Disorders of Brain and Mind 2

Edited by

Maria A. Ron and
Trevor W. Robbins

CAMBRIDGE
UNIVERSITY PRESS

PUBLISHED BY THE PRESS SYNDICATE OF THE UNIVERSITY OF CAMBRIDGE
The Pitt Building, Trumpington Street, Cambridge, United Kingdom

CAMBRIDGE UNIVERSITY PRESS
The Edinburgh Building, Cambridge CB2 2RU, UK
40 West 20th Street, New York, NY 10011-4211, USA
477 Williamstown Road, Port Melbourne, VIC 3207, Australia
Ruiz de Alarcón 13, 28014 Madrid, Spain
Dock House, The Waterfront, Cape Town 8001, South Africa

http://www.cambridge.org

First published 2003

Printed in the United Kingdom at the University Press, Cambridge

Typefaces Minion 10.5/14 pt, Formata and Formata BQ *System* LaTeX 2_ε [TB]

A catalogue record for this book is available from the British Library

Library of Congress Cataloguing in Publication data

ISBN 0 521 00456 X paperback

Contents

1

Genes and behaviour: cognitive abilities and disabilities in normal populations

Robert Plomin

Institute of Psychiatry, King's College, London, UK

Introduction

During the past three decades, the behavioural sciences have emerged from an era of strict environmental explanations for differences in behaviour to a more balanced view that recognizes the importance of nature (genetics) as well as nurture (environment). This shift occurred first for behavioural disorders, including rare disorders such as autism (0.001 incidence), more common disorders such as schizophrenia (0.01), and very common disorders such as reading disability (0.05). More recently it has become increasingly accepted that genetic variation contributes importantly to differences among individuals in the normal range of variability as well as for abnormal behaviour. Moreover, many behavioural disorders, especially common ones, may represent the quantitative extreme of the same genetic and environmental factors responsible for variation in the normal range. That is, genetic influence on disorders such as reading disability may not be due to genes specific to the disorder but rather to genes that contribute to the normal range of individual differences in reading ability. This view, known as the quantitative trait locus (QTL) perspective, has important conceptual implications because it implies that some common disorders may not be disorders at all but rather the extremes of normal distributions. This QTL perspective has far-reaching implications for molecular genetics and for neuroscience. If many genes of small effect are involved, it will be much more difficult to find them. It will also be much more difficult to explore the brain mechanisms that mediate genetic effects on behaviour.

These issues are the topic of this chapter, which focuses on cognitive abilities and disabilities. Basic introductions to quantitative genetics (such as twin and adoption designs), molecular genetics and research that uses these genetic methods to investigate behaviour are available elsewhere (Plomin et al. 2001*a*), as are more

detailed discussions of genetic research in neuroscience (Crusio and Gerlai 1999; Pfaff et al. 2000).

The very standard deviation

It is important to begin with a discussion of the different perspectives or levels of analysis used to investigate behaviour because so much follows conceptually as well as methodologically from these differences in perspective (Figure 1.1). Research on cognitive abilities and disabilities focuses on within-species interindividual differences – for example, why some children are reading disabled and others are not. In contrast, textbooks in cognitive neuroscience seldom mention individual differences and concentrate instead on species-universal or species-typical (normative) aspects of cognitive functioning (Gazzaniga 2000; Thompson 2000). Neuroscience has focused on understanding how the brain works on average – for example, which bits of the brain light up under neuroimaging for particular tasks. Until now, genetics has entered neuroscience largely in relation to gene targeting in mice in which mutations are induced that break down normal brain processes. In humans, rare single-gene mutations are the centre of attention. This approach tends to treat all members of the species as if they were genetically the same except for a few rogue mutations that disrupt normal processes. In this sense, the species-typical perspective of neuroscience assumes that mental illness is a broken brain. In contrast, the individual-differences perspective considers variation as normal – the very standard deviation. Common mental illness is thought to be the quantitative extreme of the normal distribution.

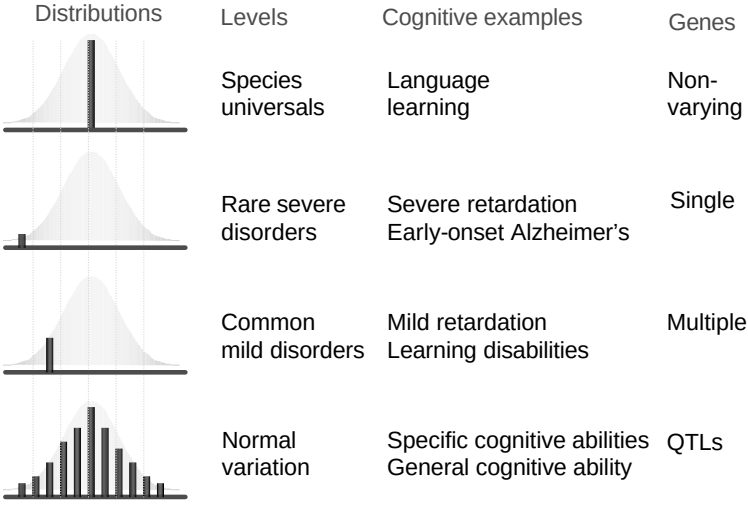

Figure 1.1 Levels of analysis.

Although perspectives are not right or wrong – just more or less useful for particular purposes – the species-typical perspective and the individual-differences perspective can arrive at different answers because they ask different questions. The distinction between the two perspectives is in essence the difference between means and variances. There is no necessary connection between means and variances, either descriptively or aetiologically. Despite its name, analysis of variance, the most widely used statistical analysis in the life sciences, is actually an analysis of mean effects in which individual differences are literally called the *error term*. Instead of treating differences between individuals as error, and averaging individuals across groups as in analysis of variance, individual-differences research focuses on these interindividual differences. Variation is distributed continuously, often in the shape of the familiar bell curve, and is indexed by variance (the sum of squared deviations from the mean), or the square of variance, which is called the standard deviation.

