PhD DISSERTATIONS OF
NORTHWESTERN ITALIAN PHILOSOPHY CONSORTIUM – FINO

A PHILOSOPHICAL ANALYSIS OF GENERAL INTELLIGENCE: BIOLOGICAL, COGNITIVE AND ONTOLOGICAL ASPECTS

Davide Serpico

2018

Examination Committee:

John Dupré
Andrea Borghini
Marcello Frixione

UNIVERSITÀ DEGLI STUDI DI GENOVA
UNIVERSITÀ DEGLI STUDI DI TORINO
UPO
UNIVERSITÀ DEGLI STUDI DI PAVIA

Index

Introduction

The notion of intelligence, although widely exploited both in science and in everyday life, is shrouded in a very controversial aura. People have personal intuitions about what intelligence is, and how to use this term in the appropriate context. Nevertheless, they acknowledge that the meaning of intelligence may change depending on the circumstances. 'Smart', 'clever', 'wise', 'creative', 'perceptive', 'bright', 'astute', 'lectured', 'responsive', and so forth: all these labels refer to some aspects of the intelligent behavior characterizing human beings. Some of these terms identify people capable of finding original solutions to problems, or of "thinking outside the box", whereas some others denote individuals who employ sophisticated reasoning or concepts. Some of them refer to people who are particularly good at handling social circumstances, thanks to their pronounced empathic abilities, and some others just denote talented individuals in specific fields, such as, for example, mathematics or physics.

The scientific research about intelligence has often accepted all these folk definitions of intelligence, but it has also tried to rule out potential confounding factors. Indeed, some of those definitions might allude to other aspects of human psychology, rather than to intelligence itself, e.g., learning or creativity; or they might simply be wrong. By contrast, science is expected to utilize the proper terms at the right time.

About a century ago, with the aim of achieving a scientific definition of human intelligence, scholars described intelligence as a mental characteristic involving some sort of cognitive abilities, setting aside practical, emotional, and social skills from the picture. Psychometrics represents this scientific tradition devoted to the role of cognitive aspects of intelligence and defines intelligence as a *general mental ability*.

What type of mental ability do psychometricians refer to? At the beginning, it was thought that more intelligent people do better than less intelligent ones in simple tasks related to the discrimination of perceptual stimuli. Moreover, it was believed that this differential capability in stimuli-response was associated with academic achievements and with the individuals' socioeconomic status. But at some point, scholars understood that perceptual abilities were not the right way to account for such complex social outcomes. Intelligence has been then defined as the ability to solve cognitive tasks of various sorts, including mathematical, logical and verbal ones. These abilities, indeed, were supposed to be better related to the individual success throughout the lifespan.

For the sake of pragmatic reasons, psychometricians developed tools—intelligence or IQ tests—to quantify over individual intellectual differences. The aim of these tests was to assess individual intelligence levels, and thus to provide a single score suitable to compare individuals according to a single "dimension", namely, *general intelligence*.

Having a technical tool for categorization has had great impact on several social areas, including education, job recruiting, and clinical practices.

Although the psychometric tradition is often considered solid in its methods and theoretical models, it is nonetheless afflicted by remarkable contrasts. This is often related to the fact that psychometricians have frequently worked alongside with behavioral geneticists, who are interested in the hereditary bases of intelligence. Therefore, the debate over the psychometric approach to intelligence has not only included psychology, but also psychiatry, biology, philosophy, sociology, anthropology, and pedagogy. Controversies are often due to abuses of psychometric tests for political motivations, which have led empirical research to be frequently biased by ideologies. Often, IQ-tests scores have been conceived as *essential* properties of people depending on hereditary bases. If intelligence is highly heritable, as thought by many geneticists, then the intellectual gaps between individuals cannot be simply reduced by educational or by social countermeasures. This has been a widespread intuition during the twentieth century as described in some examples that follow.

For the advocates of the eugenics tradition, inaugurated by Francis Galton during the nineteenth century, assessing intelligence was a central step to identify the best strategy to improve the "quality" of the human species. Whenever necessary, that step was ideally followed by controlling mating. Inspired by the eugenics ideology, immigration laws and sterilization programs arose in the United States during the first half of the twentieth century, especially in relation to the works of Henry Goddard, Lewis Terman, and Robert Yerkes. However, many scholars argued against eugenics and the related genetic determinist thought by pointing at the role of environment and experiences in individuals' development.

Similar issues arose when scholars assessed intelligence in different ethnic groups, generally finding out that Western people were statistically more intelligent than, for instance, Africans. It might seem reasonable to think that Western people achieve higher IQ scores because intelligence tests are designed within the Western society, where IQ-test abilities are pivotal in high education. By contrast, in many cultural contexts, intelligence can be considered as a quite different thing, not just involving cognitive aspects. However, several authors have not considered this option as plausible; rather, they said, intelligence differences in races reflect differences in their genetic makeups.

Intelligence tests have had an important impact in education. For instance, during the 1970s Cyril Burt designed a test aimed at addressing English eleven-years-old children to specific educational paths, in accordance with their "natural" endowments. If the IQ score was too low, the kid was forbidden to access higher education in, for instance, natural sciences. However, one might think that there is something wrong with this approach to education. Indeed, even kids who obtain low IQ scores when they were eleven can eventually become expert scientists later in the adulthood: rather than "missing on intelligence", perhaps they missed good teachers, motivations, or a fertile social

environment.

Clinical applications of IQ tests have mostly been introduced around the 1940s, especially inspired by the works of David Wechsler, in relation to the study of intellectual disability. In this context, the relevance of IQ tests is especially related to behavioral genetics research: *general cognitive disability* mostly concerns low IQ and represents the behavioral phenotype for which genetic associations are sought. However, several authors have highlighted limits in the clinical utility of IQ tests. Indeed, they are unsuited for evaluating specific developmental problems, which are likely the real cause of low scoring.

As Naglieri and Das wrote in 2015, there is a considerable empirical support in favor of the concept of general intelligence as measured by IQ tests. Perhaps one of the most important evidence is the fact that IQ scores are a good predictor of school achievement. However, they say, there is circularity in this logic since the tests used to measure intelligence are remarkably similar to achievement tests. If there is not a clear distinction between mental ability and achievement, then any child who does not have an adequately enriched educational experience will be disadvantaged when assessed with a so-called "ability" test.

As the reader might notice, many of these controversies are related to two aspects: the presumed generality and high heritability of intelligence. In fact, the psychometric approach to intelligence has been at its worst in conjunction with the behavioral genetics tradition. This is the reason why I believe psychometrics and behavioral genetics cannot be analyzed separately within the intelligence debate. Rather, they are part of the same scientific enterprise.

In a nutshell, my thesis faces the psychometric-genetic theory of intelligence, according to which: a) the term "intelligence" refers to cognitive aspects, rather than to practical and social skills, or creativity; b) IQ tests approximately measure individual levels of intelligence; c) intelligent behavior is related to a single underlying general cognitive ability; d) intelligence is highly heritable and relatively stable during the lifetime; e) intelligence is related the small effect of several genes.

The controversies about human intelligence have never really been solved. No shortage of criticisms: several theories have been proposed to clarify whether intelligence is a general ability or whether it is rather a bundle of distinct cognitive phenomena; several criticisms have been raised against a strong interpretation of genetic data on the IQ; several attempts have been made to find more comprehensive definitions of this complex psychological trait. However, the psychometric-genetic theory still represents the most important framework in human intelligence studies. Alternative theories have been proposed, of course, but they all aim to contrast the main approach.

It is not my purpose to provide with my thesis a comprehensive review of the critical positions, nor of the several alternative theories available. Rather, I want to shed light on the reasons why the critics of the psychometric-genetic approach hit the mark. Let me

explain. Several scholars criticized intelligence research as regard methodological shortcomings. However, the advocates of the psychometric-genetic approach have frequently replied to critics by just improving their technical methods. Other critics have argued in favor of alternative conceptualizations of intelligence, and in doing so they radically departed from the psychometric view. For instance, it has been proposed that intelligence, instead of being related to IQ-test skills only, involves artistic, emotional, practical and social skills. Nevertheless, psychometricians have rarely reexamined their theoretical assumptions; rather, they have often argued that alternative theories of intelligence are not really theories *of intelligence*, because they look at other psychological aspects than the ones measured by IQ tests—and, after all, psychometricians seem to have a point.

Some critics are, however, optimistic. For instance, Naglieri and Das are confident that in the next fifty years intelligence research will usher a new age in our understanding, evaluating, and enhancing the intellectual development. I believe that much more work must be done for reaching such a point. In particular, we do not need an analysis limited to *what* is wrong with the psychometric-genetic approach, but we are rather in need of a detailed analysis of *why* it is wrong. I share the view that the psychometric-genetic approach does not properly account for human intelligence, and I believe there is a specific reason why criticisms about methodologies have been successful. However, methodological remarks are not enough: a theoretical analysis is required to highlight the *profound* limits of such an approach. As I argue in my thesis, this analysis must involve both biological and ontological aspects of the very nature of general intelligence, if something like this does exist.

Psychometricians barely care about biology, and frequently their approach is quite distant from the one adopted by cognitive neuroscientists. Even models coming from behavioral genetics present incompatibilities with biological and cognitive research (e.g., with developmental biology and neurobiology). In other words, both cognitive and biological sciences, with their heterogeneity, have something to say about human intelligence, and this is quite different from what the psychometric-genetic approach says. For instance, cognitive scientists do not refer to any general cognitive ability, but rather to specific cognitive processes. Biological sciences, in turn, are not directly interested in high-level phenomena like human intelligence. Therefore, what we know from, e.g., developmental biology, is unable to directly solve issues about complex psychological traits. Nonetheless, there are many things that *should be derived* from what we know about development. As I shall show, behavioral genetics rarely accounts for those aspects.

On the other hand, alternative theories of intelligence abound and thrive, often in contrast with the psychometric one. However, those theories are often neutral with respect to biological data. At present, we need a comprehensive theoretical framework for evaluating which is "the best" theory in relation to biological and ontological aspects. Thus, a major point which characterizes my thesis is that a meaningful theory of intelligence

cannot overlook biological sciences. For instance, a detailed knowledge of the role of genes in biological systems is necessary to constrain hypotheses about how genotypes bring about cognitive and behavioral phenotypes.

Another important point is that theories of intelligence should be evaluated by focusing on their ontological commitment to mental and cognitive "entities", and by taking into consideration their causal claims. What can an ontological analysis say about general intelligence? Psychometricians and behavioral geneticists rarely uncover their ontological commitment, frequently bouncing from instrumentalist to realist theories of intelligence. As a consequence, a clarification is needed to fully understand the limits of the psychometric-genetic approach to intelligence. I contend that this approach is capable of *carving the psychological nature at its joints*. As I shall argue, this is the reason why it is so vulnerable to criticisms and did not achieve a proper understanding of the biological bases of human intelligence.

In Chapter 1, I introduce in more details the psychometric-genetic (PSY-GEN) approach to the study of human intelligence, both from a historical and a theoretical viewpoint. I mainly focus on three central aspects. The first aspect pertains to the fact that intelligence is considered a *general* cognitive ability—namely, the *g* factor—instead of being composed of multiple cognitive processes. The second one regards the assumption that intelligence is a *quantitative* phenotypic trait. Finally, the third aspect concerns empirical findings in genetics research, which attest that intelligence is *largely inherited* and related to the *additive effect* of hundreds, if not thousands, of alleles. These three aspects represent the theoretical core of the contemporary theory of general intelligence.

In Chapter 2, I clarify what geneticists mean when they say that intelligence is a quantitative phenotypic trait, and thus, I analyze the distinction between quantitative and qualitative traits. I also investigate the historical roots of genetics research about intelligence, especially in relation to the wide framework of quantitative genetics. In 1918, Ronal Fisher provided a quantitative-additive model of complex phenotypic traits. This model represents the cornerstone of quantitative genetics, and it has been widely adopted by behavioral geneticists both for empirical and theoretical reasons. However, as I shall argue, the statistical approach characterizing Fisher's model cannot account for many aspects of biological systems. Conversely, developmental biology seems to represent a better framework for studying complex traits like human behaviors.

The two following chapters represent a detour from the main topic, which is necessary to clarify the risks that a statistical approach to complex traits can involve. In Chapter 3, I analyze theoretical problems in molecular research, especially the so-called 'missing heritability problem'. In Chapter 4, I focus on the heritability debate. Both these two chapters deepen theoretical assumptions in quantitative genetics and their problems. As I shall argue, a different scientific framework is required for studying human behaviors.

Chapter 5 provides guidelines for such an alternative framework. By focusing on statistical analyses, I argue, behavioral geneticists have worked against their very aim,

that is, clarifying the relationship between genes and behaviors. Doing so requires an analysis of genetic causation, which is more achievable in developmental and in systems biology. In the light of the guidelines I propose in that chapter, I argue that IQ should not be understood as a phenotypic trait in the narrow sense. Rather, it represents a general quantification over different aspects of human cognition.

In Chapter 6, I appeal to the natural kinds theory to prove an ontological argument against general intelligence. Frequently philosophy of life sciences assumes a link between scientific realism and natural kinds. Hence, introducing natural kinds allows us to offer a framework within which we can evaluate the ontological status of general intelligence. As I shall show, an ontological commitment to general intelligence requires the existence of an underlying causal mechanism capable of connecting biological, cognitive and behavioral aspects. In order to make sense of the PSY-GEN model of intelligence in an ontological manner, I detect this hypothetical causal nexus in the *g* factor, which is frequently conceived as a mental ability underpinning intelligent behavior.

In Chapter 7, I show that the psychobiological nature of the *g* factor is far from being clear, and thus it has been subject of discussions since many decades, especially at the crossroads of psychometrics and cognitive sciences. According to several authors, there is nothing like *g* in the human organism. Therefore, I suggest, theories of intelligence devoted to the role of a general mental ability should be revised or dismissed. In this final chapter, I also provide a developmental explanation of the general factor of intelligence, which does not include, however, the existence of any general mental ability.

Even if *g* does not represent a genuine biological aspect of human beings, one might ask whether we could keep it in our vocabulary. Indeed, general intelligence might have some sort of instrumental value, such as being useful for pragmatic aims related to clinical, educational, or social decisions. In the Conclusions, I shall focus on this possibility. However, I contend this is a promising approach. Indeed, the employment of the concept of general intelligence might have a negative impact on clinical and educational practices, and within the political sphere too.

Acknowledgements: I thank my supervisors Marcello Frixione and Cristina Meini for careful reading of earlier drafts of this thesis. For their comments about specific aspects of my work, I thank Diego Marconi, Cristina Amoretti, Andrea Borghini, Roberto Casati, David Ludwig, Carlo Chiorri, and Francesco Bottaccioli. This thesis has gained improvements in many respects thanks to the conversations I had with John Dupré, Staffan Muller-Wille, and the members of the Process Biology Group at the University of Exeter. My collaboration with Yukinori Onishi has played a similar, helpful role during my visiting at the Graduate University for Advances Studies, in Hayama. Last but not least, my sincere gratitude goes to Luca Costa, Flavia Fabris, and Marco Viola for being so valid past, present, and future collaborators, and to Andrea Raimondi and Elisa Frasnelli for their criticisms on both formal and substantial aspects of the thesis. The most valuable part of my work rests upon all these voices.

Chapter 1.
The Psychometric-Genetic Theory of Intelligence

Over the last century, the study of human intelligence has been mainly developed within two scientific areas: psychometrics and behavioral genetics. Although the two are driven by different research purposes, both have provided a quantitative analysis of intelligence.

Psychometrics studies intelligence in the light of two theoretical constructs: the intelligent quotient (IQ) and the general factor of intelligence (the *g* factor or simply *g*). Since IQ and *g* are strongly correlated to each other, the relative terms are sometimes treated as synonymous. Nonetheless, they should be taken as conceptually distinct. Given a population, the IQ level changes among individuals according to a bell curve. Thus, IQ stands for the individual intelligence level assessed by specific tests. Instead, *g* stands for two different things: on the one hand, it derives from a factor analysis as the outcome of correlation matrices of cognitive test scores (say, *psychometric g*); on the other hand, it is, broadly speaking, the IQ's psychological explanation (say, *psychobiological g*). In this second sense, *g* is conceived as a general cognitive ability that underlies individuals' performance to tests. According to this approach, whatever intelligence is, it is measurable by IQ tests, while *g* explains individual intellectual differences.[1]

Behavioral genetics, in turn, investigates intelligence by means of methods deriving from the genetics of quantitative traits to understand the role of inheritance in individual differences. Like other quantitative traits, such as height and weight, intelligence occurs as a continuous range of variation within populations. In order to account for these traits, geneticists adopted the additive-polygenic model of inheritance: accordingly, several genes act additively on the phenotype—in our case, on the intellectual level. For many decades, the main goal of behavioral genetics has been to understand how relevant inheritance is in the explanation of individual differences and similarities, by the adoption of the psychometric IQ as a good 'index' of individual intelligence.

Over time the purposes of both areas have converged into a model which is still influential, which Douglas Wahlsten (2002) called *the theory of biological intelligence*. After the adoption of the *g* factor by genetic research, a consensus has been reached around a model of intelligence as a highly heritable general cognitive ability. The psychometric-genetic approach (henceforth PSY-GEN) takes *g* as a prominent psychological variable (Detterman, 2002; Jensen, 2002). According to Plomin and

[1] Drawing a distinction between the psychometric *g* and the psychobiological *g* makes sense of a fundamental conceptual difference between two possible ways of conceiving the *g* factor. Unfortunately, as I will show, the distinction between the two is not always emphasized within the scientific literature.

colleagues (2013, p. 187), *g* is one of the most reliable and valid measures in the behavioral domain.

The PSY-GEN model of intelligence roots in a long tradition, roughly started with the work of Francis Galton and then developed by authors from several research areas, ranging from psychology to biology. Although this tradition is often considered solid in its methods and theoretical models, it is nonetheless afflicted by remarkable contrasts. Several theories have been proposed to clarify whether intelligence can be conceived as a general ability or as a bundle of distinct cognitive phenomena. Several criticisms have been raised against a strong interpretation of genetic data on the IQ. Several attempts have been made to find more comprehensive definitions of this complex psychological trait.

So, long story short, the psychobiological meaning of *g* is today still controversial. This is reflected by the fact that neither intelligence nor the *g* factor has been allocated a clear position, if any, in the agenda of cognitive sciences. Since the IQ test skills clearly belong to the cognitive domain, this divergence might strike one as surprising. Nevertheless, the attempt to bridge psychometrics and cognitive sciences is relatively recent—especially considering that the psychometric approach to intelligence was born in the early twentieth century. As I shall show, the two traditions of psychometrics and cognitive psychology landed to quite different views of intelligence.

It is worth noting that there has been a constant dialogue between psychometricians and geneticists throughout the last century. This dialogue has been important for two reasons, at least. The first one refers to methodological concerns. Psychometrics managed to quantify over human intelligence as a unified phenomenon. This is related to specific theoretical viewpoints on the nature of intelligence and to social, educative or clinical goals—which depend on the historical moment. This has led, in turn, to the possibility of studying intelligence by means of quantitative genetics analysis. That is, if intelligence is a general cognitive ability, then it is possible to study it as a unified phenotypic trait; conversely, if intelligence was a compound of different entities related to each other in a complex manner, it would be very hard to study it in the way several authors aim to (why scholars study intelligence in such a way? As I will explain, the reasons depend also in this case on the author's goals).

The second reason why the dialogue between psychometricians and geneticists has been important regards ontological aspects. The ontological reality of general intelligence has been above all attested in relation to its biological bases. Roughly speaking, it has been assumed (explicitly or not) that, if general intelligence has strong biological correlates (often genetic correlates), then it must be "real" in some sense.

In this chapter, I introduce the most important aspects of the PSY-GEN theory of intelligence. Such aspects regard three central properties attributed to human intelligence. The first one pertains to the fact that intelligence is considered a general cognitive ability. I analyze the historical origins of IQ and *g* factor. In this respect, I delineate what beliefs and aims led scholars to develop psychometric tests to assess intellectual individual

differences. I also explain what theoretical models have been proposed to accommodate tests data. Furthermore, I deepen why, and for what purposes, psychometricians have conceived intelligence as a general phenomenon instead of as composed by many distinct cognitive processes. I also introduce some worries that arose when psychometrics met the newer psychological tradition that led to cognitive sciences around the 1980s.

The second central aspect of the PSY-GEN theory of intelligence regards the widespread assumption that intelligence is a quantitative phenotypic trait. This is particularly important insofar as it makes sense of the employment of quantitative methodologies for the study of the intelligent behavior. Consequently, the third aspect concerns empirical findings and controversies related to the application of genetics methods like heritability analysis and molecular techniques. These findings seem to attest that intelligence is largely inherited.[2]

Down the chapter, I introduce the major methodological and theoretical issues highlighted by the critics of the PSY-GEN approach to intelligence, which will be further explored in the following chapters. I conclude this chapter by submitting that the problem of intelligence is an ontological one, involving both philosophical, psychological, and biological explanations. As I shall show, by working under the aegis of statistical analyses, the PSY-GEN approach has arguably missed its own achievements.

1. Psychometric Intelligence

In the first part of this chapter, I analyze the historical origins of the IQ and the *g* factor. In §1.1, I delineate what beliefs and aims led scholars to develop psychometric tests to assess intellectual individual differences. In §1.2, I show how complex the relationship between tests and theories of intelligence may be. In §1.3, I deepen why psychometricians assumed that intelligence is a general cognitive phenomenon. In §1.4, I oppose the single-factor to the multiple-factor theories of intelligence, introducing also a sort of mismatch existing among psychometrics and cognitive psychology.

[2] We should, however, bear in mind that both psychometrics and behavioral genetics are quite heterogeneous scientific fields; it will be hence necessary to simplify some aspects of the PSY-GEN approach to intelligence to embrace as many viewpoints as possible. About behavioral genetics, I mainly refer to Plomin and colleagues (2001, 2013), representing a synthesis of the genetic approach to behaviors. About psychometrics and cognitive sciences, I frequently appeal to the contributes included in Sternberg & Grigorenko (2002), Sternberg & Pretz (2005) and Goldstein, Princiotta & Naglieri (2015)—three companions which summarize contrasting positions among psychometricians and cognitive scientists.

1.1. Testing intelligence: a brief history

The psychometric tradition has developed various methodologies to quantify over people's intellectual behavior. Albeit practical applications have largely changed depending on the social circumstances, the chief aim of testing intelligence has ever been to *measure*, while IQ represents a number useful in sorting or *categorizing* individuals according to their intellectual capability. Most intelligence tests refer to an underlying psychological construct which is the real object of interest, namely general intelligence. Roughly speaking, IQ is supposed to represent intelligence as assessed by tests.[3]

Having a technical tool for categorization has had implications across several social areas, from education to job recruiting. One of the most important aspects of contemporary psychology concerns the clinical impact of IQ tests, which is related to the study of intellectual disability. The most recent versions of DSM (*Diagnostic and Statistical Manual of Mental Disorders*) categorize intellectual disability as a clinical picture related to various diagnostic criteria, among which the first pertains to low IQ levels. Other criteria address adaptive functioning for social standards and intellectual and adaptive deficits during development, too. So, the relevance of IQ testing especially arises for behavioral genetics research. In fact, geneticists are not interested in the clinical condition itself, but rather in the so-called "general cognitive disability", which mostly concerns low IQ (see Plomin et al., 2013, p. 163).

