.2018.0039.**This study presents a framework for examining the effects of copy number variants on general cognition.**

35. *Stefansson H, Meyer-Lindenberg A, Steinberg S, Magnusdottir B, Morgen K, Arnarsdottir S et al. CNVs conferring risk of autism or schizophrenia affect cognition in controls. Nature. 2013;505:361. doi:10.1038/nature12818 **The study of population isolates can highlight the role of CNVs in cognition.**

36. Schumann G, Loth E, Banaschewski T, Barbot A, Barker G, Büchel C et al. The IMAGEN study: reinforcement-related behaviour in normal brain function and psychopathology. Molecular Psychiatry. 2010;15:1128. doi:10.1038/mp.2010.4

37. Pausova Z, Paus T, Abrahamowicz M, Bernard M, Gaudet D, Leonard G et al. Cohort Profile: The Saguenay Youth Study (SYS). International journal of epidemiology. 2017;46(2):e19-e. doi:10.1093/ije/dyw023.

38. Kendall KM, Bracher-Smith M, Fitzpatrick H, Lynham A, Rees E, Escott-Price V et al. Cognitive performance and functional outcomes of carriers of pathogenic copy number variants: analysis of the UK Biobank. The British Journal of Psychiatry. 2019;214(5):297-304. doi:10.1192/bjp.2018.301.

39. Warland A, Kendall KM, Rees E, Kirov G, Caseras X. Schizophrenia-associated genomic copy number variants and subcortical brain volumes in the UK Biobank. Molecular Psychiatry. 2019. doi:10.1038/s41380-019-0355-y.

40. **Ganna A, Genovese G, Howrigan DP, Byrnes A, Kurki MI, Zekavat SM et al. Ultra-rare disruptive and damaging mutations influence educational attainment in the general population. Nature Neuroscience. 2016;19:1563. doi:10.1038/nn.4404 **Important paper highlighting the role of rare mutations in intelligence using EA as a proxy measure.**

41. *Hill WD, Arslan RC, Xia C, Luciano M, Amador C, Navarro P et al. Genomic analysis of family data reveals additional genetic effects on intelligence and personality. Molecular Psychiatry. 2018;23(12):2347-62. doi:10.1038/s41380-017-0005-1. **This study shows how using family based data can resolve some of the missing hertiability in cognition.**

42. Smith BH, Campbell H, Blackwood D, Connell J, Connor M, Deary IJ et al. Generation Scotland: the Scottish Family Health Study; a new resource for researching genes and heritability. BMC Medical Genetics. 2006;7(1):74. doi:10.1186/1471-2350-7-74.

43. Zaitlen N, Kraft P, Patterson N, Pasaniuc B, Bhatia G, Pollack S et al. Using Extended Genealogy to Estimate Components of Heritability for 23 Quantitative and Dichotomous Traits. PLOS Genetics. 2013;9(5):e1003520. doi:10.1371/journal.pgen.1003520.

44. Xia C, Amador C, Huffman J, Trochet H, Campbell A, Porteous D et al. Pedigree- and SNP-Associated Genetics and Recent Environment are the Major Contributors to Anthropometric and Cardiometabolic Trait Variation. PLOS Genetics. 2016;12(2):e1005804. doi:10.1371/journal.pgen.1005804.

45. Shafee R, Nanda P, Padmanabhan JL, Tandon N, Alliey-Rodriguez N, Kalapurakkel S et al. Polygenic risk for schizophrenia and measured domains of cognition in individuals with psychosis and controls. Translational psychiatry. 2018;8(1):78. doi:10.1038/s41398-018-0124-8.

46. *Hasan A, Afzal M. Gene and environment interplay in cognition: Evidence from twin and molecular studies, future directions and suggestions for effective candidate gene x environment (cGxE) research. Multiple Sclerosis and Related Disorders. 2019;33:121-30. doi:https://doi.org/10.1016/j.msard.2019.05.005. **This article reviews the interplay between the enviroment and cognition and explores the need to examine these effects in future studies.**

47. Cheesman R, Hunjan A, Coleman JRI, Ahmadzadeh Y, Plomin R, McAdams TA et al. Comparison of adopted and non-adopted individuals reveals gene-environment interplay for education in the UK Biobank. bioRxiv. 2019:707695. doi:10.1101/707695.

48. Kong A, Thorleifsson G, Frigge ML, Vilhjalmsson BJ, Young AI, Thorgeirsson TE et al. The nature of nurture: Effects of parental genotypes. Science. 2018;359(6374):424. doi:10.1126/science.aan6877.

49. Selzam S, Ritchie SJ, Pingault J-B, Reynolds CA, O'Reilly PF, Plomin R. Comparing within- and between-family polygenic score prediction. bioRxiv. 2019:605006. doi:10.1101/605006.

50. Toulopoulou T, Goldberg T, Rebollo Mesa I, Picchioni M, Rijsdijk F, Stahl D et al. Impaired Intellect and Memory A Missing Link Between Genetic Risk and Schizophrenia? Archives of general psychiatry. 2010;67:905-13. doi:10.1001/archgenpsychiatry.2010.99.