The two perspectives also differ methodologically. Most species-typical research is experimental in the sense that subjects are randomly assigned to conditions which consist of manipulating something such as genes, lesions, drugs and tasks. The dependent variable is the average effect of the manipulation on outcome measures such as single-cell recordings of synaptic plasticity, activation of brain regions assessed by neuroimaging, or performance on cognitive tests. Such experiments ask whether such manipulations *can* have an effect on average in a species. For example, a gene knock-out study investigates whether an experimental group of mice who inherit a gene that has been made dysfunctional differs, for example in learning or memory, from a control group with a normal copy of the gene. A less obvious example can be seen in recent experimental research that manipulated tasks and found that average blood flow assessed by positron emission tomography (PET) in the human species is greater in the prefrontal cortex for high-intelligence tasks than for low-intelligence tasks (Duncan et al. 2000).

In contrast, rather than creating differences between experimental and control groups through manipulations, the individual-differences perspective focuses on naturally occurring differences between individuals. One of the factors that makes individuals different is genetics. The individual-differences perspective is the foundation for quantitative genetics, which focuses on naturally occurring genetic variation, the stuff of heredity. Although 99.9% of the human DNA sequence is identical for all human beings, the 0.1% that differs – 3 million base pairs (enough for every gene for each of us to be different) – is ultimately responsible for the ubiquitous genetic influence found for all individual-differences traits including cognitive abilities and disabilities (Plomin et al. 2001*a*). Individual-differences research is correlational in the sense that it investigates factors that *do* have an effect in the world outside the laboratory. Continuing with the previous examples, an individual-differences approach would ask whether naturally occurring genetic

variation in mice is associated with individual differences in mouse learning and memory. Genes can be knocked out and shown to have major effects on learning and memory but this does not imply that the gene has anything to do with the naturally occurring genetic variation that is responsible for hereditary transmission of individual differences in performance on learning and memory tasks. The PET experiment that compared average performance on high- and low-intelligence tasks could be addressed from an individual-differences perspective by comparing cortical blood flow in high- and low-intelligence individuals rather than comparing average performance on tasks (Duncan et al. 2000).

Other perspectives or levels of analysis lie in between these two extremes of species universals and normal variation. The effects of rare severe disorders caused by a single gene are dramatic. For example, mutations in the gene that codes for the enzyme phenylalanine hydroxylase, if untreated, cause phenylketonuria (PKU) that is associated with a severe form of mental retardation. This inherited condition occurs in 1 in 10 000 births. At least 100 other rare single-gene disorders include mental retardation as part of the syndrome (Wahlström 1990). Such rare single-gene disorders can be viewed as aberrations from the species type, exceptions to the species rule. In contrast, common disorders – such as mild mental retardation and learning disabilities – seldom show any sign of single-gene effects and appear to be caused by multiple genes as well as by multiple environmental factors. Indeed, quantitative genetic research suggests that such common disorders are usually the quantitative extreme of the same genes responsible for variation throughout the distribution. Genes in such multiple-gene (polygenic) systems are called quantitative trait loci (QTL) because they are likely to result in dimensions (quantitative continua) rather than disorders (qualitative dichotomies). For example, as discussed later, reading disability has been linked to the short arm of chromosome 6 (6p21) in several QTL analyses (Willcutt et al., in press). When the gene responsible for this linkage is isolated, the QTL prediction is that it will not reveal a gene for reading disability per se. Rather, the gene is one of many that are expected to contribute quantitatively to reading performance throughout the distribution. In other words, in terms of the genetic aetiology of common disorders, there may be no disorders, just dimensions. Other than simple and rare single-gene or chromosomal disorders, mental illness may represent the extreme of normal variation.

In summary, the individual-differences perspective views variation as normal and distributed continuously; common disorders are viewed as the extremes of these continuous distributions. As indicated at the outset, perspectives are not right or wrong. But they are different, and the proper interpretation of genetic research depends on understanding these differences. The perspectives are complementary in the sense that a full understanding of behaviour requires integration across all levels of analysis.

Quantitative genetics

Although much human genetic research focuses on rare single-gene disorders (see Chapter 2 by Skuse and Baker) and much genetic research using animal models focuses on gene knockouts (see Chapter 19 by Stephens et al.), the present chapter concentrates on the genetics of individual differences, both quantitative genetics and molecular genetics. Quantitative genetic research such as twin and adoption studies is hardly needed any longer merely to ask whether and how much genetic factors influence behavioural traits, because the answers are 'yes', and 'a lot' for nearly all dimensions and disorders that have been studied (Plomin et al. 2001*a*). However, new quantitative genetic techniques make it possible to go beyond these rudimentary questions to investigate how genes and environment affect developmental change and continuity, comorbidity and heterogeneity, and the links between disorders and normal variation. Using these techniques, quantitative genetic research can lead to better diagnoses based in part on aetiology rather than solely on symptomatology. They can also chart the course for molecular genetic studies by identifying the most heritable components and constellations of disorders as they develop and as genetic vulnerabilities correlate and interact with the environment. The future of genetic research on cognitive abilities and disabilities lies with molecular genetic research that attempts to identify specific genes responsible for heritability. Although progress in identifying genes for complex traits such as cognitive abilities and disabilities has been slow, when such genes are found the next step will be to understand the brain mechanisms that mediate genetic effects on behaviour.

This chapter focuses on genetics but it should be mentioned at the outset that quantitative genetic research is at least as informative about nurture as it is about nature. In the first instance, it provides the best available evidence for the importance of the environment, in that the heritability of complex traits is seldom greater than 50%. In other words, about half of the variance cannot be explained by genetic factors. In addition, two of the most important findings about environmental influences on behaviour have come from genetic research. The first finding is that, contrary to socialization theories from Freud onwards, environmental influences operate to make children growing up in the same family as different as children growing up in different families, which is called nonshared environment (reviewed by Plomin et al. 2001*b*). The second finding, called the nature of nurture, is that genetic factors influence the way we experience our environments (reviewed by Plomin 1994). For this reason, most measures of the environment used in behavioural research show genetic influence. For the same reason, associations between environmental measures and behavioural outcome measures are often substantially mediated genetically. The way forward in research is to bring together genetic and environmental strategies, for example, using environmental measures in genetically sensitive

designs to investigate interactions and correlations between nature and nurture. The present chapter's focus on genetics is not intended to denigrate the importance of environmental influences or to imply biological determinism.