However, testing intelligence served to quite distant goals over the past century, especially within the eugenics tradition inaugurated by Francis Galton at the end of the nineteenth century. The biometric approach introduced by Galton was not solely linked to epistemological concerns, e.g., discovering what intelligence is. Rather, it was related to understanding the inheritable bases of individual intellectual differences.[4]

The modern history of intelligence officially starts with Galton himself, who is remembered as the first who tried to study intelligence rigorously and empirically (at the time, no such things as IQ or intelligence tests existed).[5] The author pursued many types

[3] While I first focus on IQ tests, I move towards the problem of defining intelligence only at a later stage. I explain the reasons for this choice in §1.2.

[4] Within the eugenics tradition, assessing intelligence was a central step to identify the best strategy to improve the "quality" of the human population. That step was ideally followed by sorting people according to their natural intellectual gifts and, whenever necessary, by controlling mating. Immigration laws and sterilization programs arose in the US during the first half of the twentieth century, inspired by eugenics principles (see Eysenck & Kamin, 1981; Greenwood, 2015; Gould, 1981; Rose et al., 1984). Clinical purposes have mostly been adumbrated after the 1940s (Benisz et al., 2015). During the 1970s, an important relapse of tests was in education—e.g., the 11+ test, proposed by C. Burt, was aimed at addressing children to their "natural" educational path (see Eysenck & Kamin, 1981). Several scholars assessed differences among ethnic groups (e.g., Herrnstein & Murray, 1994; for criticisms, see Block, 1995; Cole, 2006; Cole et al., 1971). The importance of IQ tests in educational and clinical context is nowadays still present (Kaufman et al., 2013; Plomin et al., 2013). For some reviews, see Callier & Bonham, 2015; Panofsky, 2014.

[5] Galton was passionately interested in measuring everything to such an extent that in 1874 he expressed his belief that any aspect is describable as a parametric value (see Boakes, 1984; Richardson, 1999). It was

of research on intelligence: by means of craniometrics methods, he tried to find correlations among cranial dimensions and intelligence; another type of research was aimed at analyzing the relationship between sensory discrimination and intelligence. For both craniometrics and sensitivity tests, professional success and academic achievement were assumed as indexes of intelligence.

"Galton (1883) believed that smarter people have more acute senses (sight, hearing, touch, etc.), notice things more, and, having more information available to them, are better able to compete and succeed. He, therefore, created several tests of psychological sensitivity, such as one called 'weight discrimination', [where] examinees were blindfolded, given three identical objects, and were required to arrange the objects in order of increasing weight. A person's sensitivity to weight was determined by the finest difference among the three weights he could discriminate" (Cianciolo & Sternberg, 2004, p. 33).

The first proper reference to mental testing dates back to 1890, when James Cattell, a student of Galton, published *Mental Tests and Measurements*, bringing Galton's practices in the United States. The tests battery introduced by Cattell included both mental and physical tasks—e.g., physical strength, the speed of movement, sensory capacities, reaction time, and memory—in the light of the Galtonian hypothesis according to which sensory discrimination accounts for intelligence.

Both the methodologies proposed by Galton and Cattell could not achieve the expected results. Galton's theory turned out to be almost completely wrong (Feldman, 2015, p. 270): the correlations between sensory abilities, cranial dimensions, and professional success are very weak.[6] This also applies to Cattell's methods of assessing the correlations between academic achievements and sensory discrimination: a large-scale evaluation of Cattell's tests (Wissler, 1901) did not produce positive results (see Cianciolo & Sternberg, 2004; Greenwood, 2015).

A general dissatisfaction shrouded the attempts to measure intelligence until Alfred Binet, director of the laboratory of psychology at the Sorbonne, was recruited in 1904 by the French Minister of Education. He was asked to develop a practical guide for identifying children whose poor performance indicated a need for special education. Binet thought that, on average, older kids can solve harder problems than younger children. Therefore, his test contained a series of increasing difficulty tasks. For instance, if Pierre can solve 8-years-related problems—but not the 9-related ones—, Pierre's mental age is 8. The intellectual level was calculated by subtracting the kid's mental age from her chronological age, and a special educational program was to be planned for those children

presumably this passion for numbers that led Galton to develop remarkable statistic methodologies which made him the father of biometrics.

[6] The quest for brain size-IQ correlations is still pursued (see Chapter 7). However, as Cianciolo & Sternberg notice (2004, p. 12), "it is unclear […] whether brain volume should be considered a cause of greater intelligence or whether factors giving rise to greater intelligence, such as having experienced a larger set of intellectually demanding events, contribute to greater brain volume (e.g., see Garlick, 2002). In any case, the association between brain volume and intelligence appears weak enough to justify searching in other places for the biological basis of intelligence". See also Gould (1981).

for which a significant gap between the two ages was attested (see Binet, 1909).

According to Binet and his collaborator Theodore Simon, testing simple psychological processes, like Galton did, is a waste of time. Rather, they related intelligence to a set of judgment skills, including being able to direct one's thought to the steps that must be taken to complete a task, to adapt one's strategy during task performance, and to accurately monitor one's performance. To assess these judgment skills, Binet and Simon tested higher level cognitive functioning, such as verbal skills and social comprehension (see Binet & Simon, 1905).

Before the passing of Binet in 1911, three versions of the so-called *Binet-Simon Scale* were published. In 1912, William L. Stern proposed the term 'Intelligence Quotient' for replacing the concept of level introduced by Binet. Stern proposed that mental age was to be divided by (instead of subtracted from) chronological age and that the result was to be multiplied by 100 to avoid decimals. Now, those children on the average had an IQ = 100, while retarded children obtained a score below 100.

The prototypic form of psychometric test we use today was born around 1916, when Lewis Terman published the so-called *Stanford-Binet Scale*, using Stern's IQ for the first time as an index of individual intelligence. That test was mainly addressed by Robert Yerkes to army recruitment during the First World War, leading to a major shift in intelligence testing. In order to meet the United States Army's demands for the rapid testing of a large number of men, this test was briefly presented with written items instead of more complicated tasks requiring detailed instructions; the judgment of an examiner was replaced by right/wrong scoring techniques; time limits for test completion was imposed; and test problems appropriate for adults were developed. The Army Alpha—an adaptation of the Stanford-Binet scale provided by Arthur Otis—assessed reasoning skills and cultural knowledge. Due to the high rates of illiteracy among military recruits, the Army Beta was designed to assess intelligence without requiring knowledge of English language; pictorial instructions were used. The Army Beta assessed perceptual speed, memory, and reasoning with pictures. Both the Army tests provided a single score for each participant.

A major contributor to intelligence testing was David Wechsler, who developed a series of intelligence scales since 1939. This individually administered scales featured both verbal and performance tests, reflecting Wechsler's belief that intelligence is expressed in both verbal and nonverbal ways. Wechsler created tests for both adults (Wechsler Adult Intelligence Scale, WAIS) and children (Wechsler Intelligence Scale for Children, WISC). These tests are aimed at measuring verbal comprehension, perceptual reasoning, working memory, and processing speed (see Table 1.1).

It is worth noticing that many subtests included in WAIS derive from preexistent tests, e.g., from the Stanford-Binet and the Army Tests (Benisz et al., 2015; Kaufman, 2009; Wechsler, 1939) (see Table 1.2).

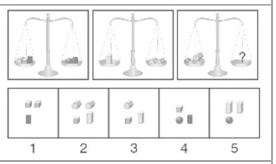
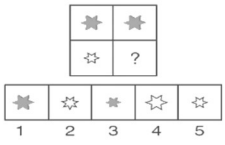
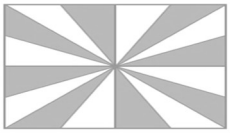
Name	Example
Information	Who wrote *Tom Sawyer*?
Comprehension	Why is copper often used for electrical wires?
Arithmetic	Three women divided 18 golf balls equally among themselves. How many golf balls did each person receive?
Similarities	In what way are a circle and a tringle alike?
Figure Weight	Problems require test-taker to determine which possibility balances the final scale. 
Matrix Reasoning	Test-taker must decide which of the five possibilities replaces the question mark and completes the sequence. 
Block Design	Problems require test-takers to reproduce a design in fixed amount of time. 

Table 1.1: Typical items of the Wechsler Adult Intelligence Scales (WAIS-IV).
From Feldman, 2015, p. 273.

Verbal Tasks	
Information	Army Alpha
Digit Span	Stanford-Binet
Vocabulary	Stanford-Binet
Arithmetic	Stanford-Binet
Comprehension	Stanford-Binet
Similarities	Stanford-Binet

Performance Tasks	
Picture Completion	Army Beta
Picture Arrangement	Army Performance Scale
Block Design	Kohs' Blocks Tests
Object Assembly	Army Performance Scale
Digit Symbol	Army Beta

Table 1.2: Historical origins of the subtests included in WAIS.

Wechsler also introduced a new concept of IQ. As it has been said, Stern-IQ was a ratio between mental and chronological age. This index of individual intelligence, albeit initially successful, did not allow a proper comparison between individuals at different age stages. Such a limitation led several authors to look for a newer definition of IQ.

For assessing IQ in adults with his WAIS, Wechsler proposed an alternative concept of IQ, that is, an index of the individual intellectual efficiency within the peer group. Wechsler-IQ is determined by converting raw scores into a normal distribution of scores in a population—a bell curve with a mean of 100 and a standard deviation of 15. This process is called 'normalization'. Statistically speaking, an individual who obtains an IQ

score of 100 will have as many individuals "better" than her as "worst". By means of Wechsler-IQ, it was possible to categorize any individual, regardless her age (see Benisz et al., 2015; Kaufman, 2009).

With the work of Wechsler, we land to contemporaneity: his revised and expanded scales have been widely used and are today still employed (Benisz et al., 2015; Cianciolo & Sternberg, 2004, p. 36). In fact, the contemporary versions of WAIS and WISC are the most widespread tests in the United States (Feldman, 2015). Furthermore, Wechsler-IQ represents the notion of IQ we employ today. Of course, over time Wechsler revised some aspects of his scales to reflect more recent psychological theories which he embraced later. However, some fundamental aspects of the original generalist model are today as they were a century ago (see §1.3).

In §1.2, I analyze some technical details about psychometric methods and I introduce the distinction between instrumentalist and realist positions about the psychometric approach to intelligence.

1.2. Test theory: just instrumentalism?

After the 1930s, the development of IQ tests has been characterized by a more systematic evaluation of their accuracy, adequacy, coherence, and consistency. Indeed, intelligence theories began to multiply, so did the need to evaluate them (Cianciolo & Sternberg, 2004, p. 39). However, the relationship between test practice and theories of intelligence has never been easy to deal with. According to contemporary test theory, psychometricians must deal with technical concerns such as tests standardization, reliability, and validity:

> "The fundamental goal of test theory is to inform test development […]. Test developers can be more certain that they are measuring the intellectual abilities they think they are measuring and can, therefore, draw more valid conclusions based on test results" (Cianciolo & Sternberg, 2004, p. 38).

As a start, IQ tests are not universally applicable: they are rather standardized for specific populations. Every population for which an IQ test has been designed, the IQ scores follow a bell curve: the average value is 100; the standard deviation is 15. The standardized score put an individual in a specific position in relation to other individuals of the population. There is nothing more than this in tests because they are designed and fixed for the precise aim of generating a normal distribution (see Lewontin, 1982; Rose et al., 1984).[7]

The standardization of an IQ test is however just a first step towards the aim of

[7] Once a test is standardized for a population, and IQ is normalized, it is possible to generalize the obtained results to a wider population. This requires that the examined population is in fact representative of the target population—and this should be demonstrated.

measuring intelligence; to ensure that IQ represents an individual's level of intelligence, several further steps are necessary. In sum, one must convert raw test scores into something which is psychologically meaningful. A raw score is not an absolute measure: its meaning refers to the specific population for which the test has been standardized. So, an IQ score does not mean anything without a specific control, that is, both internal and external criteria are required to attest that it represents the psychological variable under examination, i.e., intelligence.

In this respect, it is important to account for construct validity, content validity, test reliability, and external validity. The construct validity pertains to the fact that a test measures what it is supposed to measure, and only that—a test should not include abilities outside the psychological construct under examination. The content validity ensures that a test adequately samples the aspects of the ability one wishes to measure. The reliability of a test is essentially its coherence and consistency: the items must represent a coherent sampling of the intellectual ability of interest (they all measure the same ability), and the results of the test should be relatively consistent over repeated administrations (the rank order of people's scores should not fluctuate greatly). The external validity represents a match between test performance and external criteria generally associated with the ability one wants to measure.[8]

Psychometricians know how important is for tests to be in a strict relationship with well-developed theories. Otherwise, there is the risk of distorting results by overestimating the suitability of technical tools at the expense of the underlying theories. The problematic side of the coin lies exactly in this relationship.

It is for the sake of a precise reason that I decided to start my discussion by talking about intelligence tests instead of asking what intelligence is: if one looks at the history of intelligence, one might notice that the problem of measuring intelligence has been raised before the problem of defining intelligence. As Cianciolo and Sternberg highlight,

> "the boom in intelligence-test construction following the First World War reflected the promise that tests held for matching educational and occupational opportunities to people with particular intellectual capabilities. Although test developers had sometimes quite detailed notions of intellectual capability, tests at this time generally were not created to extend scientific thinking about intelligence. Perhaps for this reason, test developers during this period did not appear to question very frequently whether the tests they had created actually measured what they were supposed to. These tests demonstrated notable practical utility for occupational and educational placement, so it may have been of secondary concern whether they could be considered valid measures of some specific, theory-based notion of intelligence" (Cianciolo & Sternberg, 2004, pp. 37-38).

The fact that testing intelligence has somehow preceded the definition of intelligence leads us to remarkable issues concerning the reliability of the PSY-GEN theory. Of course,

[8] For instance, Galton looked at professional success for this role. Binet, instead, worked in relation to teachers' opinion of their students: if a child obtains a very low mental age, but her teachers think that she is very smart, then the test does not meet the external validity requirement (see Lewontin, 1982).

some notion about intelligence was available before the invention of tests. For instance, Galton and Cattell thought that intelligence might be defined as sensory-motor discrimination and reaction time. Thorndike supposed that an individual's level of intelligence is determined by the number of connections the individual can make (see Greenwood, 2015). Binet and Simon assumed a relationship between intelligence and judgment skills. However, the problem of what intelligence really is has played a secondary role in the psychometrics tradition; rather, the notion of intelligence was pre-theoretical or, at most, instrumental.

"We did not start with a clear definition of general intelligence [but] borrowed from every-day life a vague term implying all-round ability and […] we [are] still attempting to define it more sharply and endow it with a stricter scientific connotation" (Pintner, 1923, p. 53).

"The first tests of intellectual development, those imagined by Binet or Wechsler, were not based on very elaborated theories of intelligence. The approach of these pioneers of psychometrics was, of course, inspired by some general ideas on intelligence, but the way in which they searched for tasks likely to measure it was very empirical. Binet, for example, tried various items and retained those that discriminated well between mentally retarded and non-retarded children, between older from younger children, and good from not so good students. The construction of tests was guided by their empirical validity, in particular relating to criteria like academic performance, more than by their theoretical validity" (Lautrey, 2002, p. 117).

Roughly speaking, everything was for the sake of assessing intelligence, even approximately, to serve some practical aim. This may explain why, in several situations, intelligence has been defined as "what intelligence tests measure" (definition well known as 'the Boring's tautology'). This operative definition led many scholars to think of IQ and intelligence as coinciding with each other. As Naglieri says,

"there was no theory of intelligence that guided the selection or development the Army Alpha and Beta tests. These tests have been accepted as measures of intelligence and in fact the IQ score has become synonymous with the term intelligence" (Naglieri, 2015, p. 313).

However, psychometricians tend to defend such an operative definition. For instance, Hans Eysenck writes:

"Psychologists, when asked what intelligence is, sometimes say, with tongue only partly in cheek, that it is what intelligence tests measure. This often produces amusement among listeners not trained in science, for it seems to be nothing more than a tautology. However, in science definitions of this kind—so-called *operational definitions*—are quite common; indeed, many scientists believe they are the only kind of scientific definition which is acceptable. You define a concept in terms of the ways in which you measure it and the measurements achieved. This is not tautological because the measurements are derived from a theory and can be used to verify or invalidate it. The statement that intelligence is what IQ tests measure is not circular because it stands to be disproved by IQ measurements themselves" (Eysenck & Kamin, 1981, p. 25).

To rephrase Eysenck's thought, the notion of intelligence has been taken instrumentally, rather than ontologically. This instrumental approach is often adopted by psychometricians, but unfortunately there is a tendency to shift towards a sort of realist account. In other words, one might distinguish between the *psychometric* and *psychobiological* intelligence: the former is an abstract entity, just hypothesized for the sake of approaching the intelligence problem without any explanatory purpose about "how things work"; the latter expresses an ontological commitment about general intelligence. This distinction reflects an important disagreement in psychological research about what intelligence really is.

It is important to notice that, far from being just a scholastic and pointless controversy, this disagreement includes test practice. Many of the intelligence tests currently employed for diagnostic and placement purposes date back to the early twentieth century. Even though they have been revised to a less or greater extent, they still marry the core features of the original psychometrics viewpoint, which holds that intelligence is a high-level general ability that people have to a "different degree": those who have "more" of it have also an advantage in most tasks over those who have "less" (for similar concerns, see Cianciolo & Sternberg, 2004, pp. 53-54).

Other theories proposed in recent decades lead to the view that intelligence is not completely captured by the dominant IQ tests, or that general intelligence does not represent a good posit for psychological research.

Everything will be clear in relation to Spearman's theory, which represents the core of the PSY-GEN approach to intelligence: in a sense, as I shall show, Spearman's *g* represents indeed its ontological foundation.

1.3. The general factor of intelligence

As it has been mentioned, psychometricians study intelligence by means of several constructs, among which the most important are the IQ and the *g* factor. IQ represents the individual intelligence level assessed by tests. Instead, *g* stands for two different things: on the one hand, it derives from a factor analysis as the outcome of correlation matrices of test scores (psychometric *g*); on the other hand, it is, broadly speaking, the psychological explanation of the results of tests (psychobiological *g*).

Historically speaking, the existence of a general factor of intelligence has been first hypothesized by Charles Spearman (1904, 1923) in parallel with the rising of the test tradition. The reason why the *g* factor has been assumed is straightforward: intelligence measurements are positively correlated to each other. Although to varying degrees, if one shows good performance on a given task, one tends to show good performance also in other tasks. This empirical phenomenon is called 'positive manifold'. Thus, *g* is a summary index of a correlation matrix, representing what cognitive tests have in common

and explaining ~40% of their variance (Plomin et al., 2013, p. 210). In this respect, *g* is relatively uncontroversial. The subject of the controversies lies in the psychobiological nature of *g*.[9]

Before proceeding towards issues related to the psychobiological *g*, it is worth summarizing the reasons why the psychometric *g* raised. I opted to treat *g* separately from IQ because *g* is not strictly related to tests. Rather, its origins lie in the factor analysis, which is a statistical method that was born at the end of the nineteenth century to achieve data simplification.

Galton first raised the issue of how to reduce the number of the anthropometric measures of a large sample of criminal profiles. Francis Edgeworth (1893) solved the issue by calculating linear orthogonal functions of those variables suitable to summarize information within a smaller range of variables, much easier to deal with. So, the aim of factor analysis was to simplify the interpretation of many variables (see Di Franco & Marradi, 2003). During the first decades of the twentieth century, Spearman inaugurated an innovative factorial method capable of identifying a general factor which connects individual tests scores, reducing in such a way data complexity. In particular, Spearman found that scores on all mental tests (regardless of the testes domain) tend to load on one major factor.

However, Spearman's analysis was not solely aimed at simplifying data, but also at corroborating theoretical hypotheses. That is, Spearman had in mind a specific conception of intelligence before interpreting data. In such a way, factor analysis assumed a *corroborative role* which several psychologists (e.g., Cattell, Thurstone, and Eysenck) considered attractive for its ability to ground theories on data (see Di Franco & Marradi, 2003, p. 19). Long story short, Spearman managed, at a stroke, to summarize information about IQ tests in a single variable and to justify the existence of a general cognitive ability, which he called *g*.

As it has been said, the subject of controversy about the PSY-GEN approach lies in the psychobiological nature of *g*. The advocates of PSY-GEN do not attempt to understand *g* in any strong ontological sense. As it is often the case in psychometrics, a clear-cut distinction between methodological purposes and the reality of a psychological construct is endorsed. In other words, it does not pertain to psychometrics to explore *g* ontologically; it is sufficient to ensure that IQ tests can evaluate intelligence—whatever intelligence is.

Nonetheless, in the light of the positive manifold several psychologists have accepted the existence of an underlying general mental ability (see Garlick, 2002; Van der Maas et al., 2006). Hence, some ontologically driven hypotheses have been adumbrated,

[9] Some criticisms have been raised in relation to technical aspects of factor analysis (e.g., Gardner, 1983; Gould, 1981). In fact, this statistical method may lead to a *g* factor or not depending on methodological choices. Although most scholars hold that *g* is a well-established psychometric entity, those critical appraisals are worth recalling despite the general tendency in the PSY-GEN framework is to simply to overlook them (e.g., Plomin et al., 2001). I deal with criticisms about the psychometric *g* later.

especially when psychometrics came face to face with biological sciences. The advocates of *g* conceive this factor as a cognitive phenomenon responsible for individual differences in test performances.[10]

The degree of the realist commitment is different depending on the authors' standpoint, but most of them admit *g* as a psychobiological characteristic, or as a neural mechanism, which influences intelligent behavior. For instance, Spearman described *g* as a form of mental energy. More recently, to guarantee that *g* is a valid measure of intelligence, some scholars have tried to relate it to more reliable constructs, like the ones coming from cognitive neuroscience—e.g., working memory, processing speed, neural efficiency and brain size (for some reviews, see Cianciolo & Sternberg, 2004; Gray & Thompson, 2004; Pretz & Sternberg, 2005; Williams et al., 2008).

Briefly, *g* must exist somehow: we can look for its biological correlates to ground it in other cognitive phenomena which seem to exist since they do not derive from mere statistical research.

By contrast, several authors cast doubt on a strong interpretation of *g*. Humphreys and Stark, although psychometrician, do not accept *g* as a fixed biological capacity, arguing against reification:

"The *g* factor is interpreted much too freely as an entity, such as a fixed capacity, by psychologists and people in general. Spearman started the reification of the general factor in describing his own research. He defined intelligence as 'mental energy'. It seems to us that Jensen also reifies *g*. We do not agree with Jensen's concentration on neural correlates" (Humphreys & Stark, 2002, p. 98).[11]

About reification, it is worth noting that Binet and Simon themselves did not suppose that they were measuring a unitary capacity, far less one that is innately determined.

"Binet and Simon were careful to stress the limitations of their scales, given their belief in the malleability of intelligence and the inherent margin of error. However, their cautious approach was discarded when Goddard and Terman brought the Binet-Simon scales to the United States—both followed Galton, Cattell, and Spearman in supposing that intelligence was a unitary ability, which was largely determined by heredity" (Greenwood, 2015, pp. 129-130).

Binet warned posterity about the risks involved in psychometric research. As Gould noticed,

[10] It is worth drawing the distinction between factor analysis and principal components analysis. As Van der Maas et al. (2014, pp. 1-2) summarize, "the factor model is a reflective latent variable model, in which the factor is a hypothesized entity that is posited to provide a putative explanation for the positive manifold. The principal components model is a formative model, in which the components are conveniently weighted total scores; these are composites of the observed data, which do not provide an explanation of the positive manifold, but rather inherit their structure entirely from the data. Thus, the factor model embodies the idea that there is a common cause 'out there' that we 'detect' using factor analysis, and that should have an independently ascertainable identity in the form of, say, a variable defined on some biological substrate".