51. Fowler D, Hodgekins J, Garety P, Freeman D, Kuipers E, Dunn G et al. Negative cognition, depressed mood, and paranoia: a longitudinal pathway analysis using structural equation modeling. Schizophrenia bulletin. 2012;38(5):1063-73. doi:10.1093/schbul/sbr019.

52. Blokland GAM, Del Re EC, Mesholam-Gately RI, Jovicich J, Trampush JW, Keshavan MS et al. The Genetics of Endophenotypes of Neurofunction to Understand Schizophrenia (GENUS) consortium: A collaborative cognitive and neuroimaging genetics project. Schizophrenia research. 2018;195:306-17. doi:10.1016/j.schres.2017.09.024.

53. Lencz T, Knowles E, Davies G, Guha S, Liewald DC, Starr JM et al. Molecular genetic evidence for overlap between general cognitive ability and risk for schizophrenia: a report from the Cognitive Genomics consorTium (COGENT). Molecular psychiatry. 2014;19(2):168-74. doi:10.1038/mp.2013.166.

54. Hubbard L, Tansey KE, Rai D, Jones P, Ripke S, Chambert KD et al. Evidence of Common Genetic Overlap Between Schizophrenia and Cognition. Schizophrenia Bulletin. 2015;42(3):832-42. doi:10.1093/schbul/sbv168.

55. Hagenaars SP, Harris SE, Davies G, Hill WD, Liewald DC, Ritchie SJ et al. Shared genetic aetiology between cognitive functions and physical and mental health in UK

Biobank (N=112 151) and 24 GWAS consortia. Mol Psychiatry. 2016;21(11):1624-32. doi:10.1038/mp.2015.225.

56. Fanous AH, Zhou B, Aggen SH, Bergen SE, Amdur RL, Duan J et al. Genome-wide association study of clinical dimensions of schizophrenia: polygenic effect on disorganized symptoms. Am J Psychiatry. 2012;169(12):1309-17. doi:10.1176/appi.ajp.2012.12020218.

57. **Richards AL, Pardiñas AF, Frizzati A, Tansey KE, Lynham AJ, Holmans P et al. The Relationship Between Polygenic Risk Scores and Cognition in Schizophrenia. Schizophrenia Bulletin. 2019. doi:10.1093/schbul/sbz061. **This study provides a current estimate of the genetic correlation between cognitive function and schizophrenia susceptibility.**

58. Green MF. Impact of cognitive and social cognitive impairment on functional outcomes in patients with schizophrenia. J Clin Psychiatry. 2016;77 Suppl 2:8-11. doi:10.4088/jcp.14074su1c.02.

59. World Health O. World report on ageing and health. Geneva: World Health Organization; 2015.

60. Andrews SJ, Das D, Cherbuin N, Anstey KJ, Easteal S. Association of genetic risk factors with cognitive decline: the PATH through life project. Neurobiol Aging. 2016;41:150-8. doi:10.1016/j.neurobiolaging.2016.02.016.

61. Tucker-Drob EM, Brandmaier AM, Lindenberger U. Coupled cognitive changes in adulthood: A meta-analysis. Psychol Bull. 2019. doi:10.1037/bul0000179.

62. **Cabeza R, Albert M, Belleville S, Craik FIM, Duarte A, Grady CL et al. Maintenance, reserve and compensation: the cognitive neuroscience of healthy ageing. Nat Rev Neurosci. 2018;19(11):701-10. doi:10.1038/s41583-018-0068-2. **This article presents a clear theortitical model to explain the factors involved in cognitive decline.**

63. Stern Y, Arenaza-Urquijo EM, Bartrés-Faz D, Belleville S, Cantilon M, Chetelat G et al. Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. Alzheimer's & Dementia. 2018. doi:https://doi.org/10.1016/j.jalz.2018.07.219.

64. *Ritchie SJ, Hill WD, Marioni RE, Davies G, Hagenaars SP, Harris SE et al. Polygenic predictors of age-related decline in cognitive ability. Molecular Psychiatry. 2019.

doi:10.1038/s41380-019-0372-x. **This study explores the use of PGS in understanding the genetics of cognitive decline and highlights the need for further work.**

65. Deary IJ, Yang J, Davies G, Harris SE, Tenesa A, Liewald D et al. Genetic contributions to stability and change in intelligence from childhood to old age. Nature. 2012;482(7384):212-5. doi:10.1038/nature10781.

66. Plassman BL, Williams JW, Jr., Burke JR, Holsinger T, Benjamin S. Systematic Review: Factors Associated With Risk for and Possible Prevention of Cognitive Decline in Later Life. Annals of Internal Medicine. 2010;153(3):182-93. doi:10.7326/0003-4819-153-3-201008030-00258.