General cognitive ability

General cognitive ability (*g*) is a highly heritable quantitative trait that varies from a low end of mild mental retardation to a high end of gifted individuals (Plomin 1999*a*). One of the most consistent findings from individual-differences research on human cognitive abilities and disabilities during the past century is that diverse cognitive processes intercorrelate. Despite the diversity of cognitive tests, individuals who perform well on one test tend to do well on other tests. In a meta-analysis of 322 studies that included hundreds of different kinds of cognitive tests, the average correlation among the tests was about 0.30 (Carroll 1993). A technique called factor analysis, in which a composite score is created that represents what is shared in common among the measures, indicates that *g* accounts for about 40% of the total variance of cognitive tests (Jensen 1998). However, *g* is not just a statistical abstraction – one can simply look at a matrix of correlations among such measures and see that there is a positive manifold among all tests and that some measures (such as spatial and verbal ability) intercorrelate more highly on average than do other measures (such as nonverbal memory tests). Because all of these measures intercorrelate to some extent, *g* is also indexed reasonably well by a simple total score on a diverse set of cognitive measures, as is done in IQ tests. This overlap emerges not only for traditional measures of reasoning, spatial, verbal and memory abilities such as those mentioned above but also for information-processing tasks that rely on reaction time and other cognitive tasks used to assess, for example, working memory (Anderson 1992; Stauffer et al. 1996; Baddeley and Gathercole 1999; Deary 2000).

General cognitive ability was recognized nearly a century ago by Charles Spearman (1904, 1927), who used *g* as a neutral signifier that avoided the many connotations of the word *intelligence*. *g* is one of the most reliable and valid traits in the behavioural domain (Jensen 1998). Its long-term stability after childhood is greater than for any other behavioural trait (Deary et al. 2000), it predicts important social outcomes such as educational and occupational levels far better than any other trait (Gottfredson 1997), and it is a key factor in cognitive ageing (Salthouse and Czaja 2000). There are of course many other important noncognitive abilities, such as athletic ability, but there seems to be nothing to be gained by lumping all such abilities together as is done with the popular notion of 'multiple intelligences' (Gardner 1983). Also, *g* by no means guarantees success either in school or in the workplace – achievement also requires personality, motivation and social skills, currently referred to as 'emotional intelligence' (Goleman 1995).

Although the concept of *g* is widely accepted (Neisser et al. 1996; Carroll 1997; Snyderman and Rothman 1987), acceptance is not universal. The arguments against *g* have been reviewed (Jensen 1998). They include ideological issues such as political concerns and the notion that *g* merely reflects knowledge and skills that happen to be valued by the dominant culture (Gould 1996). Objections of a more scientific nature include theories that focus on specific abilities (Gardner 1983; Sternberg 1985). However, when these theories are examined empirically, *g* shines through. For example, one of the major advocates of a 'componential' view to cognitive processing conceded that 'We interpret the preponderance of evidence as overwhelmingly supporting the existence of some kind of general factor in human intelligence. Indeed, we are unable to find any convincing evidence at all that militates against this view' (Sternberg and Gardner 1983). *g* is not the whole story – group factors representing specific abilities are also important levels of analysis – but trying to tell the story of cognitive abilities without *g* loses the plot entirely.

The existence of *g* appears to go against the tide of current cognitive neuroscience which considers cognitive processes as 'modular' – specific and independent (Fodor 1983; Pinker 1994). However, as mentioned earlier, research in cognitive neuroscience focuses on species-typical processes. *g* is not about average performance – it is about individual differences in performance, and the fact that individuals who perform well on some tasks tend to perform well on most tasks. The investigation of individual differences represents a different level of analysis where the data clearly point to *g*. The existence of *g* does not imply that the source of *g* must be a single general physical (dendritic complexity, myelinization), physiological (synaptic plasticity, speed of nerve conduction) or psychological (working memory, executive function) process (Deary 2000). It seems more reasonable to suppose that *g* represents a concatenation of such physical, physiological and psychological processes that are all enlisted to solve functional problems. As an analogy, athletic ability depends on psychological (motivation), physiological (oxygen transport) and physical (bone structure) processes. Athletic ability is not one of these things, it is all of these things.

There is more research addressing the genetics of *g* than any other human characteristic. Dozens of studies including more than 8000 parent–offspring pairs, 25 000 pairs of siblings, 10 000 twin pairs and hundreds of adoptive families all converge on the conclusion that genetic factors contribute substantially to *g* (Plomin et al. 2001*a*). Estimates of the effect size, called heritability, vary from 40–80% but estimates based on the entire body of data are about 50%, indicating that genes account for about half of the variance in *g*. Sorting the results by age indicates that heritability increases from about 0.20 in infancy to about 0.40 in childhood and to 0.60 or higher later in life (McGue et al. 1993), even for individuals 80+ years old (McClearn et al. 1997). This increase in heritability throughout the lifespan is

interesting, because it is counterintuitive in relation to Shakespeare's 'slings and arrows of outrageous fortune' accumulating over time. This finding suggests that people actively select, modify and even create environments conducive to the development of their genetic proclivities. For this reason, it may be more appropriate to think about *g* as an appetite rather than an aptitude.