[11] The reification of *g* is stronger in those authors who emphasize the role of biological sciences. For loose and strict interpretations of *g*, see Kray & Frensch (2002).

"Binet declined to define and speculate upon the meaning of the score he assigned to each child. Intelligence, Binet proclaimed, is too complex to capture with a single number. This number, later called IQ, is only a rough, empirical guide constructed for a limited, practical purpose: The scale, properly speaking, does not permit the measure of the intelligence, because intellectual qualities are not superposable, and therefore cannot be measured as linear surfaces are measured (1905, p. 40). Moreover, the number is only an average of many performances, not an entity unto itself. Intelligence, Binet reminds us, is not a single, scalable thing like height. 'We feel it necessary to insist on this fact' Binet cautions, 'because later, for the sake of simplicity of statement, we will speak of a child of 8 years having the intelligence of a child of 7 or 9 years; these expressions, if accepted arbitrarily, may give place to illusions.' Binet was too good a theoretician to fall into the logical error that John Stuart Mill had identified—to believe that whatever received a name must be an entity or being, having an independent existence of its own" (Gould, 1981, p. 181).[12]

Authors from cognitive sciences are even more adamant. According to Kray and Frensch (2002), there is no convincing empirical evidence that supports the existence of *g*. For Stankov (2002, p. 35), there is no single cognitive process that can explain the presence of *g*: rather, it is a mixture of many different processes (including non-cognitive influences) that change during development (see also Hampshire et al., 2012; Naglieri & Das, 2002; Ramus, 2017; Van der Maas et al., 2006, 2014).

In the light of these concerns, many scholars developed tests and theories suitable to account for the non-generality of intelligence. I introduce them in the next paragraph.

1.4. Multiple-factor theories and the role of cognitive sciences

The disagreement about the generality of intelligence is sometimes known as *the one-many problem* (Brody, 2000; Furnham, 2015). Roughly speaking, this controversy opposes two different types of theories of intelligence: on one side, the single-factor theories, according to which intelligence is a general cognitive ability; on the other side, the multiple-factor theories, where intelligence represents the compound of several cognitive abilities, from which general intelligence arises as an abstract entity.[13]

The single-factor theory can be traced in Spearman's thought and in the following hereditarianism tradition. It is also traceable in Wechsler's work: both WAIS and WISC measure intelligence as a unified phenomenon and produce a single IQ score which is supposed to represent general intelligence.

---

[12] Binet also rejected the idea of an 'innate intelligence' by highlighting how kids improved their levels whenever subject to special educative forms: "the point of such courses was to *increase* the intelligence of children who had scored low on IQ tests. Binet's attitude is clear: he firmly rebuked those who believed that 'the intelligence of an individual is a fixed quantity'" (Eysenck & Kamin, 1981, p. 91).

[13] This debate looks like the modularity of mind problem, albeit the connection between the two should be clarified. Garlick (2002) discusses modularity and the intelligence debate conjointly. Anderson (2005) proposes a theory of intelligence which adopts Fodor's distinction between central processes and dedicated processing input modules. Burkart et al. (2017) interpret the *g* factor as a domain-general cognitive mechanism compatible with the presence of domain-specific modules. See Chapter 7 for more details.

Within this generalist framework, the psychobiological reality of *g* has been stated by minimizing the role of domain-specific cognitive abilities. Actually, Spearman adopted a two-factors view, concerning both *g*, which intervenes in every task, and *s*, which intervenes in specific cognitive tasks.

"Spearman claimed that *g* is a single mental capability measured by all intelligence tests and that it is some form of generalized mental energy. Specific abilities are capabilities uniquely measured by a particular mental test, for example, mathematical computation. Spearman was interested primarily in what is common among various types of intellectual abilities, rather than in what makes each one unique. He believed that specific abilities do not capture the essence of intelligence and instead proposed that important differences in people's mental test scores are due to just one intellectual capability, mental energy" (Cianciolo & Sternberg, 2004, p. 3).

In the light of Spearman's theses, over time the focus has turned more on the general factor than on specific factors, leading to the view that specific abilities play a secondary role insofar as they are strongly influenced by *g*. In this respect, the advocates of the single-factor view often propose hierarchical models aimed at describing how *g* influences other cognitive abilities (for some review, see Jensen, 2002; Kray & Frensch, 2002; Schneider & Flanagan, 2015). For instance, Carroll's Three-Stratum Theory (1993) sets *g* at the top of the pyramid. As another example, Horn and Cattell's theory (1966) presents nine abilities at the top of the hierarchy, but the most important is fluid intelligence, often equated with Spearman's *g* (see Cianciolo & Sternberg, 2004, p. 7; Gray & Thompson, 2004) (see Figure 1.1).

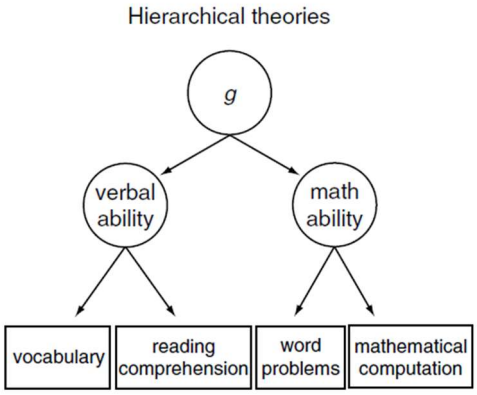


Figure 1.1: Graphical representation of a generic hierarchical theory of intelligence.
From Cianciolo & Sternberg, 2004, p. 5.

The revised versions of Wechsler's tests reflect some aspects of the hierarchical theories of intelligence. Even though these tests seem to accept the importance of specific cognitive processes, they are mostly addressed to different aspects of the same phenomenon—Wechsler believed that different aspects of intelligence can be measured

by means of different subtests (see Benisz et al., 2015).

Even if many hierarchical models take into account both general intelligence and specific abilities, *g* often represents the real explanatory entity and the depositary of the causal efficacy on intelligent behavior. In this respect, I shall contrast single-factor theories with multiple-factor theories by not considering hierarchical theories a separate category. Indeed, hierarchical theories are different to each other in relation to the commitment about *g*. In other words, hierarchical theories can imply either a generalist or a "multiple" view of intelligence (see Chapter 7 for more details).

Multiple-factor theories do not accept the reliability of *g* or underestimate its explanatory power. The advocates of this view focus their attention on the so-called group-factors, that in factor analysis are shared by some tests only. In multiple-factors tests, the variety of the items may depend on the specific idea one has in mind about intelligence; generally, researchers "extract" several factors of intelligence.[14] Here, *g* is supposed to be an emergent phenomenon whose ontological reality it is not worth assuming. Briefly, what distinguishes the two theories is the ontological commitment about general intelligence.

Multiple-factors theories have a long tradition which originates from the work of Godfrey Thomson and, especially, of Louis Leon Thurstone. Thomson (1939) proposed that *g*, instead of being a sort of mental energy, consists of many different intellectual capabilities, plus skills and motivation, which operate simultaneously in tasks solving (see Cianciolo & Sternberg, 2004, p. 3). Thurstone was attracted by Spearman's factor analysis but believed that no such thing as a general intelligence really exists. He argued that *g* was a statistical artifact resulting from the mathematical procedures used to study it. In order to corroborate such an intuition, he developed a variant factor method aimed at ruling out the correlations between the minor factors and to summarize data form IQ tests by means of the group-factors only.[15] This new methodology allowed Thurstone to detect several factors instead of a general one. The resulting analysis was less synthetic but, conversely, much more information was retained: seven factors of intelligence were clearly identifiable—the so-called *Primary Mental Abilities*, PMA (Thurstone, 1935, 1938).[16] These factors were neither general across all tests nor specific to each test. In other words, those abilities are involved in many tasks (see Schneider & Flanagan, 2015).

[14] Examples of multiple-factors tests are the PMA (Thurstone, 1938), the CAS (Naglieri & Das, 1997), and the KAIT (Kaufman & Kaufman, 1983).

[15] Thurstone inaugurated the practice (nowadays still used) of rotating the factor axes in their geometric representation. In such a way, the axes get close to the vectors which are mostly correlated with each other. The graphical representation introduced by Thurstone makes it easier to identify several principal components that represent the variables that are more alike with each other (see Gould, 1981).

[16] Word fluency, verbal comprehension, spatial visualization, number facility, associative memory, reasoning and perceptual speed.

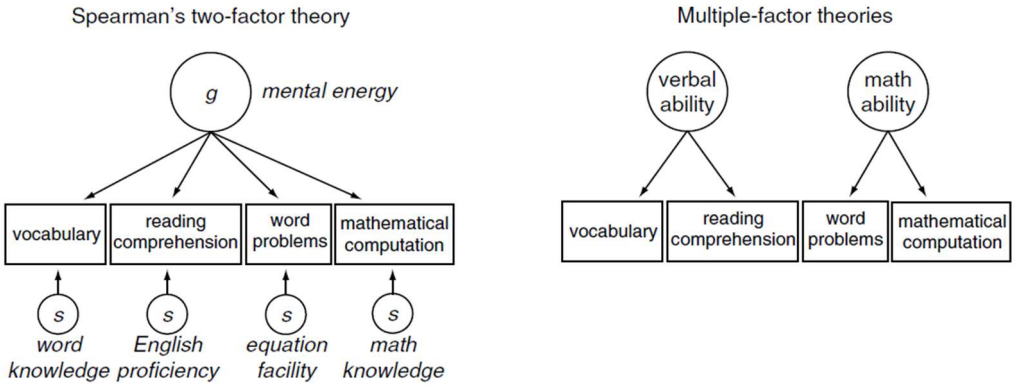


Figure 1.2: Graphical representation of two alternative theories of intelligence.
From Cianciolo & Sternberg, 2004, p. 5.

As Plucker summarizes (2016), when Thurstone analyzed test data from samples of people with similar IQ scores, he found that they had different PMA profiles, further supporting his theoretical model and suggesting that his work had more clinical utility than Spearman's unitary theory. However, when Thurstone tested an intellectually heterogeneous group of children, he did not find that the seven PMA were entirely separate; rather, he found evidence of the existence of the *g* factor as well. Thurstone managed a mathematical solution to make sense of those apparently contradictory results. The final version of his theory accounted for the presence of both a general factor and the seven specific abilities. This paved the way for the so-called hierarchical theories, where Spearman's *g* has been rehabilitated as the top of a hierarchy including group-factors. Therefore, despite Thurstone's attempt, the positive manifold is today still understood as reflecting a general cognitive ability.

However, things cannot stop at this point. The debate between Spearman and Thurstone should not be understood as a psychological controversy in a strict sense: any problem related to factor analysis concerns statistical methodologies to analyze tests data and the correlations among test scores. It is worth noting that Thurstone's theoretical intuition could have been supported by factor analysis in the same way Spearman's hypothesis was supported. In fact, depending on the nature of the variables one wants to model, some factor models can represent data better than others (see Jensen, 2002). A principal component like *g* represents information about mental tests within a unified form: a single number stands for a great number of correlational data. But, on the other side, whenever one projects several factors on a single component, one loses some information: the more is the simplification, the more is the information loss. Remarkably, this can include the theoretical meaning of correlations and variables.

These facts about psychometrics analyses attest a distinction between the validity of a factor analysis and the reliability of the related theoretical concepts: both Spearman and Thurstone used valid statistical techniques, but the two methods endorsed quite different

conceptions of intelligence.

This distinction is important insofar as it says that Thurstone conception of intelligence was not necessarily wrong: he simply found correlations among tests, getting caught in the positive manifold. As Kray & Frensch notice (2002), it is quite hard to set up cognitive tests capable of separating cognitive processes in such a way to obtain no correlations among them. For instance, tasks requiring visual processing will correlate with each other, but no one would say that general intelligence should be reduced to visual processing. Furthermore, several alternative explanations can be advanced to make sense of the positive manifold without including the psychobiological *g* (see Chapter 7).

Thurstone's attempt inspired a generation of psychologists who approached the problem of intelligence by starting from a non-psychometric viewpoint (e.g., Gardner).[17] What cognitive science tells us about intelligence? Does cognitive science account for any general cognitive ability? Strikingly, the psychometric view has not really raised these questions until the last decades. It is a fact that cognitive sciences are far more recent than psychometrics. Therefore, it is not surprising that the two, being historically and theoretically separated from each other, have pursued different questions and methods. This is likely the reason why there is no concept in cognitive sciences that could be compared to the *g* factor. Rather, cognitive sciences have mainly focused on the functional segmentation of the neurocognitive architecture rather than on any general cognitive ability (see Naglieri & Das, 2002). And this is likely the reason why, when cognitive psychologists faced the problem of testing intelligence, they often landed to quite innovative strategies aimed at disentangling different cognitive processes and skills from each other (e.g., the so-called 'second-generation tests'). The outcome has frequently been the exclusion of general intelligence from the ontological catalog.

Despite such proposals, single-factor tests are still the most important intelligence test in the field; IQ and *g* still represent the most important definitions of what intelligence is; the two still have a central role in clinical practices; they represent the theoretical constructs adopted in genetics research. As Schneider and Flanagan notice,

"The focus on the general factor of intelligence is alive and well both in clinical research and in practice. Many scholars believe that the clinical use of intelligence tests should focus primarily on the general factor of intelligence as measured by overall IQ. These scholars do not deny that other factors exist or that they are associated with important outcomes (Glutting et al., 2006). However, they argue that we measure few abilities with sufficient validity to use such measurements to make helpful decisions about individuals (Canivez, 2013). These conclusions are, of course, passionately disputed by many scholars and practitioners in the field. Nevertheless, despite the vituperations of partisans, the research is still ambiguous enough that

[17] Thurstone also influenced the psychometrician Joy Guilford, who developed a theory of intelligence which does not involve any general factor, but rather 120-180 factors (Guilford, 1967, 1982). The related abilities included different contents (figural, symbolic, semantic, behavioral), cognitive products (units, classes, relations, systems, transformations, implications), and mental operations (cognition, memory, divergent production, convergent production, evaluation). The problem with Guilford's theory was the ubiquitous correlations among factors (see Cianciolo & Sternberg, 2004, p. 6).

either position is still intellectually respectable (Schneider 2013)" (Schneider & Flanagan, 2015, p. 318).

Often, all these things are due to the fact *g* grounds in factor analysis. However, a general factor is not the inevitable outcome of any correlation matrix. What is the correct viewpoint on intelligence from a psychobiological viewpoint? This is the question on which I am mostly interested and on which I mostly focus on in Chapters 6 and 7.

Psychometricians often agree on the fact that nobody really knows why intelligence tests are related to each other. It stands to reason that such a question cannot be addressed within psychometrics research (see Jensen, 2002; Gray & Thomson, 2004). Rather, this is a matter of ontological studies: it entails both a definition of what intelligence really is and empirical research to understand its "nature", e.g., its biological bases, the underlying mechanisms involved, the multi-leveled processes from which it comes out. In fact, whether intelligence is a general phenomenon, or the compound of distinct cognitive processes, should be taken as a question about *how things work*. This is an ontological question which is likely untreatable within statistical research. Indeed, the problem concerns a choice between instrumentalist and realist conceptions of intelligence.

Like almost any scholar, I believe that cognitive sciences deserve a role in the inquiry on the psychobiological nature of intelligence. In Chapter 6, I provide a philosophical argument to shed light on how important this contribution is. However, in Chapter 7 I show that the inclusion of cognitive sicentists in the debate about intelligence has left many issues unsolved.

My position is that, by starting from cognitive and biological sciences, there is no way to support the PSY-GEN approach to intelligence. Nevertheless, much more philosophical work has to be done to firmly attest this strong conclusion. Because of this, my discussion will not begin with the analysis of the psychometric approach to intelligence, but rather by discussing the role of behavioral genetics in the research on intelligence. Being psychometricians aware that psychometric methods are not enough to attest the reliability of general intelligence, findings coming from genetics have often been welcome for the sake of bridging the gap between statistics and ontology. In the second part of this chapter, I introduce the reader to the literature about genes and intelligence.

2. Genetic Intelligence

I shall now consider the other two features of human intelligence as understood in the psychometric-genetic framework: intelligence as a quantitative phenotypic trait (§2.1) and as highly inheritable (§2.2 and §2.3). In §2.4, I focus on molecular research on intelligence. I also introduce theoretical issues related to these aspects.

2.1. A quantitative trait

Since the nineteenth century studies of Galton, intelligence has been considered a quantitative trait. This definition relies upon an important distinction between qualitative and quantitative traits that was born at the dawn of modern genetics and which is still widespread in contemporary research.

During the second half of the 19th century, Galton and Mendel developed quite different views of inheritance and biological variability. These views came to be characterized as two opposing theoretical traditions: a quantitative tradition, represented by the biometric approach, and a qualitative tradition, advanced by Mendelism.

Biometricians conceptualized phenotypic variation in a quantitative way, by focusing on phenotypic traits as related to several hereditary factors of small effect size. Quantitative traits—or complex traits, as they are often called—vary continuously over populations according to a bell curve. Height, weight, skin color, bloody pressure, IQ, and many other biological features occur in populations with a continuous range of variation. By the way, they are not something that individuals can "have or not": they are shared by all the individuals of a given species but expressed differently among them.

By contrast, Mendelians conceived phenotypic variation in a qualitative way, by focusing on discrete phenotypes. Traits like a specific pea color or a disease are either present or not in individual organisms—indeed, they are often called 'yes-no' or 'simple' traits. Furthermore, they are different from quantitative traits in their underlying factors: indeed, they are supposed to be causally related to single generative factors.

The debate among biometricians and Mendelians, which is known as the Mendel wars, last until the advent of a compromise officially signed by Ronald Fisher. In his popular paper *The Correlation between Relatives on the Supposition of Mendelian Inheritance* (1918), Fisher denied the validity of the qualitative approach, starting to treat the discrete phenotypic variation as a limiting case mainly concerning experimental conditions. He proposed a quantitative-additive view, namely the infinitesimal model (Nelson et al., 2013) or the polygenic model of inheritance (Mather, 1941, 1943). According to this framework: a) in natural populations, phenotypic traits vary continuously and are influenced by several Mendelians factors, or alleles; b) the inheritance of each allele is explained by Mendel's laws; c) the effect of each allele is small and accumulates with other genetic effects and with environmental influences as well. I will refer to this model as the *QuAd model* to emphasize the two aspects of quantification and additivity.

With the constitution of the Modern Synthesis, the quantitative view has in practice triumphed over the qualitative one. Indeed, Fisher's model has been enthusiastically welcomed by both sides for its ability to provide a unified explanatory framework based on statistical analyses. Aside from being a theoretical unification of several central questions, it was also a valuable help for practical interests in plants and animal breeding. The biometric tradition has been thus implemented in genetics, implying the adoption of useful

statistical tools. This paved the way for heritability analysis, a central methodology both in quantitative genetics and in breeding practices.

Nowadays, Fisher's model is still pivotal in genetics research. In fact, it is at the core of the study of human complex traits via heritability analysis and genome-wide association studies. Moreover, it represents the theoretical foundation of behavioral genetics' theory of intelligence. Unfortunately, it is not entirely clear how to interpret this quantitative framing of complex phenotypes. What does it mean that intelligence is a quantitative trait? How exactly could genes influence phenotypes quantitatively? It is uncontroversial that complex traits vary continuously within populations. This is an empirical fact. But does it follow that they are related to the small, equal, and additive effect of many alleles?

With his model, Fisher made a remarkable number of assumptions for the sake of simplification. A mathematical model of the genotype/phenotype relationship has been settled, in line with the guidelines required by population geneticists. Nevertheless, it should be noticed that, in its original formulation, the model was not designed to involve any strong ontological commitment. Despite this, Fisher's assumptions today represent important ontological aspects in quantitative genetics, not only related to statistical inquiries but rather understood as biological principles.

As I shall show in Chapter 2, behavioral genetics adopts the quantitative view not only to carry on statistical analyses but also to provide a quantitative explanation of the genotype/phenotype relationship (G-P map). I will formalize this attempt in relation to two distinct models traceable within genetics research, often tacitly adopted. The first one (the alleles-units model), states that each allele brings about specific "units" of phenotypes—e.g., specific IQ points (see Rietveld et al., 2014). The second one (the multilevel quantitative model) assumes additivity as a multilevel feature of biological systems: additive effects of alleles influence the phenotype passing through proteins and biochemical processes (see Plomin et al., 2009; Plomin et al., 2013).

It is important to underlie the dialogue between geneticists and psychometricians to which I previously referred. Thinking of intelligence as a general mental ability allows one to analyze it as a general biological phenomenon and to quantify over intelligent behavior. However, in Chapter 5 and 6 I highlight that this approach to human intelligence is very problematic and I point out that some quantitative traits (e.g., IQ) represent a quantification over other phenotypes (e.g., specific cognitive components).

The quantification on phenotypic traits has several purposes in genetics research. Perhaps, the most important one concerns the aim of separating the genetic and environmental influences on IQ individual differences in human populations.

2.2. Nature and Nurture

Measuring intelligence attracted the attention of behavioral geneticists from the beginning. As I mentioned in §1.1, this was not only related to noble epistemological aspirations but also to the nascent eugenics: the newborn field inaugurated by Galton required some tool to deal with intellectual differences in a quantitative manner. But let us procede step by step.

The English scientist read with a great deal of sympathy *On the Origin of Species*, written by its cousin Charles Darwin. The ongoing debate on evolutionary biology persuaded him that every characteristic of living beings should have some hereditary basis. And not only that: he also believed that individual differences should be due to heredity (see Boakes, 1984). It is worth noticing that none of the genetics principle we know today was available at that time. Nevertheless, even though neither Darwin nor Galton came in contact with Mendel's work (its rediscovery dates to the early twentieth century), a long-standing debate about inheritance and biological variability was already going on. Darwin was in fact embedded in a flourishing dialogue with other authors. At some point, the main problem of his theory became to explain the origin of variation among individuals, that is, the raw material on which natural selection acts. Darwin adopted numerous hypotheses—ranging from the pangenesis theory to the Lamarckian theory of acquired characters.

Galton has been sometimes attracted by such hypotheses (see Ferraguti & Castellacci, 2011, pp. 320-329), but he eventually developed a personal viewpoint. He adopted an empirical approach involving statistical analyses of resemblances among relatives and was convinced that the most valuable way to study heredity must entail quantitative methods. His concept of heredity is tied to the Ancestors Law: each parent contributes 1/4 of the heritage of an individual, each grandparent 1/16, and so on. This implies that infinitely distant ancestors contribute to similarities and differences among contiguous generations. Hence, heredity is susceptible to measurement because it is a matter of the physical appearance of individuals (see Griffiths & Stotz, 2013, pp. 9-14; Cowan, 1972a, pp. 407-408).

The biometric approach to biological variability has been later developed, among others, by Karl Pearson, Raphael Weldon, and George Yule. A quantitative tradition originated from Galton's studies and led to the ANalysis Of VAriance (ANOVA), and subsequently to heritability studies. The variance of a trait, introduced by Fisher (1918), measures how much the trait differs in a population. It is calculated by measuring the difference between the value of the trait in each individual and the population average value; the variance is the mean of the squares of those differences. As Griffiths and Stotz explain (2013, p. 182), Fisher used the concept of variance to quantify genetic and environmental differences and to establish how much they contribute to the phenotypic differences in a population. This focus on the relative contributions of genetics and

environment became the defining methodological mark of traditional behavioral genetics.