67. **Boldrini M, Fulmore CA, Tartt AN, Simeon LR, Pavlova I, Poposka V et al. Human Hippocampal Neurogenesis Persists throughout Aging. Cell Stem Cell. 2018;22(4):589-99.e5. doi:10.1016/j.stem.2018.03.015. **This research explores new frontiers in the understanding of cognitive variance in aging and proposes novel concepts involving neurogenesis.**

68. Van Hout CV, Tachmazidou I, Backman JD, Hoffman JX, Ye B, Pandey AK et al. Whole exome sequencing and characterization of coding variation in 49,960 individuals in the UK Biobank. bioRxiv. 2019:572347. doi:10.1101/572347.

69. *Eichler EE. Genetic Variation, Comparative Genomics, and the Diagnosis of Disease. New England Journal of Medicine. 2019;381(1):64-74. doi:10.1056/NEJMra1809315. **Very good review of the challenges facing genetic studies into the future.**

70. McCarthy S, Das S, Kretzschmar W, Delaneau O, Wood AR, Teumer A et al. A reference panel of 64,976 haplotypes for genotype imputation. Nature genetics. 2016;48(10):1279-83. doi:10.1038/ng.3643.

71. McClellan JM, Lehner T, King M-C. Gene Discovery for Complex Traits: Lessons from Africa. Cell. 2017;171(2):261-4. doi:https://doi.org/10.1016/j.cell.2017.09.037.

72. Rietveld CA, Esko T, Davies G, Pers TH, Turley P, Benyamin B et al. Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. Proceedings of the National Academy of Sciences. 2014;111(38):13790. doi:10.1073/pnas.1404623111.

73. Demontis D, Walters RK, Martin J, Mattheisen M, Als TD, Agerbo E et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. Nature Genetics. 2019;51(1):63-75. doi:10.1038/s41588-018-0269-7.

74. Grove J, Ripke S, Als TD, Mattheisen M, Walters RK, Won H et al. Identification of common genetic risk variants for autism spectrum disorder. Nature Genetics. 2019;51(3):431-44. doi:10.1038/s41588-019-0344-8.

75. Ruderfer DM, Ripke S, McQuillin A, Boocock J, Stahl EA, Pavlides JMW et al. Genomic Dissection of Bipolar Disorder and Schizophrenia, Including 28 Subphenotypes. Cell. 2018;173(7):1705-15.e16. doi:https://doi.org/10.1016/j.cell.2018.05.046.

76. Wray NR, Ripke S, Mattheisen M, Trzaskowski M, Byrne EM, Abdellaoui A et al. Genome-wide association analyses identify 44 risk variants and refine the genetic architecture of major depression. Nature Genetics. 2018;50(5):668-81. doi:10.1038/s41588-018-0090-3.

77. Pardiñas AF, Holmans P, Pocklington AJ, Escott-Price V, Ripke S, Carrera N et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. Nature Genetics. 2018;50(3):381-9. doi:10.1038/s41588-018-0059-2.

78. Watanabe K, Stringer S, Frei O, Umićević Mirkov M, de Leeuw C, Polderman TJC et al. A global overview of pleiotropy and genetic architecture in complex traits. Nature Genetics. 2019;51(9):1339-48. doi:10.1038/s41588-019-0481-0.

Figure 1: Plot of number of lead SNPs from GWAS and MTAG analysis of *g* and EA showing increases in the number of significant findings with increasing samples sizes. Each study is numbered on the graph and included for g are (A) Sniekers et al [17], n=18 lead SNPs, (B) Savage et al [24], n=242 lead SNPs, (C) Davies et al [3], n=434 lead SNPs. (D) MTAG analysis by Hill et al [27], combining GWAS of g by Sniekers et al [17] and EA by Okbay et al [14], resulted in 564 lead SNPs. Included for EA are (E) Rietveld et al [72], n=69 lead SNPs, (F) Okbay et al [14], n=74 lead SNPs and (G) Lee et al [7], n=1,271 lead SNPs. MTAG analysis by Lee et al combining GWAS of EA with cognitive performance, self-reported math ability and highest math class taken achieved an increase to n=1,624 lead SNPs.

Figure 2: Genetic correlations between different phenotypes based on the GWAS Atlas (GA; https://atlas.ctglab.nl/ [78]). Phenotypes included are educational attainment (EA) GAID # 4066 [7] and intelligence (g) GAID # 3785 [24] and the psychiatric disorders of attention deficit hyperactivity disorder (ADHD) GAID # 3 [73], autism spectrum disorder (ASD) GAID # 4037 [74], bipolar disorder (BD) GAID # 4039 [75], major depressive disorder (MDD) GAID # 4014 [76] and schizophrenia (SCZ) GAID # 3982 [77] . Figure 2a: Heatmap of genetic correlations between different GWAS. Significant correlations after Bonferroni correction (<0.05) are labelled with "*". Figure 2b: Heatmap of overlapping genome-wide significant genes ($P < 2.5 \times 10^{-6}$) between different GWAS. The number of significant genes per individual GWAS is highlighted in blue.