The most important finding comes from multivariate genetic analysis, which is used to examine covariance among specific cognitive abilities, rather than the variance of each trait considered separately. It yields a statistic called the *genetic correlation*, which is an estimate of the extent to which genetic effects on one trait correlate with genetic effects on another trait independent of the heritability of the two traits. That is, although all cognitive abilities are moderately heritable, the genetic correlations between them could be anywhere from 0.0, indicating complete independence, to 1.0, indicating that the same genes influence different cognitive abilities. In the case of cognitive abilities, multivariate genetic analyses have consistently found that genetic correlations among cognitive abilities are very high – close to 1.0 (Petrill 1997). In other words, if a gene were found that is associated with a particular cognitive ability, the same gene would be expected to be associated with all other cognitive abilities as well. As noted earlier, *g* accounts for about 40% of the total variance of cognitive tests. In contrast, multivariate genetic research indicates that *g* accounts for nearly *all* of the genetic variance of cognitive tests. That is, what is in common among cognitive abilities is almost completely genetic in origin. This finding has the interesting converse implication that what is specific to each cognitive test is largely environmental – what makes us good at all tests is largely genetic but what makes us better at some tests than others is largely environmental.

This finding from multivariate genetic research provides clues for understanding how the brain works from an individual-differences perspective. Spearman, who first described *g* in 1904, noted that ultimate understanding of *g* 'must needs come from the most profound and detailed study of the human brain in its purely physical and chemical aspects' (Spearman 1927, p. 403). The simplest brain model of genetic *g* is that there is a single fundamental brain process that permeates all other brain processing such as neural speed (e.g. myelinization), power (e.g. number of neurons) or fidelity (e.g. density of dendritic spines). The opposite model is that there are many brain processes that are uncorrelated phenotypically and genetically, but lead to a genetic correlation in performance on cognitive tasks because all of these brain processes are enlisted by the cognitive tasks. A middle position is that multiple brain processes underlie *g* in cognitive tasks but these processes are correlated phenotypically and genetically. That is, genetic *g* might exist in the brain as well as the mind. To test these different models about brain mechanisms responsible for *g*, it is necessary to identify reliable individual differences in brain processes and investigate the phenotypic and genetic relationships among these processes.

Human research on *g* can make progress towards understanding brain mechanisms using neuroimaging techniques (Kosslyn and Plomin 2001). However, mouse models of *g* would facilitate the precise analysis of basic brain mechanisms using techniques such as single cell recordings, micro-stimulation, targeted gene mutations, antisense DNA that disrupts gene transcription and DNA expression studies. Clearly, there are major differences in brain and mind between the human species and other animals, most notably in the use of language and the highly developed prefrontal cortex in the human species. However, *g* in humans does not depend on the use of language – a strong *g* factor emerges from a battery of completely nonverbal tests (Jensen 1998) – and low-level tasks such as information-processing tasks assessed by reaction time contribute to *g* (Deary 2000). Indeed, *g* can be used as a criterion to identify animal models of individual differences in cognitive processes. If *g* represents the way in which genetically driven components of the brain work together to solve problems, it would not be unreasonable to hypothesize that *g* exists in all animals (Anderson 2000). Although much less well documented than *g* in humans, increasing evidence exists for a *g* factor in mice across diverse tasks of learning, memory and problem solving (Plomin 2001).

Specific cognitive abilities

Although *g* is important, there is much more to cognitive functioning. Cognitive abilities are usually considered in a hierarchical model (Figure 1.2). General cognitive ability is at the top of the hierarchy, representing what all tests of cognitive ability have in common. Below general cognitive ability in the hierarchy are broad factors of specific cognitive abilities, such as verbal ability, spatial ability, memory and speed of processing. These broad factors are indexed by several tests, shown at the bottom of the hierarchy in Figure 1.2. In addition to specific tests, the bottom of the hierarchy can also be considered in terms of elementary processes thought to be involved in information processing.

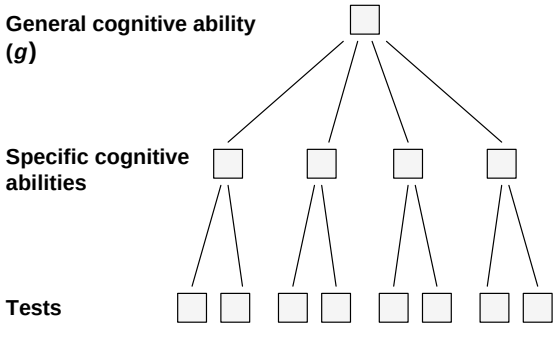

Figure 1.2 Hierarchical model of cognitive abilities.

Many specific cognitive abilities show genetic influence in twin studies, although the magnitude of the genetic effect is generally lower than that for general cognitive ability (Plomin and DeFries 1998). Family and twin studies suggest that the genetic contribution may be stronger for some cognitive abilities such as verbal and spatial than for other abilities, especially nonverbal memory. Recent studies of twins reared apart generally confirm these findings. Developmental genetic analyses indicate that genetically distinct specific cognitive abilities can be found as early as 3 years of age and show increasing genetic differentiation from early to middle childhood. Twin studies also indicate genetic influence on information-processing measures and brain-wave measures of EEG and event-related potentials (Plomin et al. 2001*a*).

Mental retardation

If genetics substantially influences general cognitive ability, one might expect that low IQ scores are also due to genetic factors. However, this conclusion does not necessarily follow. For example, mental retardation can be caused by environmental trauma, such as birth problems, nutritional deficiencies and head injuries. Given the importance of mental retardation, it is surprising that no twin or adoption studies of diagnosed mental retardation have been reported. Nonetheless, one sibling study suggests that moderate and severe mental retardation may be due largely to nonheritable factors. In a study of over 17 000 white children, 0.5% were moderately to severely retarded (Nichols 1984). The siblings of these retarded children were not retarded – the siblings' average IQ was 103 and ranged from 85–125. In other words, moderate to severe mental retardation showed no familial resemblance, which implies that mental retardation is not heritable. Although most moderate and severe mental retardation may not be inherited from generation to generation, it is often caused by noninherited DNA events, such as new gene mutations and new chromosomal abnormalities such as Down's syndrome (see Chapter 2 by Skuse and Baker). This suggestion of low overall heritability for moderate to severe mental retardation does not contradict the finding that mental retardation is a symptom for some rare single-gene syndromes such as phenylketonuria (1 in 10 000 births).