ANOVA partitions the total phenotypic variance for a trait (V_P) into the additive contribution of genetic variance (V_G) and environmental variance (V_E):

$$V_P = V_G + V_E$$

Heritability analyses came later. The term heritability has been introduced by Lush as a measure of how quickly a population reacts to selection (Lush 1945, 1949). The author distinguished two types of heritability: broad-sense heritability (H^2) and narrow-sense heritability (h^2), estimated by means of the following formula:

$$H^2 = V_G/V_P$$

$$h^2 = V_A/V_P$$

Broad-sense heritability includes the independent effect of each allele on a trait and the interactions between alleles, such as dominance and epistasis, which generate non-additive variance. Hence H^2 refers to the entire genotypic variance. Instead, h^2 refers to the additive genetic variance only (V_A), and it is considered more interesting to plant and animal breeders.[18]

It is not so trivial to assess what type of heritability is more interesting to behavioral geneticists: broad-sense heritability is what one would really like to calculate, but it is an elusive target in human populations; therefore, narrow-sense heritability has been in general more attractive to behavioral geneticists for methodological purposes. Narrow heritability is, however, more attractive for those who adhere to the QuAd model. In fact, by focusing on additive genetic effects, one can separate genetic and environmental effects, that is, estimating the magnitude of nature and nurture. Conversely, if one focus on broad-sense heritability, various types of interactions will make it hard to achieve such a separation of causes (see Chapter 4).

It is worthwhile to consider that in human populations one cannot apply heritability analysis as in artificial populations. Then, various kinds of alternative experimental designs have been developed to disentangle genetic and environmental influences on phenotypic variance. These methods are generally clustered under the label ‘family studies’. They include the adoption studies and the twin studies—e.g., the analysis of monozygotic twins (MZ) reared apart and the comparison of monozygotic and dizygotic twins (DZ). When two kids live in the same family with their natural parents, they share both genetic

[18] Narrow heritability serves as a breeding value coefficient to predict how much a population will change over generations because it represents the genetic source of the resemblance between parents and offspring. For instance, if h^2 is 100% for height (i.e., all variation is genetic), then the value of the trait will be midway between the parents’ trait values in the offspring (see Schaffner, 2016, p. 23).

determination and environmental experiences. Conversely, when two kids share only the environment but not their genes (i.e., when at least one of them has been adopted), one can, in theory, disentangle genetic and environmental effects on their behavioral traits. Again, when two identical twins are reared in different families, a sort of separation between genetic and environmental effects is possible insofar as the two are genetically very similar (approximately identical) but they are subject to different environmental experiences.[19]

That said, how is it possible to calculate a trait's heritability starting from phenotypic similarities and differences among relatives? Broadly speaking, behavioral geneticists contrast phenotypic similarity (e.g., IQ scores) with genetic similarity. The phenotypic similarity is estimated by means of different tools depending on the trait (e.g., IQ tests represent the way by which intelligence is evaluated). The genetic similarity, instead, is statistically inferred from basic principles of genetics and it is addressed by the coefficient of genetic relatedness.

Related Pair	**Proportion of Additive Genetic Variation Shared**	**Proportion of Dominance Genetic Variation Shared**
Parent and Offspring	1/2	0
Half Siblings	1/4	0
Full Siblings	1/2	1/4
Non-identical Twins	1/2	1/4
Identical Twins	1	1

Table 1.3: Coefficients of Genetic Relatedness. From Purcell, 2013, p. 380.

For instance, DZ twins share, on average, half of the additive genetic variance. MZ twins, instead, share all their genetic makeup. If individuals who are more closely genetically related tend to be phenotypically more similar with each other than others, then this tendency is evidence for that trait being heritable—that is, the trait is at least partially influenced by genes. Take IQ. One could calculate the covariance between the individual's IQ and the sibling's IQ. If the covariance is greater than zero, this implies that "smarter" individuals tend to have "smarter" brothers and sisters. Heritability is twice the difference between the correlations observed for MZ and DZ twin pairs:

$$H^2 = 2(r_{MZ} - r_{DZ})$$

[19] Unfortunately, the separation of genetic and environmental effects in family studies is not so easy. The literature is plenty of criticisms to some assumptions which are necessary to divide nature from nurture. See Chapter 4.

For example, if the correlation of IQ is 0.64 in MZ twins and 0.44 in DZ twins, heritability is calculated by taking twice the difference between the correlations, that is, 2 × (0.64 – 0.44). In this case, 40% of the phenotypic variance of IQ in the population is attributable to the additive effects of genes (see Purcell, 2013, for a detailed explanation). The remaining portion of the variance for the trait is attributable to environmental influences and obtained by subtraction.

But what are *nature* and *nurture*? Nature is a general term that refers to the genotype, the set of factors that are inherited from parents. Nurture is, instead, everything which does not fall within that set. While behavioral sciences define it simply as the environment (Plomin et al., 2013, p. 73), in quantitative genetics the environment includes all influences other than inheritance, including, for instance, prenatal events, nutrition, illness and social factors (Plomin et al., 2013, p. 106).[20]

Environmental influences are, by the way, distinguished in shared and non-shared. Therefore, according to behavioral genetics the sources of variance in a population are not just nature and nurture, but rather genes, common environment, and non-shared environment. The ACE model summarizes this point by partitioning the phenotypic variance in three sources of variance: additive genetic effects (A), common environment (C) and non-shared environment (E) (see Figure 1.3).

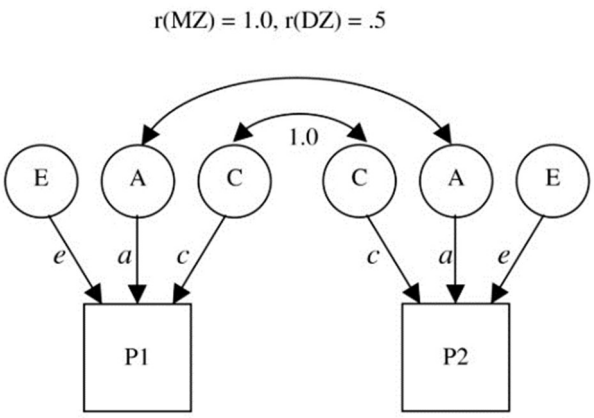


Figure 1.3: The ACE Model.

Both the notion of nature and nurture hide remarkable theoretical issues. For instance, in the ACE model, narrow heritability is germane, insofar as it represents additive genetic influence (Schaffner, 2016, pp. 26-27). Why is nature reduced to additive genetic variance? As I explain in the next chapters, non-additive genetic effects are frequently ruled

[20] To define the environment is however not a simple task, especially if one takes into account the eco-devo and the niche-construction models. For instance, the definition of environment changes depending on the level of the organism considered (e.g., cell, organism, holobiont). It also makes difference if one separates environment, experience, and learning from each other (see Michel, 2010).

out in quantitative genetics both for methodological and theoretical reasons. But if one accepts the genomic complexity of living beings, then it is straightforward that nature represents much more than additive allelic effects.

Environment presumably entails even more complexity. Unfortunately, genetics did not focus on the environment as it did on genes. In fact, the former represents, in a sense, the negative pole of the latter. Hence, we did not yet achieve a full understanding of what environmental influences really are or of how they interact with the outstanding genomic complexity uncovered by molecular and developmental research.

Plomin and colleagues (2013, pp. 105-106) stress how genetics is changing the way we think about the environment. According to the authors, three of the most surprising discoveries from behavioral genetics are about nurture rather than about nature. The first discovery they refer to is that non-shared environmental influences are surprisingly important in explaining individual differences. The second discovery concerns genotype-environment correlation (*r* G×E): many environmental measures widely used in the behavioral sciences show genetic influence, suggesting that people create their own experiences, in part for genetic reasons. The third discovery concerns genes-environment interactions (G×E): the effects of the environment can depend on genetics and that the effects of genetics can depend on the environment.[21] So, we might say whit Griffiths & Stotz (2013, pp. 183-186) that phenotypic variance must include genotype-environment interactions in such a way:

$$V_P = V_G + V_E + V_{G\times E}$$

However, as I show in Chapter 4, developmental biologists have pointed out that the genes-environment interplay represents an unmanageable issue to quantitative genetics. In fact, roughly speaking, if interactions were important, it would be hard to evaluate *quantitatively* the role played by nature and nurture in phenotypic variation, which is the chief aim of heritability research (e.g., Lewontin, 1974; Hood et al., 2010; Jacquard, 1983; Kempthorne, 1978; Michel, 2010; Wahlsten, 1990, 1994).

2.3. Genes, behaviors and the IQ controversy

Classical quantitative genetics is not interested in identifying specific genetic or environmental influences: heritability is just a general evaluation of whether the two has a role in phenotypic variation and a method to measure the relative contribution of two sources of phenotypic variance in populations. Therefore, the analysis of variance and

[21] These findings derive from environmentality analyses and from the study of genes-environment interplay which are counterparts of heritability in family studies. For more details, see Plomin et al., 2013, pp. 95-101 and 105-125.

heritability achieved Galton's dream of establishing the role of nature and nurture on intelligence.[22]

By taking biometrics and psychometrics conjointly, one has a fairly simple way to estimate *how much* genes count for intelligence. Broadly speaking, one has, first, to measure intelligence in the human population; second, one must apply correlation analyses to related individuals. In the case of heritability, the output of such steps is a percentile index of how much genes influence the intellectual individual differences.

Once applied to humans, high heritability indexes have been estimated for almost any behavioral trait. Indeed, Turkheimer (2000) has labeled this discovery as *the first law of behavioral genetics*: all human behavioral traits are heritable. According to this law, for any behavioral trait which varies in human populations and for which data are available, a portion of the phenotypic variance is statistically associated with genetic variance.

Personality traits, IQ, and psychopathological traits attracted the attention of the first generation of scholars, who estimated high heritability rates for all those behaviors (see Newman et al., 1937; Juel-Nielsen, 1965; Eysenck & Kamin, 1981; Shields, 1962). The IQ controversy is an important case of disagreement triggered by such data at the beginning of the 1980s. At the dawn of the controversy, the hereditarians (as authors like Cyril Burt, Arthur Jensen, and Hans Eysenck were called) argued that the high heritability rates of IQ stipulated the failure of the typical approach adopted within social sciences, according to which, roughly speaking, the way we are depends on *how* we live (i.e., the family in which we are born, the school we attend, the experience we are subject to, etc.). Rather, hereditarians said, the way we are depends on heredity. Several authors belonging to the environmentalist side criticized such a sentence. Among them, Leon Kamin, Richard Lewontin, Steven Rose, and Stephen Gould have been particularly critical of the theoretical assumptions made by behavioral geneticists. The dispute has never really ended up, emerging again during the 1990s after the publication of *The Bell Curve* by Herrnstein and Murray (1994) (see Block, 1995; Panofsky, 2014).

The IQ controversy led behavioral geneticists to improve their methods (i.e., trying to avoid bias in mental testing and heritability analyses) and to reassess the role of environment in human behaviors. However, nowadays, heritability estimates for IQ still hold, generally ranging between 50% and 80%, and low IQ is as heritable as IQ in the normal range (see Bouchard, 2004; Plomin et al., 2013, p. 165). Over time, heritability analyses have been applied to an increasing range of human behaviors. For instance, a popular longitudinal project called MISTRA (*Minnesota Study of Twins Reared Apart*; Bouchard et al., 1990; Bouchard & McGue, 2003), headed by Thomas Bouchard,

[22] It is important to bear in mind that Galton was interested in separating nature from nurture because of eugenics. As Morrison underlies, if one could resolve observed phenotypic variance into different fractions (i.e., expressing these fractions as functions of observed correlations) then one could easily determine the extent to which nature dominates over nurture (Morrison, 2007, p. 322).

analyzed in many different aspects several MZ twins. According to Lawrence Write's report (1997), meeting again as adults, these twins revealed very interesting resemblances: not only similar IQs or personality traits, but also shared hobbies, political preferences, religious tendencies, hobbies, habits, nervous tics, similar partners, life events (e.g., divorce), similar pets, and so on.

It is not entirely clear how to interpret such findings. Looking at social and psychological sciences, those data might strike one. In fact, we naturally tend to think that environmental influences could say anything that matters for many behaviors. As Ramus (2006) notices, a cognitive scientist's first encounter with behavioral genetics must no doubt come as a shock.

From a biological viewpoint, instead, there are two major concerns. First, one might ask how genes could matter for (at least some of) those traits. How can there be a gene which raises the risk of divorce? How can there be a gene which influences a political preference? How could genes, in general, exert any sort of influence on such complex traits? Second, one might wonder how heritability data could match with what we know from other biological sciences.

Let us take Wahlsten's reconstruction of the theory of biological intelligence (2002): genes specify the structure and physiology of the nervous system; the structure and physiology of the nervous system determines cognitive and learning abilities that constitute genuine intelligence; because a person's heredity is a constant throughout life, her intelligence relative to same-age peers is essentially fixed and psychological changes related to the environment do not reflect changes in intelligence. These theoretical principles might seem to mismatch with several facts which apply to living beings pointed out by developmental biology, systems biology, and neurobiology. For instance, I am thinking about epigenetic mechanisms and neural plasticity. Broadly speaking, biological sciences account for a non-deterministic and non-reductionist view, where genes are just one of the actors in the complex causal network which "makes" an organism.[23]

Different strategies have been adopted to deal with these issues. As far as I see, none of them is completely convincing. The first strategy takes data as they come. If one asks how to explain those data by starting from psychological sciences, one would say that genes matter for any behavior because empirical research says so. One would also say that classical theories from humanities—ranging from psychoanalysis and phenomenology to sociology—should be revised in the light of genetics discoveries. The degree of deterministic commitment varies depending on the author. For instance, Arthur Jensen (1969) has been particularly adamant in his conclusions. Conversely, Robert Plomin concedes to the environment an important role in the behavioral domain. Some psychologists appeal to these recent concessions to make sense of their clinical practice (e.g., Fonagy et al., 2005). The problem with this strategy is that, by taking data as they

[23] I further analyze this point in Chapter 4 and 5. For the present moment, we might say that behavioral genetics is faced with the so-called 'developmentalist challenge' (see Schaffner, 2016).

come, one should reject findings from several biological sciences: if complex behavioral traits are strongly determined by genes as behavioral geneticists tend to think, then something seems to be wrong in the general way by which biologists frame organisms.

A moderate strategy consists in weakening the significance of heritability findings for our scientific and social purposes. As Gottlieb stated (1992, p. 147), genes are an inextricable component of any developmental system and thus it is trivially true that genes are involved in all traits. This sounds nowadays like a truism which does not say anything interesting about biological systems. Therefore, heritability findings have been branded as practically useless for the very aim of genetics research: that is, developing intervention strategies in the social and clinical contexts, or explaining genetic causation. This strategy cannot make sense of several claims made within heritability research, which is supposed to be able to say something important about those aspects.

Other authors almost completely reject heritability data (see e.g., Joseph, 2004, 2006) by looking at the methodological weaknesses which characterize these studies. This strategy has been widely adopted by those authors who, from the early 1980s, vigorously tackled the heritability tradition. The renowned IQ controversy originated from this attitude. This strategy seems to be insufficient insofar as behavioral geneticists firmly state that those methodological weaknesses have been fixed.

In Chapter 4, I explore the debate about heritability in much more details. For the present moment, we might say that high heritability rates for IQ represent the reason why, according to behavioral geneticists, it is indubitable that inheritance plays a key role in intellectual development. Therefore, a relationship between inheritance and heritability is assumed. By contrast, several authors criticize the application of heritability on human behavior. Indeed, some fundamental assumptions in quantitative genetics remain in the same way they were a century ago; it is important to explore such original assumptions in quantitative genetics and what consequences they have on contemporary research. As I will show, numerous doubts have been raised about the domain of applicability of heritability analysis; the link between heritability and heredity; the improper inference from heritability to probability to inherit a trait; the biological meaning of the term heritability; the lack of mechanistic and developmental explanations; the social and philosophical consequences of the related empirical findings (see Barnes & Dupré, 2013; Block, 1995; Downes, 2015; Jacquard, 1983; Joseph, 2001, 2004; Kempthorne, 1978, 1997; Kitcher, 1985; Lewontin 1974, 1982; Northcott, 2006; Rose et al., 1984; Rose, 1997; Sober, 1988; Wahlsten, 1990, 1994).

In Chapter 5, I point out that, if one asks how to explain heritability data from a biological perspective, the situation seems to be quite complex. I shall submit that it is unlikely that genetics has something interesting to say about complex behaviors. Human behavioral traits are often very general and poorly defined; conversely, genetic influence is very specific, involving the encoding of RNA and proteins. It is very hard to couple these two things, not only because the relationship between genes and behaviors has not

been clarified yet, but also because the putative explanation should entail much more than DNA. In other words, the first law of behavioral genetics might conceal nothing more than *non-specific negligible correlations* between molecular and higher-level phenomena. But all in good time.

Let us now have a few steps forward by analyzing recent methodologies adopted in quantitative genetics.

2.4. Identifying the genes for intelligence

Until the 1990s, psychologists and geneticists tried to understand the role of inheritance by means of heritability analysis but, as I explained, this method does not detect specific genes related to phenotypes. The aim of identifying genes associated with phenotypic traits characterizes various types of experimental designs deriving from molecular research. In human research, where molecular intervention is not feasible, molecular genetics assumes statistical methods broadly aimed at identifying stable associations between different groups of people and genotypic characteristics.

But first, what about the relationship between genes and intelligence? Since intelligence is a quantitative trait, it has been generally assumed that several genes provide a small contribution to IQ level—namely, the polygenic model of inheritance (see Fisher, 1918; Mather, 1941, 1943, 1964; Plomin et al., 2013; Snustad & Simmons, 2012). Hence, the quest for the genetic bases of intelligence does not pertain to a specific "gene for intelligence", since the effect of every gene involved in the phenomenon, if taken in isolation, is not decisive.

Nevertheless, this has not always been the case. According to Plomin and colleagues (2013, pp. 128-129), contemporary quantitative molecular methods derive from the marriage between two different traditions: on the one hand, quantitative genetics (say, biometrics); on the other hand, molecular genetics (say, Mendelism). During the 1980s, the two have come together to identify genes for polygenic traits, technically called *Quantitative Trait Loci* (QTL). But before this reconciliation, the molecular inquiry over complex behaviors was frequently characterized by a Mendelian approach where the relationship between complex behaviors and genotype was supposed to be more linear and simple than it is thought today.[24] At that time, two approaches were widely adopted to associate single or a few genes and phenotypes: linkage analysis and the candidate-gene approach.

Linkage represents the co-segregation of a genetic marker with a disease within families. If an allele is more common in affected than in unaffected family members, then the genetic marker and the disease are said to be linked with each other and the location of a

[24] For a discussion about the relationship between Mendelian and molecular genetics, see Chapter 5.

possible defective genetic variant is established.[25] This method is ideal for genetic mapping of single-gene (rare Mendelian) disorders such as Huntington Chorea and phenylketonuria (PKU). In such diseases, the causing gene or deletion is generally present in those individuals with the disease and absent in those without the disease. Linkage analysis is considered nowadays too weak for its low statistical power and for its unsuitability for studying complex traits. Indeed, it does not work for traits where many genes are involved and where no one gene is necessary or sufficient to cause the disorder (see Jordan, 2000; Eley & Craig, 2005).

Similar considerations apply to the candidate-gene approach. This is an extremely powerful and economic method for discovering genes related to behaviors. Nevertheless, it is largely limited by its reliance on existing knowledge about the presumed biology of the phenotype under investigation. In other words, this method does not represent a tool for scanning systematically the genome (Plomin et al., 2013, pp. 137-139). Indeed, even if genes are successfully identified as related to diseases or traits (read: replicable findings), the detailed molecular features of most biological traits remain unknown (see Zhu & Zhao, 2007).

Over time, both the linkage and the candidate-gene approaches became to be considered unreliable and ill-suited for their aims, insofar as their results were small, genetically unspecific, and unreliable in replication (see Turkheimer, 2011). Association studies—among which the most important is the genome-wide association (GWA)—came to represent the most promising methodology for seeking genes related to complex phenotypes. These methodologies compare allelic frequencies for a group of individuals carrying the trait (e.g., a mental disorder) versus controls. In the case of IQ, they compare low-scoring versus high-scoring individuals. GWAS focus on single nucleotide polymorphisms (SNPs), individual segments of DNA nucleotides for which variation among individuals only includes two alleles from the available four (A, T, C, or G). SNPs are, then, indicators of the genetic variability.

Association is considered a very powerful method to investigate the genetic architecture of complex traits and diseases because many genetic variants can be assayed in thousands of individuals. However, several difficulties afflict the application of GWAS on complex human behaviors. The most important problem depends on the fact that the genetic variants associated with complex traits frequently account for only a little percentage of their overall genetic variance and of their estimated heritability. The mismatch between the heritability due to detected variants and the overall heritability is known as 'missing heritability'. Several scholars have taken this sort of (missing) finding as an evidence for the validity of the quantitative-additive genetic model: complex traits are related to thousands of alleles of small effect (smaller than previously thought!) and hence

[25] The marker may not itself be involved in the etiology of the disorder, but may be nearby and inherited along with the disease-causing locus.

none of them accounts in isolation for an appreciable proportion of the heritability estimated by classical quantitative genetics. It does follow that the right way to seek the genes for intelligence is merely improving molecular methods to eventually find reliable association between alleles and behaviors.

As I show in Chapter 3, this conclusion is not necessarily appropriate. I will argue against this idea by listing the major attempts aimed at tackling the missing heritability problem and by showing that this problem is likely due to unjustified theoretical assumptions in quantitative genetics.

Conclusion: What Is the Problem with Statistical Approaches?

The problem with psychometrics and factor analysis is that defining intelligence entails theoretical and empirical research. Finding correlations between variables is not enough to ensure that a general cognitive phenomenon is at stake behind the human intelligent behavior. As I show in Chapter 6, what we need is an analysis of intelligence where an ontological commitment is explicitly exerted. The tendency of PSY-GEN to continuously shifting from psychometrics to psychobiology reveals the lack of clarity about ontological aspects and, at the same time, leads to confusion about the very questions in intelligence research: How human brain works? What are the components of human cognition? Is intelligence a valuable posit for scientific research?

As for the case of psychometrics, part of the debate about behavioral genetics might be summarized by saying that its statistical approach is insufficient to achieve the stated objectives, that is, to make sense of the relationship between genes and behaviors. This is a biological problem; therefore, it involves ontological questions. Is statistics suitable to account for biology? I would say no.

My first worry is related to the mismatch between theories and methods. Statistical methodologies are frequently not accompanied by articulated theoretical backgrounds and are frequently said to be 'theory-free' or instrumental. For example, factor analysis can be understood as an explorative method capable of highlighting correlations among variables and indicating how to proceed with further psychological analyses. In Chapter 6 and 7, I show that it is misleading to lean on general concepts like the *g* factor as if they were real entities in the outside world. This is, by definition, beyond the territory of the factor analysis.

Genome-wide association studies represent another example of how the mismatch between theories and methods can be problematic. GWAS are understood as systematic methods to analyze the whole genome—something apparently doable without any underlying theory. This huge scanning can suggest new lines of research on specific genomic portions, but we must bear in mind that how to "zoom" into the genome is a theory-laden choice, despite behavioral geneticists tend to believe the opposite (see Chapter 3).