In contrast, in this same study, siblings of mildly retarded children (1.2% of the sample) showed lower than average IQ scores. The average IQ for these siblings of mildly retarded children was only 85. These important findings – that mild mental retardation is familial whereas moderate and severe retardation is not familial – also emerged from the largest family study of mild mental retardation, which considered 80 000 relatives of 289 mentally retarded individuals (Reed and Reed 1965). This parent–offspring family study showed that mild mental retardation is very strongly familial. If one parent is mildly retarded, the risk for retardation in their children is about 20%. If both parents are retarded, the risk is nearly 50%.

Although mild mental retardation runs in families, it could do so for reasons of nurture rather than nature. Twin and adoption studies of mild mental retardation are needed to disentangle the relative roles of nature and nurture. Although no proper twin or adoption studies of diagnosed mild mental retardation have been reported, three small twin studies suggest that low IQ is at least as heritable as IQ in the normal range (Plomin 1999*a*). These studies also suggest that mild mental retardation may be the lower end of the distribution of the same genetic and environmental factors that are responsible for general cognitive ability.

Reading disability

Reading is the primary problem in about 80% of children with a diagnosed learning disorder. As many as 10% of children have difficulty learning to read. For some, specific causes can be identified, such as mental retardation, brain damage, sensory problems and deprivation. However, many children without such problems find it difficult to read. Children with a specific reading disorder (also known as *dyslexia*) read slowly, and often with poor comprehension. When reading aloud, they perform poorly.

Family studies have shown that reading disability runs in families. The largest family study included 1044 individuals in 125 families with a reading-disabled child and 125 matched control families (DeFries et al. 1986). Siblings and parents of the reading-disabled children performed significantly worse on reading tests than did siblings and parents of control children. Earlier twin studies suggested that familial resemblance for reading disability involves genetic factors (Bakwin 1973; Decker and Vandenberg 1985). Although one twin study showed little evidence of genetic influence (Stevenson et al. 1987), the largest twin study confirmed genetic influence on reading disability (DeFries et al. 1999). For more than 250 twin pairs in which at least one member of the pair was reading disabled, twin concordances were 66% for identical twins and 36% for fraternal twins, a result suggesting moderate genetic influence.

As part of this twin study, a new method was developed to estimate the genetic contribution to the mean difference between the reading-disabled probands and the mean reading ability of the population. DF extremes analysis (DeFries and Fulker 1985, 1988) takes advantage of quantitative scores of the relatives of probands rather than just assigning a dichotomous diagnosis to the relatives and comparing twin concordances for the disorder. To the extent that reading deficits of probands are heritable, the quantitative reading scores of identical co-twins will be more similar to that of the probands than will the scores of fraternal twins. In other words, the mean reading score of identical co-twins will regress less far back toward the population mean than will that of fraternal co-twins, which is the case for reading disability (DeFries and Gillis 1993). Half of the mean difference between the probands and the

population is heritable. This is called 'group heritability' to distinguish it from the usual heritability estimate, which refers to differences between individuals rather than to mean differences between groups. Results of DF extremes analysis indicates that group heritability for reading disability is moderate and similar to individual heritability estimates for reading, suggesting that reading disability is quantitatively rather than qualitatively different from the normal range of reading ability (DeFries and Gillis 1993).

Various modes of transmission have been proposed for reading disability. The autosomal dominant hypothesis takes into account the high rate of familial resemblance but fails to account for the fact that about a fifth of reading-disabled individuals do not have affected relatives. An X-linked recessive hypothesis is suggested when a disorder occurs more often for males than females, as is the case for reading disability. However, the X-linked recessive hypothesis does not work well as an explanation of reading disability. One of the hallmarks of X-linked recessive transmission is the absence of father-to-son transmission, because sons inherit their X chromosome only from their mother. Contrary to the X-linked recessive hypothesis, reading disability is transmitted from father to son as often as from mother to son. It is now generally accepted that, like most complex disorders, reading disability is caused by multiple genes as well as by multiple environmental factors.

Communication disorders

Despite the strong trend of much linguistic theorizing to invoke an innate basis for language (Pinker 1994), genetic research has been slow in coming to the field of language, but the field is making up for lost time (Gilger 1997; Plomin and Dale 2000; Rice 1996). DSM–IV (American Psychiatric Association 1994) includes four types of communication disorders: expressive language (putting thoughts into words) disorder, mixed receptive (understanding the language of others) and expressive language disorder, phonological (articulation) disorder and stuttering (speech interrupted by prolonged or repeated words, syllables or sounds). Hearing loss, mental retardation and neurological disorders are excluded.

Several family studies, examining communication disorders broadly, indicate that communication disorders are familial (Stromswold 2001). For children with communication disorders, about a quarter of their first-degree relatives report similar disorders, compared with about 5% for the relatives of controls (Felsenfeld 1994). Three twin studies of communication disorders found evidence for extremely high heritability, with average concordances of about 90% for identical twins and 50% for fraternal twins (Lewis and Thompson 1992; Bishop et al. 1995; Tomblin and Buckwalter 1998). The only adoption study of communication disorders confirms the twin results (Felsenfeld and Plomin 1997).

These disorders are frequently comorbid but little is known about the genetic and environmental links between them as they emerge in infancy and early childhood.

Multivariate genetic analysis suggests that DSM–IV diagnostic categories may not reflect the genetic origins of these disorders (Plomin and Dale 2000). For example, expressive and receptive language disorders overlap genetically, whereas genetic factors appear to be different for individuals who have articulation problems and those who do not (Bishop et al. 1995).

A large-scale study of twins in infancy and early childhood is under way in the UK to investigate the genetics of early-onset language problems and their relationship to other cognitive and behaviour problems (Plomin and Dale 2000). The study shows that vocabulary delay is highly heritable (73% group heritability using DF extremes analysis) as early as 2 years of age (Dale et al. 1998). Examples of multivariate genetic findings include a high genetic correlation between lexical (vocabulary) and grammatical (sentence complexity) development (Dale et al. 2000) and a strong genetic correlation between language and nonverbal cognitive development (Price et al. 2000).

Family studies of stuttering over the past 50 years have shown that about a third of stutterers have other stutterers in their families. The Yale Family Study of Stuttering includes nearly 600 stutterers and more than 2000 of their first-degree relatives (Kidd 1983). About 15% of the first-degree relatives reported that they had stuttered at some point in their life, about five times greater than the base rate of approximately 3% in the general population. Moreover, about half of the affected first-degree relatives were considered to be chronic stutterers. One small twin study of stuttering suggests that familial resemblance is heritable, with concordances of 77% for identical twins (17 pairs) and 32% for fraternal twins (13 pairs) (Howie 1981). A large twin study that included a single item about stuttering in a questionnaire study also found evidence for substantial genetic influence (Andrews et al. 1991). Although much remains to be learned about the genetics of stuttering, the evidence as it stands suggests substantial genetic influence (Yairi et al. 1996).

Molecular genetics

The twentieth century began with the rediscovery of Mendel's laws of heredity. The word *genetics* was only invented in 1903. Fifty years later it was understood that DNA was the mechanism of heredity. The genetic code was cracked in 1966; the 4-letter alphabet (G, A, T, C) of DNA is read as 3-letter words that code for the 20 amino acids that are the building blocks of proteins. The crowning glory of the century and a tremendous start to the new century is the Human Genome Project which has provided a working draft of the sequence of the 3 billion letters of DNA in the human genome.

When the working draft of the human genome sequence was published in February 2001, much publicity was given to the finding that there are fewer than half as many genes (30 000) in the human genome as expected – about the same

number of genes as in mice and worms. A bizarre spin in the media was that having only 30 000 genes implies that nurture must be more important than we thought. The idea that fewer genes means more free will is silly. Do flies have more free will than us because they have fewer genes? However, the finding that the human species does not have more genes than other species is important in suggesting that the number of genes is not responsible for the greater complexity of the human species. In part, the greater complexity of the human species occurs because during the process of decoding genes into proteins, human genes, more than the genes of other species, are spliced in alternative ways to create a greater variety of proteins. The greater complexity of the human species may be due to quality rather than quantity: other subtle variations in genes rather than the number of genes may be responsible for differences between mice and men. If subtle DNA differences are responsible for the differences between mice and men, even more subtle differences are likely to be responsible for individual differences within the species.

Another interesting finding from the Human Genome Project is that only 2% of the 3 billion letters in our DNA code involves genes in the traditional sense, that is, genes that code for amino-acid sequences. This 2% figure is similar in other mammals. On an evolutionary time scale, mutations are quickly weeded out from these bits of DNA that are so crucial for development. When mutations are not weeded out, they can cause one of the thousands of severe but rare single-gene disorders. However, it seems unlikely that the other 98% of DNA is just along for the ride. For example, variations in this other 98% of the DNA are known to regulate the activity of the classical genes. For this reason, the other 98% of DNA might be the place to look for genes associated with quantitative rather than qualitative effects on behavioural traits.

The most exciting development for behavioural genetics is the identification of the DNA sequences that make us different from each other. There is no human genome sequence – we each have a unique genome. Indeed, about one in every thousand DNA letters differs, about 3 million variations in total. Many of these DNA differences have already been identified. The Human Genome Project has spawned new technologies that will make it possible to investigate simultaneously thousands of DNA variants as they relate to behavioural traits. These DNA differences are responsible for the widespread heritability of psychological disorders and dimensions. That is, when we say that a trait is heritable, we mean that variations in DNA exist that cause differences in behaviour.

DNA variation has a unique causal status in explaining behaviour. When behaviour is correlated with anything else, the old adage applies that correlation does not imply causation. For example, when parenting is shown to be correlated with children's behavioural outcomes, this does not imply that the parenting caused the outcome environmentally. Indeed, it has been shown that parental behaviour to

some extent reflects genetic effects on children's behaviour (Plomin 1994). When it comes to interpreting correlations between biology and behaviour, such correlations are often mistakenly interpreted as if biology causes behaviour. For example, correlations between neurotransmitter physiology and behaviour, or between neuroimaging indices of brain activation and behaviour, are often interpreted as if brain differences cause behavioural differences. However, these correlations do not necessarily imply causation. Behavioural differences can cause brain differences. In contrast, in the case of correlations between DNA variants and behaviour, the behaviour of individuals does not change their genome. Expression of genes can be altered but the DNA sequence itself does not change. For this reason, correlations between DNA differences and behavioural differences can be interpreted causally: DNA differences cause the behavioural differences and not the other way around.

Integration of quantitative genetics and molecular genetics

Since its origins early in the twentieth century, quantitative genetics has focused on commercially valuable traits in plants and animals and socially important traits in the human species, especially behavioural dimensions and disorders. Techniques were developed such as twin and adoption designs for humans, and inbred strain and selection studies for animals, in order to investigate the extent to which genetic factors contribute to the observed differences in such complex traits. Such studies consistently pointed to an important role for genetics even for the most complex of all traits, behaviour. However, the evidence pointed to the involvement of many genes as well as many environmental factors, so that it seemed hopeless to identify specific genes responsible for the genetic contribution to most behavioural traits. In contrast to the focus of quantitative genetics on important phenotypes and on naturally occurring genetic variation responsible for phenotypic differences, molecular genetics focused on genes and techniques to create new mutations in model organisms such as the fruit fly, in order to investigate how genes work. Because of their differences in perspectives and methods, these two approaches to genetics diverged steadily during the twentieth century.

In the 1980s the development of a new generation of polymorphisms in DNA itself began to make it possible to identify genes responsible for the heritability of complex traits influenced by many genes as well as by many environmental factors. As mentioned earlier, the pace of this integration has been accelerated dramatically as a result of the Human Genome Project which has brought us to the threshold of a postgenomic world in which the genome sequence of our species and others is known. Most importantly for the analysis of complex traits, several million DNA variations (polymorphisms) in the genome sequence are being identified which are the ultimate causes of the ubiquitous heritability of complex traits.