This separation between theory and methods, typical of the PSY-GEN model of intelligence, makes it quite hard to understand those situations in which things go wrong: spurious or non-replicable associations are renowned examples, in this sense, for the genetics agenda. Here, the lack of a comprehensive theoretical perspective can confuse researchers. This fact has often led to the publication of disarticulated attempts aimed at fixing the issues, with the inevitable outcome of more confusion. In this respect, blatant cases in quantitative genetics are represented by the contemporary interpretations of the polygenic model of inheritance (see Chapter 2), the missing heritability problem (see Chapter 3), and the heritability debate (see Chapter 4). The long-standing problem of clarifying the relationship between intelligence, IQ, the *g* factor, and neurobiological variables represents the best example for the psychological side of the coin (see Chapters 6 and 7).

My second worry about statistics concerns psychological and biological explanation. To explain how intelligence works requires, first, a good definition of intelligence; second, it requires to rappel down into neurobiological mechanisms. As I show in Chapter 7, the lack of both these two aspects is an insurmountable fragility of the psychometric theory of intelligence, despite the attempts of several scholars to come up with hierarchical models of the neurocognitive architecture.

Concerning genetics research, as I mentioned, the relationship between genes and behaviors is a biological problem which concerns the causal pathways connecting the two. Therefore, it involves the study of the individual development, instead of the analysis of individual differences in populations. By the adoption of different research strategies, I submit, behavioral genetics could really open the black box of genetic causality on human behaviors, seeking molecular and developmental explanations. This is essentially the topic of Chapter 5, where I ask what requirements behavioral genetics should fulfill to account for such an important topic. I suggest that what lacks in behavioral genetics is, ironically enough, biology. In that chapter, I then propose some guidelines to proceed in the analysis of genetic causality on complex behavioral traits like intelligence. It is in the light of those guidelines that I conclude that intelligence is not a quantitative trait and that, as far as I can see, there is no quantitative behavioural trait that is worth pursuing.

To conclude this introduction, let me say that statistical analyses are not just a methodological matter. Rather, they carry with them a luggage of ontological assumptions. As I will show, in both psychometrics and behavioral genetics, statistics has led to implausible ontological principles never really placed in the spotlight. These assumptions, on the one hand, have made psychometrics' fortune: the entire field of general intelligence studies rests on the reification (legitimate or not) of what has been found with factor analysis; but, on the other hand, such ontological assumptions irreparably worked against behavioral genetics. The two fields of the PSY-GEN model of intelligence should nowadays recover their individual identity. Psychometricians should take care of their ontological assumptions and, if required, they should rule out general intelligence from their agenda as

a genuine cognitive entity. Cognitive sciences and philosophy would play a chief role in this theoretical movement. Geneticists, on their own, should disentangle cognitive phenotypes from the psychometric concept of intelligence: the latter is a harmful posit for genetics agenda.

Chapter 2.
Quantitative Genetics: A Critical Appraisal

Geneticists often assume that some phenotypic traits are qualitative while other are quantitative; albeit apparently simple, this distinction hides remarkable epistemological issues.

Most biological features—e.g., height, weight, skin color and metabolic pathways—are frequently understood as quantitative, or biometric. IQ is considered one of the most interesting and important of such traits. Since the IQ is an index of the individual intelligence assessed by psychometric tests, trying to address issues about IQ in genetics research is very important for any theoretical analysis of human intelligence.

What does it mean that IQ is a quantitative trait? During the first half of the twentieth century, geneticists have proposed the additive polygenic model of inheritance to account for quantitative traits: several genes and environmental influences act additively on the phenotype, e.g., on the total height or on intelligence. Other traits have been described as qualitative, that is, broadly speaking related to single or a few genes, e.g., pea color and monogenic pathologies. This distinction reflects a secular disagreement among biometricians and Mendelians, arose at the dawn of Modern Synthesis.

Biometricians and Mendelians have given rise to two different types of genetic analyses: quantitative genetics and molecular genetics. It is nowadays assumed that the two have been unified into contemporary quantitative genetics, which now includes molecular research (e.g., Plomin et al., 2013, p. 128). However, a philosophical conflict underlying the original debate still influences contemporary research. The fracture between biometricians and Mendelians involves the relationship between genotype and phenotype and different conceptions of what genes are. It is worth noting that the original debate started in a very different context than the current one. Thus, the quantitative/qualitative problem concerns the way in which theoretical models have been followed one after the other, throughout a century, to account for developing conceptions and methodologies.

In this chapter, I show that the conflict has been won by the quantitative side, leading most of the scholars to embrace a quantitative conception of phenotypic traits and to avoid the qualitative reasoning originally employed by Mendelians. However, I suggest, we need to keep some qualitative aspects in the biological theory, especially for complex traits like behaviors which are often framed quantitatively.

1. A Broad Model of Qualitative and Quantitative Traits

It is often assumed that the early Mendelian theory concerns the so-called qualitative traits.[26] From a population viewpoint, Mendelism focused on phenotypic traits that vary qualitatively within populations: that is, those traits which show two or more categorical (non-overlapping) alternatives. Accordingly, qualitative traits fit into discrete categories and individuals can be categorized as belonging to distinct sets. Sometimes, qualitative traits are called "yes/no" or "either-or" traits (e.g., Plomin et al., 2013, p. 422; Purcell, 2013, p. 374).

From the inheritance viewpoint, qualitative traits are related to monogenic inheritance: the phenotype is supposed to be influenced by a single or a few genes. Kenneth Mather (1941, 1943, 1964) called this type of inheritance "oligogenic" and stated that qualitative traits are generally due to one or two genes.[27] Mendel's pea color and human eyes color are two renowned examples of this sort. Some human pathologies (e.g., PKU and Huntington Chorea) are also known to be due to monogenetic conditions: inherited via oligogenic patterns and influenced by single genes. Molecular genetics assumes that it is possible and promising to seek specific genes related to qualitative traits because their individual effect is appreciable (e.g., candidate-gene approach).

Concerning the relationship between genes and other sources of variation (i.e., environmental influences) for qualitative traits, the relevance of other factors is generally small (Plomin et al., 2013, pp. 94-95), or not highlighted because of a strong genetic involvement:

"Genes can influence phenotypes through major biochemical pathways [strong genetic explanations]. This is the case with monogenic diseases and conditions that involve a small number of genes" (Dar-Nimrod & Heine, 2011, p. 5).

"[In] the traits that Mendel studied, as well as [in] Huntington disease and PKU, […] a single gene is necessary and sufficient to cause the disorder. That is, you will have Huntington disease only if you have the *H* allele (necessary); if you have the *H* allele, you will have Huntington disease (sufficient)" (Plomin et al., 2013, p. 32).

Let us now turn to quantitative traits. The quantitative approach dates back to Galton, who developed statistical methodologies suitable to study the hereditary bases of human

[26] Although this is a standard reconstruction, it is historically inaccurate. Johannsen (1903), Weinberg (1908), Nilsson-Ehle (1909) and East (1910), but also Mendel himself, proposed a way to account for quantitative traits which looks like the one proposed by biometricians. However early Mendelians did not have the mathematical tools required for dealing with quantitative variation and hence focused on the qualitative one in experimental contexts. I thank Staffan Müller-Wille for highlighting this point (personal communication, January 2017).

[27] According to Mather (1941, p. 160), trigenic or tetragenic inheritance is relatively rare. It is not relevant for the present discussion to explore in details Mather's empirical observation. For the sake of the argument, I consider oligogenic inheritance as involving single genes.

quantitative variation. Galton's work has been early developed by, e.g., Pearson, Weldon, and Yule. The newborn biometric approach became dedicated to quantitative phenotypic variation over populations.[28] In accordance with this view, phenotypic traits vary continuously over populations accordingly to a bell curve; every alternative form shades gradually into the other and all gradations are to be observed (Mather, 1941, p. 160). Hence, these traits do not fit into discrete categories: any clear-cut distinction is arbitrary. It is worth noting that quantitative traits are not yes/no traits: rather, they are carried by every member of a species, being species-specific but interspecific also: height, weight and blood pressure, are widespread in nature, while IQ is shared by every human being—and perhaps not just by humans. Hence, individuals are different from each other only quantitatively: they show a trait in different degree (e.g., higher or lower IQ).

Quantitative traits are related to polygenic inheritance (another term introduced by Mather). According to the polygenic model, quantitative traits are related to several non-epistatic genes (up to hundreds or thousands) that influence the phenotype altogether. Their individual action is additive, small, and to some degree replaceable—insofar as it is often assumed that they bring an equal contribution to phenotype (see Dobzhansky, 1970; Mather, 1941, 1943, 1964; Morrison, 2007):

"Not all alleles operate in a complete dominant or recessive manner. Many alleles are additive in that they each contribute something to the phenotype" (Plomin et al, 2013, p. 33).

"Because of this low order of effect of individual genes, it is necessary to study the action of these genes 'en masse' by statistical techniques. This obviously results in inferences about the average properties of a set of quantitative genes" (Griffing, 1949, p. 303).

At least in the most simplified models, no one gene in a polygenic system is individually necessary. Rather, two individuals might have the same phenotype having a different *allelic makeup*. For instance, two individuals with the same height, or the same IQ level, might have different allelic combinations which lead to the same "amount". Individual differences are related to two distinct sources of variation: genes and environment. Therefore, the situation is, broadly speaking, multifactorial. I will refer to this as the quantitative-additive model of inheritance (hereafter, QuAdM).

Since the effect of every gene involved, if taken in isolation, is not decisive, quantitative genetics have not pertained to scan the genome seeking specific genes associated to phenotypic traits. Rather, this genetic approach regards statistical analysis of the phenotypic variation within populations.

Galton and Mendel worked approximately simultaneously without getting to know each other. They developed quite different conceptions of inheritance and biological variability, which have been the source of two different traditions. Indeed, at the turn of the

[28] The quantitative approach to heredity was not the only way pursued by Galton. However, biometricians ignored this aspect of Galton's work (see Cock, 1973; Morrison, 2007; Norton, 1975).

nineteenth century, a debate among Mendelism and biometrics took place. At the beginning these two perspectives were in evident antithesis (e.g., see the debate among William Bateson and Karl Pearson), but over time a pluralist consensus has been reached about the existence of both qualitative and quantitative traits. Or so it seems.

Fisher is generally considered who fixed the fracture by proposing that complex traits are influenced by several alleles, each of which is inherited according to Mendel's laws (Fisher, 1918, 1919; Fisher et al., 1932; Norton, 1975). On the one hand, Fisher attributed to the early Mendelian corpus an account of the so-called qualitative phenotypic variation, according to which traits are related to single underlying generative factors—called "genes" by Wilhelm Johannsen later. On the other hand, the biometric corpus has been related to the analysis of complex traits that vary continuously over populations. The unification of these traditions led to conceive complex traits as related to several Mendelian units.

Fisher's model is nowadays accepted as the backbone of the Modern Synthesis and as a unification of the two original traditions of genetics. I shall consider this model as just *an attempt* to reconcile the two perspectives. In fact, the disagreement between Mendelian and Galtonian approaches has silently continued, leading genetics research to bounce from one side to the other depending on the period and on the specific problem under examination.[29] This has led to several misunderstandings about the relationship between genotype and phenotype. But all in a good time.

As a first step, I shall notice that there are two different interpretations of Fisher's proposal in genetic literature:

1) Broad Model: Some traits are qualitative or oligogenic while other are quantitative or polygenic. Oligogenic inheritance concerns only simple traits, while complex traits are related to polygenic inheritance. This is what I have described above as the accepted distinction between the two types of phenotypic traits.

2) Strict Model: Every phenotypic trait is related to several genes, that is, there is no trait which is related to oligogenic inheritance. Strictly speaking, there are no qualitative traits: single alleles follow Mendelian patterns, but every trait is polygenic.

Geneticists tend to conceive some traits as qualitative and other as quantitative, adopting different methodologies (e.g., statistical or molecular) for different purposes. Therefore, they often accept the broad model. However, as I show in §3, the general trend is to go towards the strict model and ruling out any qualitative framing of phenotypes. Here, not only the existence of qualitative traits is denied—at best, they are considered

[29] Sometimes these shifts have been related to different conceptions of what genes are, depending on methodological purposes (e.g., Griffiths & Stotz, 2013; Rheinberger et al., 2015). Such pluralist views, albeit compelling, risk to cover up important ontological disagreements which deserve to be considered.

special or controversial cases (e.g., monogenic pathologies)—but the same applies to qualitative phenomena such as qualitative phenotypic variation.

The shift towards the quantitative framework is especially present in behavior genetics, where complex traits and quantitative methodologies attract more attention, but also in theoretical biology and in philosophy of science. In the philosophical context, the shift towards the strict model has been welcomed for arguing that there are no simple or monogenic traits (i.e., there is not a "gene for X", where X is a given phenotype). Rather, phenotypic traits are related to the influence of several genes.

This viewpoint is widespread and persuasive. Nevertheless, the unreliability of qualitative aspects depends on the reliability of the strict model. I will show that the QuAdM, in its contemporary interpretations, is characterized by remarkable theoretical problems. I then reevaluate the role of some sorts of qualitative framing as useful and, to some extent, necessary to account for the genotype/phenotype relationship.

1.1. Variation, inheritance and genetic influence

Before approaching the issues mentioned above, I shall notice that the vocabulary involved in the broad model might lead to confusions. Instead of saying that there are two types of traits, the qualitative/quantitative distinction can be better understood by separating questions about population variation, inheritance, development, genotype-phenotype relationships, and genetic methodologies. Since it is inaccurate to talk about qualitative and quantitative traits, I will not appeal, whenever possible, to such terms. Let me explain why by drawing some definitions:

- Population Variation: A trait might vary in a population in two different ways. According to the qualitative one, there are discrete categories that cut the population into distinct sets. Individuals do manifest or not the trait, or they manifest different forms of a phenotype without overlapping cases. According to the quantitative variation, instead, phenotypes vary continuously over populations. This distinction accounts for typical dichotomies like discrete/continuous and categorical/dimensional variation. Talking about qualitative and quantitative traits often refers to population variation.

- Inheritance: Phenotypic traits pass from one generation to another according to different patterns. Oligogenic inheritance concerns traits inherited in accordance with Mendel laws. Polygenic inheritance, instead, concerns traits inherited in a different way: Fisher's model explains that every allele is inherited following Mendel laws,

but the trait is not inherited in such a way (that is, it is not transmitted from parents to offspring in the "simple" way accounted for by Mendel's studies).[30]

- Genetic influence: It concerns the relationship between genotype and phenotype (the G-P map), the genetic architecture of phenotypes and the developmental pathways that bring about particular phenotypes starting from specific genotypes. Qualitative traits, in this respect, should be called monogenic traits, where the development is influenced by one gene. Quantitative traits are in this respect polygenic traits, that develop under the influence of many genes. This distinction accounts for widespread phrasings like monogenic/polygenic and simple/complex traits.[31]

- Qualitative methodologies (e.g., analysis of pure lines, molecular methods, linkage analysis, candidate-gene approach): These methods allow to study monogenic traits and population variation which is due to oligogenic inheritance. Artificial and controlled environments are the proper contexts to apply such analyses.

- Quantitative methodologies (e.g., analysis of variance, heritability analysis, genome-wide association studies): These methods allow to study polygenic traits, the statistical association between phenotypes and genotypes, and the continuous variation within populations. Essentially, they find statistical regularities between contiguous generations in natural populations.

Inheritance, variation, and genetic influence are disentangled from each other in many respects.[32] Table 2.1 summarizes the relationships among the three.

[30] However, as Mendelians (e.g., Johannsen) stressed from the beginning, phenotypes do not "pass" from one generation to another. What is inherited, in fact, is not a trait, but rather the DNA sequences related to the trait and, at most, the developmental potential for a trait. However, both Mendelian genetics and biometrics often speak in terms of 'inherited traits'.

[31] As I show in Chapter 5, the reference to monogenic traits can lead to misunderstandings, since they often denote monogenic pathologies.

[32] For instance, even if a trait varies qualitatively over populations, it is not necessarily influenced by a single gene. Suppose that pea color behaves categorically in population: while some individuals are yellow, others are green. This is consistent both with polygenic and oligogenic inheritance. It is also worth considering that variation is not necessarily involved in any case. Consider a trait that does not vary in the population. Every human being has two legs. The genetic variability for this trait tends to zero, and we can get this point regardless any further knowledge about its inheritance patterns or about the G-P map.

	Polygenic Inheritance	Quant. Variation	Polygenic Traits
Polygenic Inheritance		U	T
Quant. Variation	U		U
Polygenic Traits	T	U	

	Oligogenic Inheritance	Qualitative Variation	Monogenic Traits
Oligogenic Inheritance		U	T
Qualitative Variation	U		U
Monogenic Traits	T	U	

	Polygenic Inheritance	Quantitative Variation	Polygenic Traits
Oligogenic Inheritance		SC2	F
Qualitative Variation	SC1		SC1
Monogenic Traits	F	SC2	

Table 2.1: Relationship between inheritance, population variation, and genetic influence.

'T' stands for combinations that are, by definition, *true*: a polygenic trait is related to polygenic inheritance and that a monogenic trait is related to oligogenic inheritance. Conversely, 'F' stands for combinations that are, by definition, *false*.[33] The broad model accounts for 'U': it is *uncontroversial* that a polygenic trait varies continuously over populations, and it is uncontroversial that a monogenic trait shows discontinuous variation. The link with inheritance is straightforward: a polygenic trait is coupled to polygenic inheritance, hence traits related to polygenic inheritance might vary quantitatively. The same applies to monogenic traits, oligogenic inheritance, and qualitative variation.

'SC' stands for *special cases*: 'SC1' says that polygenic traits, related to polygenic patterns of inheritance, might vary qualitatively within a population. Most scholars would agree on the existence of such cases in experimental conditions and in relation to threshold phenomena (see §3.1). 'SC2' says that monogenic traits might vary quantitatively within a population. Not many scholars would agree on such cases. Indeed, when we think about quantitative variation, we deal with a gradual fade of a phenotypic form into another. It seems that the only way to account for such phenomenon is appealing to QuAdM, where hundreds of genes with small effect size influence the phenotypes. This is, in fact, the reason why biometricians adopted this model. However, the strict model does not accept SC2 by principle, insofar as it does not accept the existence of monogenic traits.

[33] If we establish that a trait is related to oligogenic inheritance, we assume that no other gene influences the development of that trait. Conversely, if we think that other genes influence its development, e.g., via epistatic interactions between genes, we should say that the trait is polygenic. Similarly, it cannot be the case that a trait is inherited via polygenic patterns and it presents monogenic developmental patterns—it would not make sense to say that it is a complex trait.

Before analyzing the contemporary accounts of the qualitative/quantitative distinction, I shall look at the historical origins of the QuAdM. Indeed, as I shall show, several contemporary issues have been inherited by the early history of quantitative genetics.

2. Rooting the Disagreement: the Mendel Wars

The qualitative/quantitative problem and the success of the strict model over the broad one root in the foundation of the Modern Synthesis. In this paragraph, I interpret the original constitution of the QuAdM as a synthesis of different lines of philosophical problems, concerning Darwinism vs. Lamarckism, gradualism vs. saltation, phenotype vs. genotype, statistical methods vs. biological theories. I focus, then, on the relevant aspects of the QuAdM useful to understand contemporary quantitative genetics.

2.1. From phenotype to genotype

Galton's theory of heredity derives from the issues related to the original Darwinian theory of natural selection. The chief problem in Darwin's theory was to explain inheritance and the origin of biological variability: Why are offspring dissimilar to parents in some respects? What is the source of the variability on which natural selection acts? At that time, inheritance was a broad term to refer to every element passing from parents to offspring. Therefore, it was debated whether variation was included in inheritance: variation was sometimes understood as related to post-natal influences (environment) and sometimes as related to inheritance itself (e.g., a mixture of parental inheritance or ancestral contributions) (Cowan, 1972a, p. 394-395). Darwin described inheritance and variation as two opposing forces, capable of being stronger or weaker in individual cases. However, introducing the indirect effect of the environment via natural selection and discussing several possible theories of the origin of variation, Darwin accentuated the confusion, making it very hard to disentangle inheritance and variation.[34]

Galton has been important because he moved from the vague term of *inheritance* to the operational concept of *heredity* (Cowan, 1972a, p. 391).[35] He explained variation and inheritance as part of the same phenomenon, that is, the Law of Ancestors: each parent contributes 1/4 of the heritage of an individual, each grandparent 1/16 and so on. This implies that infinitely distant ancestors contribute to inheritance—the contribution de-

[34] Darwin proposed the pangenesis hypothesis, according to which environment can perturb reproductive organs leading to a rearrangement of gemmules. Sometimes, Darwin has been forced to admit that Lamarck was right in attributing to the environment the capacity to directly influence evolution.
[35] For a historical reconstruction of this semantic shift, see Radick (2012).

creases in a geometrical ratio—explaining both similarities and differences over contiguous generations. In this form, heredity (the compound of inheritance and variation) became susceptible to measurement, because it was "located" in the physical appearance of individuals (Cowan, 1972a, pp. 407-408).

To fully understand Galton approach, we should remind that the distinction between phenotype and genotype has been formally proposed by Johannsen in the 1910s. Therefore, heredity was not necessarily linked to underlying factors (e.g., genes). A two-layered theory had not yet been formally proposed. Let us imagine lacking any theory to distinguish genotype from phenotype. Then, one cannot appeal to recessive and dominant alleles to explain differences among generations as Mendel did. Then, one thinks of inheritance and variation as part of the theory of phenotype, the observable characteristics. This made biometrics a theory entirely focused on phenotypes:

"Galton wanted to be able to measure something, and the only thing he could measure, the only thing that was available to him as evidence of heredity, was the physical appearance of individuals. By defining heredity in terms of the measurable characteristics of a population Galton had started down the road to genetics" (Cowan, 1972a, pp. 410-411).

In such a theory, nothing concerns underlying "invisible" explanations.[36] This conception has been reinforced by the following biometricians, who developed Galton's ideas towards a purely statistical theory without any commitment to causal notions (Morrison, 2007, p. 315). Pearson, particularly influenced by positivism and empiricism, ruled out invisible entities as far as possible from genetics explanations. His approach was, therefore, purely phenotypic (Norton, 1975, p. 540).

Mendelians brought the necessity to deal with invisible factors and to look for the underlying mechanisms of heredity. Johannsen gave the name 'genotype' to the whole set of those factors.

"From the Mendelian point of view Pearson and other biometricians were looking at the whole subject of heredity in the wrong way. […] The relationship between parent and offspring phenotypes, Johannsen argued, is not the real phenomenon of heredity. It merely provides evidence that we can use to investigate the relationship between the parent's genotype and the offspring's genotype" (Griffiths & Stotz, 2013, p. 13).

Biometricians reacted in two ways. On the one hand, biometricians tackled Mendelism. At first, because of the antipathy for invisible underlying factors. Second, for evolutionary purposes: like Darwin, biometricians were committed to gradualism. Even before Mendel rediscovery, Bateson was already arguing in favor of the idea of evolution by discontinuity, then developed by Hugo de Vries:

[36] Sometimes, Galton appealed to underlying mechanism (e.g., in the pangenesis hypothesis). However, this is the most speculative part of his corpus.

"To Pearson and Weldon, busily establishing the new order, the advent of Mendelism in 1900 came as something of a shock. First, it was pioneered by Weldon's old adversary William Bateson, who associated it with the doctrine of evolutionary discontinuity which he had opposed to their continuous view of evolution. Second, it was a system which employed theoretical terms—that is, Mendelian 'factors ', described as 'physiological units of as yet unknown nature—in a manner not obviously compatible with Pearson's methodological tenets. […] Mendelians under Bateson now opposed a programme of genetic atomism, concentrating upon explaining the inheritance of attributes […] and upon explaining visible hereditary phenomena by reference to the segregation of unseen physiological units" (Norton, 1975, p. 541).