If many genes influence a trait, the trait is likely to be distributed quantitatively in a continuous distribution. For this reason, such genes are often referred to as quantitative trait loci (QTLs). The name implies that complex traits influenced by multiple genes are thought to be distributed as continuous, quantitative dimensions rather than as discontinuous, qualitative disorders. Unlike single-gene effects that are necessary and sufficient for the development of a disorder, QTLs act like probabilistic risk factors. Although QTLs are inherited in the same mendelian manner as single-gene effects, if many genes affect a trait then each gene is likely to have a relatively small effect. This makes it much more difficult to detect QTLs than single-gene effects but the potential availability of millions of DNA markers makes this daunting prospect possible. A revolutionary implication of the QTL perspective is that there may be no disorders from a genetic perspective. Disorders may merely be the quantitative extreme of the same genetic factors that contribute to heritability throughout the dimension (Plomin et al. 1994). That is, there may be no genes specific to a disorder – genes associated with a disorder might have the same effect throughout the distribution. In other words, a QTL associated with reading disability may actually be associated with the entire continuum of reading ability, that is, with the high end and middle of the distribution as well as the low end.

In the case of single-gene effects such as phenylketonuria that cause severe mental retardation, the gene is necessary and sufficient for the development of the disorder. In contrast, cognitive disorders in childhood such as reading disability are much more common than any known single-gene disorders, with risks often reported to be as high as 1 in 10. Traditional methods for identifying single-gene effects such as the use of large family pedigrees are unlikely to succeed in identifying QTLs because the effect size of individual QTLs will be relatively small. The earliest attempts to find genes for behavioural disorders focused on schizophrenia and manic-depressive psychosis at a time when gene-hunting techniques were limited to identifying a single gene necessary and sufficient to cause the disorder. Although there has never been any solid evidence that these disorders are caused by a single gene, this single-gene approach was all that was available at that time for finding genes. Although there are several promising leads, no clear-cut associations with schizophrenia and bipolar affective disorder have been identified (Baron 2001). We now realize that such designs are only able to detect genes of major effect size.

Identifying QTLs

Molecular genetic studies in the cognitive domain were begun relatively recently and have used QTL approaches from the start, which may contribute to the quicker successes in this domain. An example of a behavioural QTL is the association between apolipoprotein-E and late-onset dementia (Corder et al. 1993), an association that has been replicated in scores of studies and remains the only known predictor of

this common disorder in later life. Although a particular allele in this gene leads to a five-fold increased risk for dementia, it is a QTL in the sense that many people with this allele do not have dementia and most people with dementia do not have this genetic risk factor. As described in the following section, a replicated QTL linkage has been found for another cognitive disorder, reading disability.

The advantage of linkage approaches, including QTL linkage approaches, is that they can systematically scan the genome for linkages using just a few hundred DNA markers. QTL linkage designs use many small families (usually siblings) rather than a few large families. The disadvantage is that they can only detect QTLs that account for a substantial amount (perhaps 10%) of the genetic variance. In contrast, allelic association can detect QTLs that account for 1% of the variance but thousands of DNA markers are needed to screen the genome systematically for association because association can only be detected if a DNA marker is very close to a QTL. In other words, linkage is systematic but not powerful and allelic association is powerful but not systematic (Risch and Merikangas 1996). For this reason, allelic association approaches have largely been limited to studies of 'candidate' genes. If, as is usually the case, a DNA marker in or near a candidate gene is not itself functional (that is, it does not produce a coding difference), the marker may be close enough to a functional QTL to be in linkage disequilibrium with it and thus yield an indirect association with the disorder. One problem with a candidate gene approach is that, for behavioural disorders, any of the thousands of genes expressed in the brain could be viewed as candidate genes. The way out of this conundrum is to conduct a systematic scan of the genome for allelic association using many thousands of DNA markers, although tens of thousands or even hundreds of thousands of DNA markers would be needed to identify or exclude all QTL associations (Kruglyak 1999). Such systematic large-scale genome scans or scans of all known candidate gene polymorphisms are becoming possible with new technologies that can quickly genotype thousands of DNA markers (Watson and Akil 1999).

Progress in identifying genes associated with behaviour and other complex traits has been slower than expected, in part because research to date has been underpowered for finding QTLs of small-effect size, especially in linkage studies. Very large studies are needed in order to identify QTLs of small-effect size (Cardon and Bell 2001). A daunting target for molecular genetic research on complex traits such as behaviour is to design research powerful enough to detect QTLs that account for 1% of the variance, while providing protection against false positive results in genome scans of thousands of genes. In order to break the 1% QTL barrier, samples of many thousands of individuals are needed for research on disorders (comparing cases and controls) and on dimensions (assessing individual differences in a representative sample). Another factor in the slow progress to date is that only a few candidate gene markers have been examined rather than systematic scans of

gene systems or of the entire genome. The Human Genome Project will accelerate progress towards identifying all functional DNA variants expressed in the brain, especially those for entire neurotransmitter pathways.

QTLs and cognitive abilities and disabilities

Mild mental retardation and general cognitive ability

QTL linkage or association studies of mild mental retardation have not yet been reported even though quantitative genetic research mentioned earlier suggests that mild mental retardation represents the lower extreme of the same multiple genetic and environmental factors that affect cognitive functioning in the normal range. Although systematic QTL studies of mild mental retardation have not been reported, a QTL perspective suggests that QTL studies of normal IQ or even high IQ could identify QTLs that are also associated with mild mental retardation. This is part of the rationale for an allelic association study comparing high-IQ and control individuals called the IQ QTL Project (Plomin 2002). The first phase of the project employed an allelic association strategy using DNA markers in or near candidate genes likely to be relevant to neurological functioning, such as genes for neuroreceptors. Allelic association results were reported for 100 DNA markers for such candidate genes (Plomin et al. 1995). Although several significant associations were found in an original sample, only one association was replicated in an independent sample. However, a recent attempt to replicate this finding was not successful (Hill et al., in press).