But, on the other hand, biometricians proposed models capable of embracing both continuous variation and discrete hereditary factors. Yule (1902) argued that the two things were in theory compatible by assuming that many factors cooperate in determining a trait (Norton, 1975, p. 542). Pearson, although reluctant, showed that if a trait like stature depends linearly and additively on *n* independent Mendelian loci, each with complete dominance, the expectation of the distribution of stature for large *n* would be very close to the normal (Morrison, 2007, p. 320; Norton, 1975, p. 549; Roll-Hansen, 1978, p. 212).

It is worth reading these works as the first instances of the QuAdM the way Fisher has later formulated it. Indeed, to explain the continuous variation of a trait, it should be assumed that: 1) a very large number of underlying factors relates to the trait; 2) these factors are independent of each other; 3) any of these factors is completely dominant; 4) the relationship between a trait and the underlying factors is linear; and 5) that relationship is additive.

My hypothesis is that Fisher has succeeded in laying the groundwork to implement a theory of genotype in a theory which was essentially elaborated to study phenotypic variation. In his proposal, Mendelian heredity at every locus is accepted, and population variation concerns only the phenotype as it was for biometricians. By means of this shift, Mendel laws have been taken as a new way to interpret both variation and inheritance: no longer as part of the phenotype, but as something about underlying causal factors and, only secondarily, something about observable features.

2.2. The new synthetic quantitative model

A new theory was born for solving several theoretical necessities. On the one hand, the new theory was at the same time statistical, quantitative, and less dense of theoretical assumptions as possible (this conserved Galton's empirical approach, Yule's quantitative model, and Pearson's epistemology). On the other hand, this theory was devoted to the primary role of genotypes in determining phenotypes and, hence, evolutionary changes. Historically speaking, we might recognize Fisher's model as the cornerstone of Modern Synthesis (Stephens, 2008), insofar as it solved several issues coming from evolutionary theory: a) it accounted for evolution in natural populations using little random genetic

mutation—that is, phenotypes had come to have a secondary role;[37] b) gradualism has been supported by appealing to little mutations, ruling out discontinuous evolution;[38] c) Lamarckism has been ruled out;[39] d) it accounted for variability and evolution.[40]

At the end of its constitution, the new synthetic model superimposed a quantitative theory of phenotype (borrowed from biometrics) upon a theory of genotype (borrowed from Mendelism). The latter ended up being quantitative as the former. We should consider three important consequences. First, Galton's tendency to polarize variability on *nature* (see Chapter 1), led to think of the environment as a "container" that selects preexistent genetic variability. That is, the environment does not have any role in producing variability. Second, the assumption that phenotypes do not bring about any relevant information in the eyes of evolution has led to thinking that every information is stored in the genotype. Then, at least in an epistemic sense, the G-P map has been conceived as linear (one-to-one).[41]

The most important outcome concerns the abstractness of the model. Fisher pursued abstract thinking, criticizing Mendelians for their attention to details. He attributed three *separated* causes acting on the total phenotypic variance: dominance, environmental causes, and additive genetic effects. Fisher showed that the hypothesis of cumulative Mendelian factors fits in data and provided a plausible explanation for the inheritance of continuous variation—an aspect that both Pearson and Punnet criticized (see Morrison, 2007, p. 323). However, Fisher differed from biometricians with respect to specific assumptions about the very nature of Mendelian factors: they all were equally important.

"The simplifying *methodological* assumptions involving independence and an indefinitely large number of Mendelian units were based on the analogy with gas theory that Fisher alluded to in his earlier [1915] work. Essentially, he treated large numbers of genes in a way similar to the treatment of large numbers of molecules and atoms in statistical mechanics. By making these simplifying and idealizing assumptions Fisher was able to calculate statistical averages that applied to populations of genes in a way *analogous* to calculating the behavior of molecules that constitute a gas. But, it is important to stress that the analogy was a

[37] Weissman principle supported the idea that the only thing that matters for evolution is the genotype (evolutionary theory concerns only heritable variation). See Griesemer (2000).
[38] Ironically enough, Galton did not accept Darwin's gradualism and thought that quantitative variation was unable to explain major evolutionary changes (Morrison, 2007). Indeed, saltationists were inspired by Galton himself.
[39] This issue was important for Galton too: "Galton ignored the problem of embryological development and he eliminated all forms of the inheritance of acquired characters [because] he would have undercut the basis of his eugenic theory. […] Throughout his career the noninheritance of acquired characters—the impotence of nurture—remained [a] fundamental *a priori* assumption" (Cowan, 1972a, p. 409).
[40] Population geneticists were aware that a population must keep genetic variability below the watchful eye of natural selection—otherwise, selection would push genotypic frequencies to the optimum fast. This is one of the reasons why recessive alleles have been considered as important although they tend to lower the fitness. If Mendelians were right in assuming monogenic traits, then selection would stabilize every genotype on the 'best' allelic combination too fast, ruling out the possibility of any further phenotypic change. This was a very important reason to adopt the polygenic view.
[41] Crick's Central Dogma furnished a molecular explanation of this linearity later (see Griffiths & Stotz, 2013), but a linear conception mostly resides in Fisher's early work.

methodological one. If he could construe populations of genes on analogy with the way statistical mechanics" (Morrison, 2007, p. 323).

As a result, at the dawn of Modern Synthesis there was no point in modeling something but the additive effects of several alleles: ideally, everything linearly passes from genotype to phenotype. Admitting complex interactions between the genotype, the phenotype, and the environment would have been a complication. This is probably why the architects of the quantitative view weakened the role of interactions (e.g., dominance, epistatic effects and genes-environment interactions), albeit aware of their importance in real biological phenomena (e.g., Mather, 1941, 1943). The easiest way to model inheritance concerned allelic additive small effects in a one-to-one map.[42] The new model was now ready to be adopted by geneticists devoted to mathematical methods (i.e., Haldane, Wright, and Fisher himself). The model looks like the one originally hypothesized by Pearson. To explain the continuous variation of a trait, it assumes that: 1) a very large number of Mendelian factors relates to the trait; 2) those factors are independent of each other and do not interact in any way; 3) the relationship between a trait and the underlying factors is linear; 4) the relationship between a trait and the underlying factors is additive; and 5) the effects of the factors are equal and small.

Contemporary issues in genetic research are nested in the consequences of the original QuAdM I mentioned above and are related to the theoretical questions the model was supposed to fix. As I will show, behavior geneticists tend to think of behavioral traits as related to many alleles; they tend to admit a linear G-P map, albeit in newer forms; they tend to assume that interactions are often negligible; they tend to focus more on genes than on environment. Let us see how contemporary genetics has implemented the original quantitative model and its corollaries.

3. The Contemporary Quantitative Model

How does contemporary genetics account for the qualitative/quantitative distinction? Is there room for qualitative traits? One might notice that the reference to qualitative traits is still present, and not without any reason: in fact, diseases are often either present or not in individuals; sometimes, phenotypes show categorical forms (e.g., eye colors); some traits are related to the influence of single genes (e.g., monogenic pathologies). However, only qualitative variation is accepted in contemporary quantitative genetics: yes/no traits are mainly reduced to polygenic traits. There are three ways in which it is possible to account for qualitative traits in this strict quantitative context:

[42] Waddington also pointed at the G-P linearity as the core of Fisher's Darwinism—coupled with genetic atomism—to accommodate statistical methods (see Peterson, 2011).

1) Qualitative Strategies: Qualitative traits are exceptional cases, *real* monogenic traits. This strategy is quite common in medical genetics, but not so common in other contexts. In fact, monogenic traits are generally monogenic pathologies;

2) Mixed Strategies: Qualitative traits are polygenic traits subject to qualitative variation. This strategy is often adopted in two different manners: the first one concerns experimental conditions and artificial populations; the second one concerns the threshold model for genetic risk;

3) Quantitative Strategies: Qualitative traits are, in fact, polygenic traits that vary continuously within populations. This strategy is an extreme interpretation of the classical QuAdM, and it will likely be the favorite choice over the next years.

I the next paragraphs I analyze these three strategies. Then, I show how the current direction is to go towards an entirely quantitative account of phenotypes. I also show for what theoretical and empirical reasons the strict quantitative view has taken hold.

3.1. Qualitative strategies

Monogenic traits are influenced, by definition, by single genes. In several contexts, it is quite common to refer to Mendelian traits by assuming that traits such as eye color present categorical (non-overlapping) forms. However, some scholars think this is an old-fashioned perspective. For instance, John McDonald (2012) describes the idea that some phenotypes are due to single genes as a myth. In fact, many traits have been long considered monogenic while they are not—even eye and hair color have proved to be polygenic (see also Jamieson & Radick, 2012; Sturm & Frudakis, 2004).

However, some exceptions are admitted in medical genetics. Traits like albinism, brachydactyly, cystic fibrosis, Huntington Chorea, PKU and various mental retardations (like the ones related to PKU and Chorea), are considered as influenced by single genes which are necessary and sufficient conditions for one being affected. It is not straightforward how to read these traits in the general quantitative framework. For now, I shall consider them as exceptional cases, but I provide a possible interpretation in Chapter 5.

Yes/no traits are frequently accounted for by two different (mixed) strategies. The first one appeals to experimental conditions. Mather's classical conception of qualitative variation can be useful to approach the discussion. For him, every trait might be subject to both qualitative and quantitative variation, but it does not follow that there are phenotypes influenced by single genes:

"It is possible that, if some organism could be grown in a constant environment and rendered homozygous for all but one of the genes affecting a quantitative character, this one gene might be observed to segregate

and give sharply distinct classes just as a qualitative gene does. Nor do qualitative and quantitative genes affect different characters. Stature, for example, is usually a quantitative character, but in many organisms, dwarf forms are known to segregate sharply from the normal type, so falling into the qualitative class" (Mather, 1941, p. 160).

Dwarfism is a good example for non-normal conditions: qualitative variation is admitted as something concerning *only one of the several alleles* involved in the trait. In other words, the trait is polygenic even if it shows discontinuities in populations:

"Any given character may be subject to both polygenic and oligogenic variation. Thus a *Drosophila melanogaster* may be wild type and have some 18 or 20 chaetae on the ventral surface of each abdominal segment, but it may, on the other hand, show the effects of the mutant gene 'scute', in which case the number of chaetae is very much smaller. The flies of each kind are sharply distinct, for, though the chaeta number is variable, the two classes, wild type and scute, do not overlap. This is characteristic of oligogenic variation. But the precise number of chaetae on a wild-type fly is subject to the control of many genes each of small effect, as well as being influenced by environmental conditions" (Mather, 1943 p. 38).

To stress the point, such examples do not concern monogenic traits. Oligogenic inheritance is admitted insofar as it is linked to those alleles which account for qualitative variation within a specific population. But many traits, if not all, are influenced by several alleles. In fact, according to Mather, qualitative variation is relatively rare in nature: the interesting variation concerning natural populations is the quantitative one. By contrast, the Mendelian approach is limited to laboratory practices, i.e., artificial selection and analysis of pure lines.

The second mixed strategy concerns thresholds of accumulating genetic-risk factors. This strategy, often adopted in medical genetics, accounts for yes/no traits quantitatively by assuming that qualitative variation may not only concern monogenic traits but polygenic traits as well. Let us take psychiatric disorders like schizophrenia. Broadly speaking, they are classical yes/no traits, but are they monogenic? The answer, according to the model under examination, is negative. Plomin and colleagues (2013, p. 35) propose that the genetic risk is distributed according to a bell curve and thus related to polygenic systems, but they are not observable until a certain threshold of accumulating risk-factors is reached. Then, according to the threshold model, disorders *seem to be* either present or not, but their genetic architecture is, in fact, polygenic: like an on/off button, the light is turned on beyond a certain threshold—i.e., in the presence of enough pathological alleles (see Purcell, 2013, p. 362).

Briefly, the threshold model states that pathologies are polygenic but that they vary qualitatively within populations. The explanatory power of this strategy is controversial. For instance, it is not clear what the threshold represents: is it a genuine biological phenomenon (e.g., a systemic event in which the organism suddenly changes) or a threshold

superimposed by practical purposes (e.g., social or clinical)?[43] The question about the liability threshold model remains open, but we might say that, as a general trend, monogenic traits are related to the medical context, while "normal" traits are conceived as influenced purely quantitatively, in accordance with the QuAdM.[44]

In some extreme accounts, even the qualitative variation is going to be framed quantitatively. This does not only apply to height, weight, IQ, and skin color—renowned polygenic traits—but also to psychopathologies like schizophrenia, generally related to a qualitative framework because of the categorical aspect of the diagnostic process. Before proceeding in this direction, it is important to look at the methodological and theoretical background of this strict quantitative framework.

3.2. Why so strict?

Two distinct elements, at least, have led the contemporary debate to rule out the reference to qualitative traits and to adopt a strict quantitative model. The first one concerns empirical matters: the failure in finding a clear relationship between single genes and phenotypes (e.g., in candidate-gene approach) has been taken as an evidence that most traits are related to the small effect of several genes (for both classical quantitative traits and common disease). The second one concerns a more philosophically oriented literature. Here, several scholars criticized the idea of monogenic traits because it leads to thinking of a one-to-one relationship between genes and traits.

To explain the conception adopted by contemporary quantitative genetics, we must distinguish between two historical moments: before and after the application of GWAS. Before that, many clinical conditions were thought to be monogenic. It was typical to appeal to the candidate-gene approach, which allows to search for specific genes involved in specific phenotypes—starting from a clinical condition and looking for genes that are envisaged as functionally related to the trait. In most cases, failing to replicate results led to think that these methods, albeit powerful, are not systematic enough to find reliable associations or however flawed by wrong assumptions (see Chapter 1).

Philosophically speaking, such failures led to rethinking theoretical assumptions behind methods. Roughly speaking, it has been questioned the idea of a linear relationship between traits and genes. For instance, the reference to "the gene for X" has been criticized.[45] Several shortcomings are frequently deemed as outcomes of a strong commitment

[43] Thresholds phenomena might be related to what Haslam (2014, p. 14) calls "dimensional" and "practical" kinds. Here, the placement of a threshold is driven by external criteria, e.g., clinical purposes. These thresholds do not represent natural discontinuities.

[44] I cannot exclude that some non-pathological traits are supposed to be related to single genes. For example, human blood group is a case, albeit peculiar. I consider such traits as exceptional as pathologies.

[45] Linearity has not been criticized only because of the failure of molecular analyses, but also because of other important discoveries (see Griffiths & Stotz, 2013; Rheinberger et al., 2015).

in Mendelian analyses, for instance: 1) folk genetic determinism, related to the idea of dominance; 2) genetic essentialism; 3) simplistic one-to-one G-P map; 4) beanbag genetics; 5) the tendency to do not consider genetic interactions—e.g., non-Mendelians phenomena like epistasis and pleiotropy (see Bouran & Kampourakis, 2013; Dar-Nimrod & Heine, 2011; Fodor & Piattelli-Palmarini, 2010; Godfrey-Smith, 1999; Jamienson & Radick, 2013; Mayr, 1963; Ratner, 2004).[46]

However, the problem with Mendelian genetics likely concerns its relationship with experimental practices like hybridization. Since its born, Mendelian genetics has been interested in the *differential action* of genes, that is, the action of variant genes while other conditions are held constant. This approach characterized, for instance, both Mendel's research and the Morgan's research group in the 1910s. In the experimental context, where it is possible to analyze pure lines, one trait appears to have a simple correlation to one gene (linearity), as long as other genes and developmental elements of the network remain constant: isolating the action of single genes is here an explicit purpose.

"Morgan and his school were well aware that, as a rule, many genes were involved in the development of a particular trait as, e.g., eye-color, and that one gene could affect several characters. To accommodate this difficulty and in line with their experimental regime, they embraced a differential concept of the gene. What mattered to them was the relationship between a change in a gene and a change in a trait, rather than the nature of these entities themselves. Thus the alteration of a trait could be causally related to a change in (or a loss of) a single genetic factor, even if it was plausible in general that a trait like eye-color was, in fact, determined by a whole group of variously interacting genes" (Rheinberger et al., 2015).

Then, what is problematic is not the Mendelism itself, but its distortions: geneticists are (or should be) aware of the fact that there are no single genes that are sufficient conditions to develop a trait, i.e., there are no monogenic traits. The very aim of the critics of Mendelism is to avoid *simplicity*. Sensational statements in newspapers have long been criticized in this respect: there is not "a gene for intelligence", or for criminality, homosexuality, obesity, depression, and so forth. Phenotypes, and *a fortiori* complex behaviors, are not under the causal control of single genes. Rather, they are subject to the influence of several factors (including several genes). Then, there are no monogenic traits.

Recent empirical findings reinforced this view and seem to match with some assumptions of the QuAdM. By improving molecular research techniques, association studies have been proposed to search for genes in broader samples and to scan huge portions of the genome in reasonable times. The underlying assumption of molecular research was: if heritability is high for a trait, and our tools are adequate (powerful and systematics), it

[46] Jamieson and Radick (2013) suggest reforming genetic pedagogy in such a way that Mendel laws are not assumed as the core of genetics. This would make it be possible to focus on the complexity of developmental systems, dethroning Mendelian mechanisms as the standard rule, and electing the "exceptions" as protagonists (e.g., incomplete dominance, codominance, penetrance, pleiotropy, epistasis, phenotypic plasticity, epigenetic factors). Bouran and Kampourakis (2013) propose, instead, to change the wide concept of genes with the term 'genetic material'.

should be possible to identify genes capable of accounting for the heritability of that trait—it does not matter how many genes are involved in complex traits, molecular research will find them. Once applied, GWAS returns an unexpected dataset: no single gene accounts for an appreciable portion of the heritability, and the sum of the known associations between SNPs and traits explains only a small portion of the heritability (this latter is known as 'the missing heritability problem').

Geneticists interpreted the situation as a confirmation of the QuAdM and inferred that there are no individually relevant genes for quantitative traits:

"Although Fisher's 1918 paper provided the basis for reconciling the differences between Mendelians and biometricians, these two worlds of genetics drifted apart because of the differing perspectives that follow from thinking qualitatively versus thinking quantitatively. The two worlds are now being brought together by genome-wide association research (GWA research), which shows that the ubiquitous heritability of common disorders is due to multiple genes of small effect size" (Plomin et al., 2009).

"Despite an adequate sample size for detecting large effects and despite high-precision measurements, we found few associations between SNPs and traits at an appropriately stringent significance threshold. Since many of our measured phenotypes (including our behavioral phenotypes) are known to be heritable, the absence of strong associations in our data indicates that […] both physical and behavioral traits are mainly affected by numerous genes with small effects" (Chabris et al., 2013, p. 7).

Even though several alternative explanations of the missing heritability have been provided (see Chapter 3), the supposed polygenic architecture of human behaviors has been formalized as *the fourth law of behavior genetics*: human behaviors are associated with many genetic variants, each of which accounts for a very small percentage of the behavioral variability (Chabris et al., 2015, p. 305).[47]

3.3. The quantitative strategy and its problems

As I shown in §3.1, qualitative and mixed strategies have been kept in the limited domain of pathologies. However, for the reasons mentioned above, both genetics methods and theoretical thought converged in moving towards a strict quantitative model. This movement has been recently even emphasized: yes/no traits and qualitative variation are framed in a purely quantitative manner too. To make a case, Plomin, Haworth, and Davis (2009) examined the disconnection between qualitatively diagnosed common disorders and their quantitatively distributed polygenic liabilities, concluding that the qualitative variation is an extreme case of the quantitative one:

[47] The other three laws state that: 1) all human behavioral traits are heritable, i.e., they are affected to some degree by genetic variation; 2) the effect of being raised in the same family is smaller than the effect of genes; 3) a substantial portion of the variation in complex human behavioral traits is not accounted for by the effects of genes or families (see Turkheimer, 2001).

"Most GWA studies are case-control studies that focus on qualitative traits and typically compare allele frequencies for diagnosed cases versus controls. If GWA studies indicate that multiple genes affect these disorders, this implies that their genetic liability is distributed quantitatively rather than qualitatively" (Plomin et al. 2009, p. 872).

The authors state that disorders, which are generally treated as yes/no traits, are actually continuous: symptoms might increase gradually from normality to abnormality. A diagnosis occurs only when a certain level of symptomatic severity has reached. This implies that disorders are in fact related to several genes which act additively on the phenotype.

"The polygenic liabilities that emerge from GWA research will lead to common disorders being thought of as the extremes of quantitative traits and, ultimately, to a scientific focus on quantitative traits rather than disorders […]. Thinking quantitatively will be aided by speaking quantitatively—a shift in vocabulary is required so that we start talking about 'dimensions' rather than 'disorders' and about genetic 'variability' rather than genetic 'risk'" (Plomin et al., 2009, p. 872-873).

The authors seem to imply that the view accepted in medical genetics, i.e., the existence of monogenic (yes/no) disease, is misleading. Hence, seeking specific genes with appreciable effects on the phenotype is pointless—at least for common disorders.[48] This proposal is an extreme interpretation of the classical QuAdM. For instance, it is different from Mather's account of qualitative variation. Mather thought that a polygenic trait might, in fact, vary qualitatively when the differences within the population are entirely due to a single allele, being the other loci invariant among individuals. Conversely, the quantitative view is now going to be applied on traits which "seem" to vary qualitatively over populations.

I cannot disagree with those who criticized the simplicity of some interpretation of Mendelism (that is, the existence of monogenic traits). However, I do not agree on the choice of the putative alternative view, that is, the QuAdM. This model has been taken for tackling the problems of Mendelism because of its reference to a multitude of factors underlying complex traits. However, I believe that the QuAdM carries philosophical implications that are not alike to those generally attributed to Mendelism: in the nascent quantitative conception, we might find a linear G-P map and a consensus on the fact that every kind of interaction is negligible. Let us consider an example about IQ:

"A study with over 100,000 participants, allowed researchers to identify three promising genetic variants. Nonetheless, those three variants accounted for only a small fraction of the variation in intelligence; an

[48] As far as I know, this approach has not been applied on every pathological condition (e.g., rare diseases like Huntington Chorea). However, it might be adopted for them in the light of controversies. For instance, debates about Chorea's onset are still present, involving the number of repeated CAG triplets. PKU, in turn, is related to more than 500 different mutations of *PAH* gene, some of which cause milder symptoms (Plomin et al., 2013, p. 166). For an analysis of common *vs.* rare disorders, see Chapter 3.

individual who received both copies of all three variants would on average score less than two IQ points higher than someone who inherited none of them" (Tabery, 2015, p. S12).

"We identified three genetic variants associated with cognitive performance. As expected from the calculation, the effects of these variants on cognitive performance are tiny. A copy of each variant accounts for only 0.3 points on a standard IQ test (with a mean of 100 and standard deviation of 15). A person who inherits all six copies (note: one genetic variant has two copies) of increasing variants differs by 1.8 points compared to individual who inherits none" (Benyamin & Visscher, 2014; for the original research see Rietveld et al., 2014).

It is not clear how to read these results in a biological sense. One way is to read them according to a one-to-one relationship between alleles and units of quantitative phenotypes. In the case examined, alleles bring about specific points—or fractions of points—to the IQ phenotype. I call this *the alleles-units model*. Another way, explicitly proposed, is to look for quantitative phenomena over every level of the organism—I call this *the multilevel quantitative analysis*:

"These quantitative traits need not be limited to symptoms of the diagnosed disorder [e.g., schizophrenia] but can occur at any level of analysis. […] Once multiple genes are found to be associated with a disorder, understanding the mechanisms by which each gene affects the disorder leads to quantitative traits being recognized at all levels of analysis: from gene expression profiles, to other '-omic' levels of analysis, to physiology and often to the structure and function of the brain" (Plomin et al., 2009, p. 874).

In this case, it is hypothesized a linear relationship between genome (DNA), transcriptome (RNA, amino acids), proteome…up to phenotypes (brain and behavior) (Plomin et al., 2013).[49] One might notice that this idea looks like the one attributed to Mendelism and to the early molecular account of genes (see Griffiths & Stotz, 2013).

For now, we might say that GWAS seem to confirm the principles included in the QuAdM, now implemented in the strict quantitative model:

1) A very large number of Mendelian factors relates to complex trait;
2) These factors are independent of each other and do not interact in any way;
3) The relationship between a trait and the underlying factors is linear;
4) The relationship between a trait and the underlying factors is additive;
5) The effects of the factor are equal and small.

In the next paragraphs, I try to make sense of the alleles-units model and of the multilevel quantitative analysis from a biological viewpoint. I then point out their problematic aspects.

[49] A multilevel-analysis does not necessarily entail linearity. However, linearity must be hypothesized to account for the QuAdM, which rules out non-linear relationships.

4. Biological Interpretations of the Quantitative-Additive Model

Attributing a linear G-P map to quantitative genetics might strike one insofar as this map is generally attributed to Mendelism, where G-P relationships are relatively simple. This is, in fact, why several authors have criticized Mendelism and adopted a multifactorial view. However, appealing to the QuAdM does not save us from those problems usually blamed to Mendelism, i.e., linearity and overlooking interactions. As I shall show, these two aspects constitute important aspects of the contemporary, strict quantitative model.

4.1. Alleles and units: nothing more than mathematics

The alleles-units model is traceable in the classic quantitative view. Advocates of QuAdM attributed fixed hypothetical coefficients to each allele:

"Consider […] two heterozygotes with the same genotypic value *a.* If this is identical with its genetic value, that is, if the heterozygote is exactly intermediate between the corresponding homozygotes, the mean value of its immediate progeny and of subsequent generations will also be *a.* If, on the other hand, the genotypic value of the heterozygote is due, for instance, to complete dominance for high values, the two homozygotes can be scored as +*a* and -*a,* and the genetic value of the heterozygote is zero. The mean value of its progeny will now be 1/2 *a,* and this will be further halved in each subsequent generation, thus gradually approaching zero" (Panse, 1940, p. 104).

"With only three polygenes of equal effect, the genotypes AABBcc, AAbbCC and aaBBCC will, for example, give the same phenotype. This phenotype would also characterize the genotypes AaBBcc, AABbcc, AaBbcc, etc., if dominance were the rule, or AABbCc, AaBBCc, and AaBbCC in the absence of dominance. […] The allelomorphs designated by small letters are assumed to add nothing to the expression of the character, while each allelomorph designated by a capital letter adds 1 unit. […] As the number of genes involved increases, more phenotypes are possible, and the distribution becomes more nearly continuous […] as observed, for example, in human stature" (Mather, 1943, pp. 39-40).

"We need some way of specifying how much an allele affects the trait. Considering only a locus with two alleles, A_1 and A_2, we define the average value of one of the homozygotes (say, A_1A_1) as *a* and the average value of the other homozygote (A_2A_2) as *–a*. The value of the heterozygote (A_1A_2) is labelled *d* and is dependent on the mode of gene action. If there is no dominance, *d* will be zero (i.e., the midpoint of the two homozygotes' scores). If the A_1 allele is dominant to A_2, then *d* will be greater than zero. If dominance is complete (i.e., if the observed value A_1A_2 equals that of A_1A_1), then $d = +a$" (Purcell, 2013, p. 374).

All these authors refer to fixed coefficient representing the average effect of each allele upon a trait. This conception is closely reminiscent of the alleles-units model. However, the latter seems somehow stronger than Fisher's model. Indeed, the abstractness of the original QuAdM depends on the context in which it arose: at the beginning, population

genetics was characterized by a strong mathematical background where an abstract reference to allelic additive effects was possible and, as I shown in §2, necessary for the sake of simplification.

Interestingly, the architects of the Modern Synthesis aware that a purely additive relationship between alleles represents a simplification. The shift towards a developmental application of that model is more recent and leads to thinking of the quantitative model as suitable to account for the *real*, rather than abstract, G-P map—even though it was not proposed for this purpose. In such a way, an ontological commitment has been imported into the model: alleles bring about phenotypes according to fixed coefficients (equal additive effects). The alleles-units model likely relies upon such a shift.

Alongside this ontological commitment, the necessity to emphasize the pure additivity as an abstraction has gradually faded. For instance, Bouchard's review (2004) attests that there is no evidence of non-additive genetic effects for IQ and psychiatric diseases. Plomin et al. (2013, p. 199) say that the absence of important non-additive variance is very fortunate for the attempts to identify intelligence genes because this allows to study intelligence with statistical methods within a purely quantitative-additive framework.[50]

Ironically enough, the philosophical inadequacy of Mendelism (i.e., linearity and simplicity) is going to reappear in the quantitative view: appealing to the QuAdM does not safe us from those shortcomings generally attributed to Mendelism. Moreover, the strict quantitative model brings additivity as a further assumption which is absent in Mendelism.

Nowadays, we can go beyond the problems accounted for by the original Fisher's model. We are not in need of a quantitative model to account for the complexity of biological systems, because we now understand the chief role of phenotypes in evolution and why genotypes are not enough to explain biological phenomena. Therefore, we should go beyond abstract conceptualizations and analyze whether the contemporary quantitative models are plausible *biological explanations* (for similar concerns, see Nelson et al., 2013) of the G-P relationship. We should explain what is the biological basis of a polygenic trait, its genetic architecture and, as Plomin and colleagues would like to explain (2009), what are the mechanisms for which additivity of polygenic systems could possibly be expressed over every level of the organism, up to the phenotype.

[50] For the present moment, it is not my aim to evaluate these findings from an empirical perspective. In Chapter 3, I explain that they depend on choices behind quantitative genetics methods, which could make it unable to detect non-additive genetic effects.

4.2. The multilevel quantitative analysis and its biochemical explanation

For my purposes, I take the biochemical interpretation (hereafter, BioChem), proposed, for instance, by Ferraguti & Castellacci (2011, chapter 4), as a prototypic explanation of the multilevel quantitative analysis that involves two levels: the genome and the proteome. Let us consider a quantitative trait which is under the influence of four genes: **A, B, C, D** (a very simplified model, but still good for our purpose). These genes code for four different enzymes. Every gene is represented in the population by three alleles at appreciable frequencies. Consider the gene **A**. The first allele, A, is dominant. The other two, a' and a", are recessive, but a' is dominant on a". Thus, the dominance relations can be summarized as A > a' > a".

Gametes	**A**	**a'**	**a"**
A	AA	Aa'	Aa"
a'	Aa'	a'a'	a'a"
a"	Aa"	a'a"	a"a"

Table 2.2: Combinations ofgametes within BioChem.

Every individual who carries the dominant allele A has a genotype coding for a fully functional enzyme, say the wild-type, the form of which is called *A*. Genotypes a'a' and a'a" code for a functional enzyme, but not efficient as the first, called α'. This allele is characterized by a missense mutation. Homozygous a"a" code for a non-functional enzyme, called $\alpha"$. This allele is characterized by a nonsense mutation.

In other words, every enzyme has the same biological function, but a different efficiency or capacity to adhere to the target-substratum. This biochemical property of theenzymes can be expressed algebraically by the Michaelis-Menten equation (see Johnson & Goody, 2011). Accordingly, the relationship between substrate concentration and reaction rate can be expressed quantitatively (Nelson & Cox, 2013, p. 202). By using a mathematical simplification, we might say that genotypes influence phenotypes for a value between 0 and 1. The relative influence is the following:

1) Homozygotes AA and heterozygotes Aa' and Aa" influence amounts to 1, by means of the enzyme *A*;
2) Homozygotes a'a' and heterozygotes a'a" influence amounts to 0,5, by means of the enzyme α';
3) Homozygotes a"a" influence amounts to 0, by means of the enzyme $\alpha"$.

What I said for the gene **A** applies to genes **B**, **C**, and **D**, too. If we assume that every enzyme produces an additive phenotypic effect, then four genes and three enzymatic

forms produce 64 combinations. We can easily imagine a quantitative variation within a population for the trait, even more so if the genes involved are more than four (hundreds or thousands) and if we take, in addition, environmental influences.

In sum, BioChem grounds both quantitative variation and polygenic influence upon enzymes' biochemical properties. Individual variation within populations is interpreted in the light of this phenomenon concerning all the possible allelic combinations, in accordance with the QuAdM. Thus, the genotype acts additively on the phenotype by means of the quantifiable action of enzymes. In other words, if there something which is "quantitative", it is not the action of genes, but rather the *efficiency* of the enzymes they code for.

Since quantitative traits are meant to be grounded on biochemical properties of the products of several genes, what is, then, a trait that varies qualitatively over populations? Ferraguti & Castellacci (2011) seem to agree with Plomin, Haworth, and Davies (2009), saying that the qualitative variation is a limiting case of the quantitative one. To make an example, let us analyze pea color. Again, A, B, C, and D, are four involved genes. Suppose now that they all bring pigments to the phenotype additively: the more efficient are the produced enzymes, the more the color will be intense. We might think to phenotypes as such: on the one hand, the lightest pea conceivable has a total absence of pigment, having only non-functional enzymes (e.g., only recessive alleles); on the other hand, the darkest pea has the best set of enzymes imaginable for functionality (e.g., only dominant alleles). In some circumstances (e.g., experimental conditions), there are no individuals that bring mixed allelic combinations, and then one can observe qualitative variation. In other cases, peas show a mixed combination of alleles, and the population follows a bell curve for the trait. This interpretation of the QuAdM is supported by Jamieson and Radick:

> "At the molecular level, it turns out, there are not two things, a gene 'for' roundness and a gene 'for' wrinkledness. There is, as far as DNA is concerned, mainly just one thing: DNA encoding an enzyme that coverts sugar into starch. Depending on the number of functional copies of that sort of DNA in a given pea plant, the seeds on that plant will have different quantities of the enzyme, hence different quantities of starch, hence—for reasons to do with the effects on water absorption—different seed shapes. [Real pea seeds] show every gradation, from extreme wrinkledness to full roundness. That expectation would no doubt only get stronger with supplementary attention to how other genes in the pea genome, ambient temperature and pressure, mineral content in the absorbed water, and so forth also affect seed shape. And indeed, there are many degrees of wrinkledness (and of other traits) in real pea seeds" (Jamieson & Radick, 2013, p. 583-584).[51]

The biochemical interpretation of the quantitative multilevel analysis sounds quite persuasive, but it leaves open several questions. I now address three problems for BioChem concerning the additivity of small genetic-enzymatic effects, and I argue that the

[51] As the authors remind us, Weldon (1902a, 1902b) tackled Mendelism with a similar quantitative analysis of pea characters.

multilevel analysis is not a compelling from a biological point of view. Indeed, it rules out qualitative aspects which seems to be necessary.

The first problem with the BioChem regards the fact that we cannot quantify over qualities. According to the modl, we should assume the possibility to quantify over enzyme's functionality simply because we have a quantitative index (the Michaelis-Menten constant). This quantification is reasonable in some respect, but not others:

1) One can compare the efficiency of enzyme produced by different alleles of the same locus (e.g., the efficiency of the enzyme A with the efficiency of the enzyme a');
2) Ideally, one can quantify over the phenotypic effect of two enzymes produced by different alleles of the same locus;
3) One can compare the efficiency of enzymes produced by different loci (e.g., the efficiency enzyme A with the efficiency of the enzyme B);
4) One *cannot* quantify the phenotypic effect of two enzymes produced by different loci.

The last point is the problematic one: the Michaelis-Menten constant is not a measure of an enzyme's contribution to a specific phenotype. Rather, it is simply a measure of its ability to adhere to a target-substratum. What is interesting in proteins action is their "quality", *what they do* within an organism in a specific context. Thus, biochemically speaking, we cannot quantify anything but the enzyme's efficiency, because we cannot quantify over two different qualities.

The second issue of BioChem is that enzymatic actions are likely individually important even if they are small. The aim of quantifying the phenotypic overall influence of several enzymes' effects is controversial. This will be clear shortly. Let us remind the assumption of QuAdM, according to which an allele brings about *a quantity* to the phenotype (we can assume +1, +0,5 or 0). If one takes in conjunction BioChem and QuAdM, one must accept two consequences: a) the individual enzyme's function is replaceable by other enzymes capable to fulfil a similar function; and b) the enzymes' individual functions are irrelevant to the phenotype: their actions are phenotypically meaningful only if taken in conjunction, or *en masse*. The first consequence is uncontroversial: it is generally assumed a many-to-many relationship between enzymes and functions (i.e., multiple realizability). The second consequence, instead, is problematic: if genes have a small and additive effect, and if one maintains linearity, then, enzymes act additively on phenotypes as genes do. This is a very problematic conclusion: a phenotype is not the sum of some enzymatic actions; and even if a phenotypic outcome is the addition of several enzymatic effects, every effect is, in fact, individually important. Let us look again at pea color. We have previously taken A, B, C, and D as four involved loci. Then, we assumed that they all bring pigments to the phenotype additively: the more efficient are the enzymes, the more the color is intense. If every enzyme brings about a pigment to the phenotype, then

the phenotype is the sum of those pigments within a metabolic cascade. It is straightforward that every pigment is individually necessary to reach the final color, e.g., the dark green.[52]

The last argument against BioChem questions whether alleles have really equal effects. I believe that assigning qualitative effects to alleles is the only way to make sense of Mather's conception of qualitative variation. According to Mather, as I explained, traits are polygenic, but in some cases, they may vary qualitatively within a population (see §3.1). Qualitative variation can occur because a single allele is responsible for the categorical variation within a population. Often, this is related to missense or nonsense mutations, for which an allele codes for a dysfunctional or for a non-functional genetic product, and this have an appreciable effect on the phenotype. So, the only way to explain phenomena like dwarfism in Mather's vocabulary is to accept that some alleles have a qualitative effect—or, to say it differently, a large effect size. However, the strict quantitative model rules out the reference to large effect sizes.

If the proposed arguments hold, I submit we need qualitative reasoning to understand the G-P map. Someone might say that such conceptual arguments are not enough to dismiss a theoretical model which grounds on empirical evidence: we are in need empirical data in the opposing direction. I do not deny *a priori* that a trait related to small additive genetic effects might exist. But still, I believe that such a view fails in accounting for important biological aspects. As I show in the next paragraph, this is true from an empirical perspective, too.

4.3. From development to statistics and return

Conceptual problems arise when we apply the QuAdM on individual cases and to the multilevel analysis. One might expect similar results when the model is applied to developmental reasoning. As many scholars do, I assume that developmental biology has reached a deeper understanding over organisms' biology than the QuAdM. First, the latter represents a statistical account which achieves biologically plausibility by assuming molecular biology as its corroborative context; nonetheless, developmental biology can successfully incorporate molecular research programs as well. Second, thinking of genes and development as separate from each other seems to be a so old-fashioned posit to make my assumption very close to be trivial (see Chapter 4).

If I am right, the best context to assess the biological validity of the QuAdM is developmental biology. Here two issues for the QuAdM arise, concerning additivity and linearity as biological principles:

[52] This point is consistent with empirical findings provided by Chabris et al. (2013): at least in the case of eye and skin pigmentation, they found loci of large effect (see Chapter 3).

Issue #1: Interactions *versus* Additivity. The QuAdM rules out genetic interactions and the genes-environment interplay. The existence of these phenomena is acknowledged also in behavioral genetics, but they are generally not implemented in quantitative explanations for the sake of modeling (e.g., Purcell, 2013, pp. 400-401). However, assuming additivity as the core architecture of phenotypic traits does neglect the complexity arising from those interactions and, in turn, leads to omitting very important aspects about biological organisms. Indeed, developmental biology points out that the additive model fails in accounting for the G-P map. Both Waddington's epigenetics, the analysis of reaction norms, and the critics of quantitative genetics as well, attest loudly this point.

Issue #2: Non-linearity *versus* Linearity. The QuAdM implies G-P linearity, which is, as far as I know, not justified in any biological sense. This is endorsed by both molecular genetics and developmental biology. The former brought the transition from an abstract concept of genes (where linearity is acceptable) to a material one, where linearity has been ruled out: molecular biology itself has made it impossible to think of the genome simply as a set of pieces of contiguous DNA co-linear with the proteins derived from it (see Rheinberger et al., 2015). If this is true for proteins, it must be true for higher levels of a biological system as well. With respect to development, by assuming G-P linearity there is no way to explain phenotypic plasticity and robustness. Non-linearity between genotype and phenotype is what these phenomena require.

Briefly, interactions, robustness, and plasticity are not easily implemented in a strict quantitative-additive model.[53] Nevertheless, we should consider that the QuAdM has its domain of applicability, and I do not mean to throw out the champagne with the cork. To understand quantitative genetics, we should return to the very meaning of the original quantitative model: its authentic goal was to account for statistical inquiries. This statistical-oriented approach derives from the biometric, correlational analysis of the phenotypic similarities among relatives. Apparently, in this framework, there is no ontological commitment to the existence of quantitative biological phenomena in a strict sense, and then the QuAdM does not quantify over genetic effects because it is not committed to a quantitative explanation of development. In other words, the mathematical tradition where the QuAdM has been proposed was interested in genotype-to-genotype map among generations. Instead, development, interactions, robustness, and plasticity, have a part to play in the G-P map problem. This is the reason why the architects of the Modern Synthesis explicitly avoided development, focusing instead on genes as developmental invariants (see Griesemer, 2000; Lewontin, 1974; Rheinberger et al., 2015).

[53] I thoroughly explain the role of genes-environment interactions in Chapter 3 and 4. In Chapter 5 I say something about plasticity and robustness, but these phenomena cannot be fully addressed in this work (for details, see Griffiths & Stotz, 2013; Lewontin, 2001; Pigliucci, 2001).

However, behavioral geneticists have, at some point, adopted the statistical view to account for the G-P map and for development. The alleles-units model and the quantitative multilevel analysis are evident examples of this shift from statistical to developmental analyses and from populations to individuals.

"[A concern follows] not from the statistical assumptions of [Fisher's] model, which has been extraordinarily successful as a statistical method for the modeling of complex genetic effects in families, but rather from the theoretical implications of [the] model for the scientific investigation of genetic and environmental developmental processes that actually produce the complex phenomena of interest. Ironically, it is precisely the aspect of the [Fisher's] model that makes it so successful on a statistical level—the fractioning of genetic variance components into a collection of indistinguishable binomial effects—that makes it problematic as a model for the actual genetic etiology of traits. It is difficult to think of a biomedical disorder or behavioral trait with an etiology that is meaningfully described as an additive accumulation of a large number of very small independent causes, however useful such a model may be for modeling the occurrence of disorders and traits in pedigrees" (Turkheimer, 2011, p. 229).

"Monogenic theories suggest major biochemical pathways which can be uncovered, whereas polygenic models suggest a complexity of chemical interactions probably intractable to exact study. Thus if most behavior traits must be fit to polygenic models, we may be left only with statistical analyses of such problems as how many genes are involved and the specification of the almost infinite number of interactions between them. Such mathematical exercises seem to us to have only trivial importance and, furthermore, to be of small interest to most biologists and psychologists" (Fuller & Thompson, 1978, p. 438).

This topic will be central in the three next chapters, where I show that conflating statistical analysis and developmental biology has been a harmful strategy for understanding the relationship between genes and behaviors.

Conclusion: Why the Missing Heritability?

I agree with those who think that makes sense only within a form of outdated biology. However, the real problem might not be the Mendelism itself, but the simplifying assumptions in contemporary quantitative genetics. At the end, the former is not more simplistic than the latter. By the way, Mendelism recognizes qualitative variation within populations, while the contemporary quantitative view is going to apply to population variation a limitless biometric approach—which was, from the beginning, a statistical framework to study phenotypes without any commitment to developmental mechanisms. Developmental biology seems to be the only way to account for qualitative biological phenomena, arising by means of complex interactions between genes, their products, and the environment.

I argued that we need qualitative reasoning to understand polygenic traits' development. This calls seriously in question a purely quantitative-additive model. However, it is not my aim to deal with any existent phenotypic trait—the biological world is just too complex to think that a few principles could be able to address every aspect. Therefore,

in Chapter 5,by focusing on behavioral traits, I will suggest that both qualitative traits (e.g., monogenic pathologies) and quantitative traits (e.g., IQ) should not be understood as phenotypic traits in a strict sense.

Since I rejected the quantitative-additive model, one may ask why the missing heritability problem afflicts genetic research. This is a good point: indeed, if I am right, genome-wide association studies should have found genetic factors of large effect-size accounting for the high heritability of the investigated traits. As I show in the next chapter, the missing heritability problem might depend on some methodological and theoretical assumptions that, once ruled out, will assert the poor reliability of the QuAdM. Indeed, some of the solutions proposed by scholars to fix the missing heritability problem seem to fit in the analysis I proposed so far.

Chapter 3.
Making Sense of Missing Heritability

During the last two decades, genetics research has been affected by a major puzzle known as the missing heritability problem. What is missing is a large proportion of the estimated heritability of complex traits, which remains uncovered by the genetic variants associated with phenotypic traits investigated by molecular research.

For my general purposes, this is a decisive topic because it tackles my criticism against the plausibility of the biological interpretations of the quantitative-additive model. In the previous chapter, I submitted that this model is appreciable insofar as it conceives complex traits as polygenic, but it is not convincing in its attempt to explain complex traits by relating them to the equal, small, and additive effects of many alleles. This assumption, I argued, is nothing but an abstraction or a simplification. I have shown that a plausible biological explanation should admit individually relevant allelic effects—e.g., for polygenic traits like eye color—and that additivity and linearity represent unjustified ontological assumptions.

One might say that, if I am right, genetic effects of large size must exist and, hence, molecular research should have found some genetic factors of large effect accounting for the high heritability of complex traits. Then, why the missing heritability problem? It is the aim of this chapter to offer a framework in which my previous analysis could fit in with empirical data provided by molecular studies.

This chapter could seem to be a small detour from my main inquiry about human intelligence. However, since IQ and intelligence are taken as coincident by the advocates of the PSY-GEN approach, and since both have been framed quantitatively, it is not possible to give up an analysis of the missing heritability problem. This concerns the relationship between empirical data and theoretical models: molecular research provided powerful tools for behavioral genetics studies, and if one looks at the available data, one might think that geneticists are finally walking towards the right way to explain the hereditary basis of behavioral phenotypes. The missing heritability problem highlights that this assumption could be wrong, leaving room for further theoretical inquiries.

1. The Quest for Genes

Since the 1980s, quantitative genetics and molecular genetics have come together to identify genes for polygenic traits, technically called *Quantitative Trait Loci* (QTL). In the early twenty-first century, almost everyone assumed that sequencing human genome

would have paved the way to the discovery of the genes "for" the trait known to be highly heritable. Unfortunately, linkage and candidate-gene studies entailed small samples, were genetically unspecific, and unreliable in replication (see Turkheimer, 2011). Association studies (such as genome-wide association, GWAS) came to represent the most promising methodology for seeking genes "for" complex phenotypes (see Chapter 1).