As mentioned earlier, attempts to find QTL associations with complex traits have begun to go beyond candidate genes to conduct systematic genome scans using dense maps of DNA markers. As part of the IQ QTL Project, a first attempt to use this approach to identify QTLs associated with IQ focused on the long arm of chromosome 6, and found replicated associations for a DNA marker that happened to be in the gene for insulin-like growth factor-2 receptor (*IGF2R*) (Chorney et al. 1998), which has been shown to be especially active in brain regions most involved in learning and memory (Wickelgren 1998). Another polymorphism in the *IGF2R* gene has been genotyped and similar results were found for the new polymorphism in a new sample (Hill et al., in press).

The problem with using a dense map of markers for a genome scan is the amount of genotyping required. In order to scan the entire genome at 1 million DNA base-pair intervals (1 Mb), about 3500 DNA markers would need to be genotyped. This would require 700 000 genotypings in a study of 100 high '*g*' individuals and 100 controls. With markers at 1 Mb intervals, no QTL would be farther than 500 000 base pairs from a marker. Moreover, it is generally accepted that 10 to even 100 times as many markers would be needed in order to detect all QTLs (Kruglyak 1999;

Abecasis et al. 2001; Reich et al. 2001). Despite the daunting amount of genotyping required for a systematic genome scan, this approach has been fuelled by the promise of 'SNPs on chips' which can quickly genotype thousands of DNA markers of the single nucleotide polymorphism (SNP) variety.

DNA pooling, developed for use in the IQ QTL Project, provides a low-cost and flexible alternative to SNPs on chips for screening the genome for QTL associations (Daniels et al. 1998). DNA pooling greatly reduces the need for genotyping by pooling DNA from all individuals in each group and comparing the pooled groups so that only 14 000 genotypings are required to scan the genome in the previous example involving 3500 DNA markers. A scan of 1842 DNA markers using DNA pooling and a multiple-stage design found two markers that yielded significant results in two independent case–control studies but neither reached significance in a third within-family study (Plomin et al. 2001*c*). Rather than genotyping additional anonymous DNA markers, the IQ QTL Project is now focusing on functional polymorphisms such as SNPs in coding regions (cSNPs) and SNPs in regulatory regions in which the marker can be presumed to be the QTL.

Reading disability

The first QTL linked to a human behavioural disorder by a QTL linkage approach has been reported and replicated for reading disability (Cardon et al. 1994). The method used was sib-pair QTL linkage, which is conceptually similar to DF extremes analysis. Instead of comparing identical and fraternal twins, siblings are compared who share 0, 1 or 2 alleles for a particular DNA marker. If siblings who share more alleles are also more similar for a quantitative trait such as reading ability, then QTL linkage is implied. QTL linkage analysis is much more powerful when one sibling is selected on the basis of an extreme score on the quantitative trait. When one sibling was selected for reading disability, the reading ability score of the co-sibling was also lower when the two siblings shared alleles for markers in a certain region on the short arm of chromosome 6 (6p21). Significant linkage was also found for markers in this region in an independent sample of fraternal twins and in three replication studies (Grigorenko et al. 1997; Fisher et al. 1998; Gayán et al. 1999). The linkage to chromosome 6 appears for both phonological and orthographic reading measures. In 1983, linkage to chromosome 15 was reported using traditional analyses of pedigrees (Smith et al. 1983). Chromosome 15 linkage (15q21) for reading disability has also been replicated in several studies (Smith et al. 1991; Grigorenko et al. 1997; Schulte-Körne et al. 1997).

The next step is to pin down the specific genes responsible for these QTL regions (Smith et al. 1998). When the specific genes are identified (so far, the QTL linkage has only been tracked to its neighbourhood of several million base pairs of DNA rather than to a specific location), it will be of great interest to investigate the

extent to which the gene's effects are specific to reading, or affect language or other cognitive processes more broadly.

Communication disorders

A single-gene disorder with its primary effect on language has been reported, albeit for a single family (Fisher et al. 1998). This family included 15 linguistically impaired relatives whose speech has low intelligibility and who have deficits in nearly all aspects of language but especially grammar. The family showed a simple dominant mode of inheritance that could be traced to one grandmother. A region on the long arm of chromosome 7 (7q31) was found that is linked to the disorder in this family (Lai et al. 2001). The same region has also been linked with autism (International Molecular Genetics Study of Autism Consortium 1998).

QTL linkage studies of specific language impairment are under way in Oxford and Edinburgh. Quantitative genetic results can help to chart the course for molecular genetic research in this area. At the most rudimentary level, the Twins Early Development Study (TEDS), described earlier, has shown that language problems even at 2 years of age are highly heritable, suggesting that language impairment is a good target for molecular genetic research. Although it is not unreasonable to focus on specific language impairment, quantitative genetic research suggests that genetic effects on persistent language problems may be general to cognitive development rather than specific to language (Plomin and Dale 2000). For this reason, TEDS has launched an allelic-association genome scan using DNA pooling in an attempt to identify some QTLs responsible for general language impairment and general cognitive impairment and comorbidity between them.

Once QTLs are found for language disability, hypotheses derived from quantitative genetic research can be tested empirically, for example, to assess the extent to which QTLs are specific to language (or to some specific component of language) or general to cognitive impairment. Indeed, all the questions raised by quantitative genetics – about developmental change and continuity, about multivariate issues of heterogeneity and comorbidity and the links between the normal and abnormal, and about the interplay between nature and nurture – can be addressed much more precisely and profitably once specific genes are identified (Plomin and Rutter 1998).

Conclusions

Despite the slow progress to date in finding genes associated with cognitive abilities and disabilities, their substantial heritability means that DNA polymorphisms exist that affect these traits. I am confident that we will find some of them. Although attention is now focused on finding specific genes associated with complex traits, the greatest impact for the neurobehavioural sciences will come after genes have