Association is considered a very powerful method to investigate the genetic architecture of complex traits and diseases: several hundred thousand to more than a million single nucleotide polymorphisms (SNPs) can be assayed in thousands of individuals. In the past decades, association studies identified hundreds of genetic variants associated with investigated conditions.

"The genome-wide association […] method represents an important advance compared to 'candidate gene' studies, in which sample sizes are generally smaller and the [genetic] variants assayed are limited to a selected few, often on the basis of imperfect understanding of biological pathways and often yielding associations that are difficult to replicate. GWAS are also an important step beyond family-based linkage studies, in which inheritance patterns are related to several hundreds to thousands of genomic markers. Despite many clear successes in single-gene 'Mendelian' disorders, the limited success of linkage studies in complex diseases has been attributed to their low power and resolution for variants of modest effect" (Manolio et al., 2009, p. 747).

In the case of yes/no traits (e.g., schizophrenia), these methods compare allelic frequencies for groups of individuals carrying the trait versus controls. In the case of quantitative traits (e.g., IQ), they compare low-scoring versus high-scoring individuals. GWAS focus on SNPs, individual segments of DNA nucleotides for which variation among individuals only includes two alleles from the available four (A, T, C, or G). SNPs are, then, indicators of genetic variation.

"An allele in a gene is said to be associated with a trait if it occurs at a significantly higher frequency in the affected individuals compared to the control group (i.e., when the null hypothesis of equal allele frequency across groups is false). […] A positive association can occur: (1) if the allele is actually causing the disease; (2) if the allele is in linkage disequilibrium[54] with the disease-causing gene; (3) due to population admixture[55]" (Eley & Rijsdijk, 2005, p. 1043).

Hundreds of thousands of DNA markers must be genotyped to scan the genome thoroughly and find reliable associations (around 500,000 SNPs are needed on very large samples, thousands to tens of thousands of people) (Plomin et al., 2013, pp. 139-142). This is a significant point in relation to the sort of genetic variants explored by GWAS. Single-nucleotide polymorphisms are generally called 'common variants'. The majority

[54] Linkage Disequilibrium (LD) is the statistical association of alleles at different loci. Two alleles are in linkage disequilibrium if they are not inherited separately, breaking up the Mendel law about independent assortment.

[55] Population admixture is often called 'population stratification'. In mixed populations, if a trait is more frequent in one ethnic group it will show positive (spurious) associations with any allele which is more common in that group.

of the association studies starts from the common disease-common variant hypothesis (CD/CV), according to which common genetic variants (allele frequency in the population >1%) could have a role in the etiology of common diseases. The hypothesis is generally based on evolutionary reasoning:

> "Rare Mendelian diseases are almost always caused by a spectrum of rare mutations, because selection acts strongly against these alleles. […] The CD/CV hypothesis rested on the following premise: because the vast majority (~99%) of genetic variance in the population is due to common variants, the susceptibility alleles for a trait will include many common variants except if the alleles have had a large deleterious effect on reproductive fitness over long periods. For common diseases or traits, many susceptibility alleles may have been only mildly deleterious, neutral or even advantageous. Examples may include diseases of late onset, diseases resulting from recent changes in living conditions such as diabetes and heart disease, morphological traits, and alleles with pleiotropic effects that result in balancing selection" (Lander, 2011, p. 5; see also Reich & Lander, 2001).

So, the CD/CV model assumes that genes act as risk factors by means of small and additive effects. In other words, there is no single variant which is individually sufficient to cause complex common disease. This might sound familiar: in fact, McClennan and King (2010) attribute the hypothesis to the Galtonian tradition and to the following QuAdM provided by population geneticists (see Chapter 2).[56]

The literature about GWAS and common diseases concerns many diseases (for some reviews, see Chabris et al., 2013; Hirschhorn et al., 2002; Lander, 2011; and Visscher et al., 2012). Zuk et al. (2012) estimated that, at that time, GWAS identified more than 1,200 loci harboring genetic variants associated with at least 165 common human diseases and traits.

However, one might notice that the empirical findings on physical and medical traits are generally more encouraging than those on behavioral phenotypes. On the one hand, it is easier to find common variants associated with the formers. On the other hand, the identified variants account for only a little percentage of the overall genetic variance and of the heritability of the analyzed behavioral traits. The mismatch between the heritability due to detected variants and the overall heritability is known as 'missing heritability'. The missing heritability problem affects, to some extent, any genetic research which employs GWAS. Let us look at this issue in more details.

[56] Apart from evolutionary reasoning, the suitability of the CD/CV model to account for common diseases justifies a methodological choice of GWAS, such as the fact that GWAS looks for risk alleles in the common frequency spectrum (Risk Allele Frequency, RAF, .0.3). This is, however, also a technical limitation. We should also consider that to define common and rare variants is somehow arbitrary. For instance, Dickson et al. (2010) define rare variants as having risk allele frequency (RAF) 0.005–0.02 and define common SNPs to be representative of those used in GWAS studies (minor allele frequency, MAF.0.05).

2. The Missing Heritability Problem

As it has been said in Chapter 1, heritability analysis represents the most important methodology of classical quantitative genetics.[57] High heritability for a given trait is often understood in two ways: a) the evidence that genes greatly influence the examined trait; and b) a way to assess the magnitude of genetic and environmental influences on it. Therefore, high heritability represents the previous condition to make sense of the quest for genes via molecular techniques: if the link between heritability and genetic influence holds, and if heritability for a trait is high, then a great amount of the variation for a trait in a population is supposed to be due to hereditary factors. Then, it might be possible to find molecular genetic variants that influence the trait (see Chabris et al., 2012, pp. 1-2).

Despite initial enthusiasms, applying GWAS to highly heritable traits did not give the expected results. It is worth underlying two different aspects:

1) Missing Variance: Genetics analyses do not identify appreciable genetic effect sizes associated with most of the interesting traits, leaving unexplained a large proportion of the genetic variance of those traits. I call this aspect, with Gibson (see Eichler et al., 2010), 'the missing variance problem' (hereafter, MVP);

2) Missing Heritability: The SNPs associated with the analyzed trait account for a small proportion of the trait' heritability. This is the well-known missing heritability problem (hereafter, MHP).

Let us see some details about the MVP. The available technologies can place upward of a million SNPs on a single chip. As Turkheimer said,

"the availability of these inexpensive chips renewed expectations that the discovery of the allelic molecules underlying genetic variation, and thus the establishment of the genetic etiology of complex traits, was finally at hand, but […] it hasn't come to pass. What has happened instead is that for any given characteristic we have discovered a handful of SNPs that appear to be in [linkage disequilibrium] with an unknown but certainly very large number of genes, which are more or less predictably correlated with an outcome of interest. For any given SNP, effect sizes are generally less than 1% of the phenotypic variance, even for something as uncontroversially heritable as height. The genetic mechanisms of the major heritable syndromes and traits have not been found" (Turkheimer, 2011, p. 231).

Let us take human stature. The heritability of height is very high (~80%) and, among other quantitative physical traits, height is relatively well-studied. Despite this, Visscher

[57] Heritability is the proportion of phenotypic variance that is attributable to genotypic variance, a value between 0 and 1. In other words, a heritability index evaluates to what extent variation in a phenotypic trait—within a population, in a specific environment—could be explained by genetic variation among individuals. If heritability is high for a trait in a population, the population variance for that trait is statistically related to significant genetic differences among individuals in that population. If heritability is low, the variance is probably attributable to environmental variance.

et al. (2008) have found at least 40 loci associated with human height, but they explain only about 5% of phenotypic variance despite the study of tens of thousands of people. In Weedon et al. (2008) the sum of 20 SNPs that had reached the rigorous levels of significance employed in such studies accounted for 2.9% of the variation in height, or a difference of about 2 inches in the heights of individuals falling at the lowest and highest 5% of the SNP scale. Lettre et al. (2008) found 12 SNPs capable to explain only 2% of the variation and about 11/2 inches between the extremes. The three studies published by Gudbjartsson et al. (2008) Lettre et al. (2008) and Weedon et al. (2008) included, in sum, 65,000 participants, each compared in respect with a half million SNPs. None of the individual SNPs associated to height accounted for as much as 1% of the variation in the trait, and the handful of SNPs that reached statistical significance in more than one of the studies accounted for around 3% (see also Turkheimer, 2011, p. 232).

Let us now turn to the MHP, which does not regard phenotypic variance, but rather, heritability.

"Recent reports suggest that combining all known SNP associations for any trait explains a small proportion of heritability, ranging about 5 percent (Manolio et al., 2009) to, at most, 20 percent of the known heritability (Park et al., 2010). This gap between the genome-wide identified associations and heritability has become known as the missing heritability problem (Mahler, 2008)" (Plomin et al., 2013, p. 141).

"The loci found to be associated with a given trait at a strict threshold of statistical significance typically account for only a small proportion of the trait's heritability (as estimated from traditional studies of the correlations between close relatives), and this discrepancy has led to much discussion of missing heritability" (Lee & Chow, 2014).

To summarize, the MVP concerns the detected genetic influence on a trait's variance in a population. Instead, the MHP represents the fact that detected genetic variants cannot account for heritability indexes calculated by classical quantitative genetics. Since variance and heritability are two strictly related parameters (see Chapter 1), the difference between the MVP and the MHP is not generally emphasized. The following questions, then, are supposed to coincide with each other:

a) If GWAS can identify the causally relevant alleles for a trait, and a trait is related to genetic factors, why the detected SNPs do not account for a large proportion of the phenotypic variance?
b) If GWAS can identify causally relevant alleles for a trait, and if heritability is high for the trait, why the detected SNPs do not account for a large proportion of the heritability estimations in classical experimental designs (e.g., twin studies)?

Why could the two problems be (and sometimes are) treated as the same problem? The answer is straightforward: heritability and genetic influence are considered as strictly related. Moreover, a strict relationship occurs also between variance and heritability. For

now, I consider the missing heritability problem as a comprehensive problem, as it is often done.[58]

Let us deepen the situation a bit more. To begin with, some exceptions to the MHP do exist. For instance, few alleles explain a substantial proportion of variance for age-related macular degeneration (Kraft & Hunter, 2009). In general, the proportion of the explained heritability is higher for medical conditions than for stature (see Lander, 2011). Missing heritability is nonetheless a problem for the majority of the investigated heritable traits, among which medical conditions are not an exception.

"Although GWASs have proven successful in identifying regions of the genome harboring variants that contribute to complex phenotypes and diseases, for most traits the effects of all associated loci account for a small proportion of the estimated heritability. With the exception of age-related macular degeneration and type 1 diabetes, for which collectively the proportion of heritability explained to date is approximately 50% and 80%, respectively […], most complex disease variants identified to date together account for much less of the trait variance" (Stranger et al., 2011, p. 376).

But the less encouraging cases are behavioral phenotypes. Indeed, the genetic effect sizes revealed for physical traits are generally higher than, for instance, general cognitive ability (Chabris et al., 2013). Any case of pigmentation (eye darkness, hair darkness, red hair, freckling and skin darkness) presents *at least* an allele with an appreciable effects size (between .998 and .267). Stature is a borderline situation (effects size between .120 and .305). The case of behaviors is quite different: effects sizes for intelligence have been attested between .062 and .026; neuroticism and conscientiousness revealed respectively a single allele of .014 and .038 effect size. Ironically enough, liberal versus conservative attitude presents an allele of .552 effect size.

Strikingly, behavioral traits are generally as heritable as physical and medical traits, or even more (IQ is a renowned example of a very highly heritable trait). General cognitive ability presents a heritability ~80%; externalizing behavior in females, 73%; alcohol use in females, 75%. Conversely, the only physical or medical traits with a comparable heritability are height (80%) and the lipoprotein A level (95%) (see Boomsma et al., 2002; Chabris et al., 2013).

"Though most behavioral traits are moderately to highly heritable, the genes that influence them are elusive: many published genetic associations fail to replicate. With physical traits like eye color and skin pigmentation, in contrast, several genes with large effects have been discovered and replicated" (Chabris et al. 2013, p. 1).

It should not be surprising that the literature about general intelligence frequently shows non-replicable associations—both for candidate-gene and genome-wide studies—

[58] However, heritability is the proportion of phenotypic variance due to genetic variance; then, it does not coincide with variance. I later stress the fact that some issues concern the MVP, while others concern the MHP.

and missing heritability (see e.g., Butcher et al., 2008; Davis et al., 2010; Chabris et al., 2012; Plomin et al., 2013).

The MHP has been often perceived as a great failure for genetic research:

"The fact that faster, cheaper, and more powerful methods of genotyping have led to fewer, smaller, and less reliable findings on the connection between genes and behavior, despite the near-certainty that such connections exist, stands as one of the disappointments of 21st century science" (Chabris et al., 2013, p. 15).

Nevertheless, most scholars take the MHP as a supporting evidence for the QuAdM: complex traits are related to thousands of alleles of small effect (smaller than previously thought), and hence none of them accounts in isolation for an appreciable proportion of the heritability estimated by classical quantitative genetics.

"Despite an adequate sample size for detecting large effects and despite high-precision measurements, we found few associations between SNPs and traits at an appropriately stringent significance threshold. Since many of our measured phenotypes (including our behavioral phenotypes) are known to be heritable, the absence of strong associations in our data indicates that—aside from pigmentation—both physical and behavioral traits are mainly affected by numerous genes with small effects" (Chabris et al., 2013, p. 7).

However, this conclusion is not necessarily appropriate. Many attempts have been made to make sense of missing heritability. Stranger et al. (2011) summarize the most widespread proposals:

- Effect sizes of associated variants may be underestimated due to incomplete linkage disequilibrium between causal variants and marker SNPs (MVP);[59]
- Low-frequency polymorphisms or rare variants that are not captured by current genotyping platforms may contribute a portion of the unexplained heritability (MVP and MHP);
- Heritability may be overestimated, with epistasis, epigenetics, and genotype-environment interactions contributing to trait heritability (MHP);
- Many additional, currently undetected small effects may together comprise a significant contribution to heritability (MHP).

Evolutionary explanations have been also adumbrated: genetic effects size could be very small for keeping genetic variability under the eye of natural selection; any large genetic effect, it is said, would disappear (Manolio et al., 2009) (MVP).[60]

Some of these explanations appeal to mere methodological issues. For instance, many refer to the dimension of the populations (too small) or to the available genetic

[59] Some of these proposals are related to the MVP, while others to the MHP.
[60] However, the authors recognized that such an evolutionary dynamic would manifest a low, rather than missing, heritability.

markers (too few, or the wrong ones). The most explicative strategy, in this respect, reminds that GWAS is underpowered to detect very small genetic effects insofar as they do not reach the threshold for statistical significance (*p*-value):

"The distribution of effect sizes for common variants affecting human complex traits is highly skewed toward small effect sizes, and the true distribution is likely even more skewed than the empirical distribution, as GWAS are underpowered to detect small effects. The identification of additional loci of small effect will be partially addressed through meta-analysis of multiple GWAS, but given stringent significance thresholds, it is unlikely that GWAS will ever be powered to identify the full spectrum of small effects" (Stranger et al., 2011, p. 376).

Within human height research, it is renowned the attempt of Yang et al. (2010) and Visscher et al. (2010) to identify the missing heritability. The authors remind us that GWAS act under a very stringent *p*-value to reduce the occurrence of false positive associations. This may cause many real associations to be missed, especially if individual SNPs have small effects on a trait. Changing some parameters, the authors explained the whole heritability as such: less than 1% is due to any single SNP; around 10% depends on linear combinations of statistically significant SNPs; around 45% is attributable to all genetic variants in LD with hundreds or thousands of SNPs on currently available chips; the rest is (speculatively) attributable to rare genetic variants. A similar interpretation has been given by Wood et al. (2014). These proposals have been enthusiastically accepted, for instance, by Turkheimer (2011), who thanks the "Visscher's program" to have proved the validity of classical quantitative genetics, despite the numerous criticisms.[61]

However, as Hemani et al. suggest (2013), whether these explanations are the complete story deserves exploration. Indeed, a different line of thought characterizes those scholars who analyzed the MHP from a broader theoretical perspective by questioning the theoretical assumptions characterizing quantitative genetics. In other words, some of the proposed solutions do not appeal to the mere need for an improvement of molecular techniques. Rather, they try to revise the strict quantitative view that characterizes GWAS (see Nelson et al., 2013).

Two sets of the available hypotheses are particularly interesting for my aims: the first one explores the possibility that common variants are not the right kind of genetic elements to look at for understanding the genetic architecture of complex traits; the second one entails various remarks about the reliability of heritability analyses, such as the fact that this method overlooks genes-environment interactions. In the following paragraphs, I analyze these proposals and I show that quantitative genetics has imported statistical

[61] Turkheimer's argument is striking. According to the first premise, heritability analyses have been criticized. The second premise notices that the MHP has been blamed for the unsuitability of heritability analyses. The third premise states that the Visscher's program explains how it might be possible to account for the missing heritability. Conclusion: classical heritability analyses are reliable. As far as I can see, this is neither an evidence for the validity of heritability analyses nor a demonstration that the missing heritability has been finally detected.

terms and conceptions in an ontologically-oriented framework. Therefore, taking the MHP for supporting the QuAdM does not seem to be a promising strategy.

3. Rare Variants *versus* Common Variants

One might notice that the validity of the CD/CV model, and what its best domain of applicability is, are open empirical questions. Are common variants relevant to phenotypic traits? If yes, for which traits? Are all phenotypic traits similar in respect of their genetic architecture? Might it be possible to account for the missing heritability (or for the missing variance) by appealing to the presence of rare variants of large effect sizes? If yes, why GWAS did not find them?

Rare variants are usually supposed to be high-penetrant on phenotype and associated with rare diseases only. Classically, they pertain to the candidate-gene approach.[62] However, we are now dealing with a different problem: are genome-wide scans systematic enough to detect mutations that are rare in the population? The answer is likely negative (see Eyre-Walker, 2010). Rather, GWAS looks for genetic variants with an appreciable frequency in the population (polymorphisms, or common variants, according to the CV/CD theory). As Wray et al. summarize,

> "one plausible explanation [of the MHP] is that rare variants, which existing GWAS platforms are not designed to capture, make significant contributions to the heritability of many traits and diseases. It is indeed likely that many multifactorial and heterogeneous phenotypes will be influenced by a diverse array of genetic factors that span the spectrum from private mutation to common variant" (Wray et al., 2011, p. 1).

Biological explanations are required to evaluate the rare variants *versus* common variants problem. Several authors point, indeed, to a different model that we may call "common disease-rare variants" (CD/RV). Dickson and colleagues (2010) argued that rare variants might explain some of the heritability that is currently missing. Moreover, they could be the cause of a proportion of detected associations between complex traits and common SNPs from GWAS. Conversely, Wray and colleagues argued against this idea, but they recognize that:

> "undoubtedly, part of the missing heritability is explained by imperfect LD between the genotyped SNPs and causal variants, including rare causal variants and including multiple rare causal variants concentrated in relatively short genomic regions. Dickson and colleagues give six examples of known synthetic associations detectable in GWAS but generated by rare causal variants, providing compelling evidence for their existence" (Wray et al., 2011, p. 9).

[62] This approach relies on the ability of the researcher to make a good hypothesis about where to look along the genome (see Chapter 1). However, failing to replicate candidate-gene results does not imply that there are not genetic effects of large size: candidate-gene approach has just often failed in identifying them.

Both Eichler and Leal (see Eichler et al., 2010) refer to the importance of genetic variants of large impact. Those variants would be individually rare but collectively common. In the same publication, Flint advances a more elaborated argument by hypothesizing different genetic architectures for different traits. According to him, it is true just in a broad sense that complex traits are commonly under the control of additive genetic effects—this regards averages across all phenotypes in different populations. Conversely, a more fine-grained analysis would reveal that the genetic architecture of disorders like schizophrenia is not the same of height, weight, and IQ. By modelling fitness, for instance, disorders turn out to considerably lower the fitness. This can have consequences on the genetic effect sizes involved, raising doubts on the appropriateness of the SNPs-approach. Indeed, it is possible that GWAS are suitable to account for some traits (classical polygenic traits) and not for others (common diseases) (see Lander, 2011, on the results about medical and psychiatric diseases).

More radically, it is possible that every common variant has been successfully detected by GWAS but, perhaps, common variants neither represent the whole story nor an interesting part of it. Goldstein is particularly adamant in this respect:

"I assume that all SNPs yet to be discovered have weaker effect sizes than the weakest so far found. Though the strongest SNP may have been found, many SNPs could remain unidentified in the range of the lower effects that have been determined. […] The sample sizes that have been studied for height, however, range from 14,000 to 34,000. At the lower sample size, the power of detection is 90% for the largest effect size; for effect sizes as small as 0.05%, the largest sample size provides a 10% chance of detection. Even if we conservatively assume that all remaining unidentified variants influencing height each explained as much as 0.05% of the variation, 1500 such variants would be required to explain the missing heritability. These calculations also assume that the effects of 'height SNPs' are additive. If variants show meaningful interactions, a somewhat stronger genetic effect could emerge among variants with small individual effect sizes" (Goldstein, 2009, p. 1697).

The author concludes that there are probably either no more common variants to discover or no more that are worth discovering. If Goldstein was right, it would be reasonable to think that the contribution of common variants has been overestimated.

McClennan and King (2010) give an elaborated interpretation of why GWAS are unlikely suitable to detect the relevant alleles associated with diseases. They argue that human populations are very heterogeneous from a genetic point of view, and then it is very hard to find stable statistical associations between genetic factors and phenotypic traits:

"in molecular terms, we suggest that human disease is characterized by marked genetic heterogeneity, far greater than previously appreciated. Converging evidence for a wide range of common diseases indicates that heterogeneity is important at multiple levels of causation: (1) individually rare mutations collectively play a substantial role in causing complex illnesses; (2) the same gene may harbor many (hundreds or even thousands) different rare severe mutations in unrelated affected individuals; (3) the same mutation may lead to different clinical manifestations (phenotypes) in different individuals; and (4) mutations in different

genes in the same or related pathways may lead to the same disorder. […] Causality in this context can almost never be resolved by large-scale association or case-control studies" (McClennan & King, 2010, p. 210).

In order to explain individual cases, they refer to the role of rare genetic variants, rare structural genomic mutations, and *de novo* mutations:

"[Several empirical] results suggest that a substantial portion of autism and schizophrenia is caused by individually rare mutations—small and large—that disrupt the function of genes operating in critical neurodevelopmental pathways. Several genomic hotspots have been implicated in more than one psychiatric or neurocognitive phenotype (Cook and Scherer, 2008). The converse is also true: the same mutation may be associated with different psychiatric disorders or with no illness at all" (McClennan & King, 2010, p. 212).

To stress the point, phenotypic similarities could be not related to genetic similarities. The genetic heterogeneity of human populations and the multiple-realizability of phenotypic traits are compelling reasons for thinking that association studies cannot fulfill their goals. One might assume that heredity has a great impact on a disease or on a trait, but this does not grant that we can identify genetic factors with statistical methods.

This point is better clarified in relation to the discussion about linkage. To demonstrate the biological importance of risk variants detected in GWAS, it has been said that a risk variant is not itself a critical functional variant, but it is supposed to be in linkage disequilibrium with a rare mutation of clear effect. The principle is that linkage disequilibrium of risk variants with rare mutations of functional effect leads to statistical associations in genome-wide association studies. However,

"the hypothesis is reasonable if genetic heterogeneity of the disease is very low in the series of cases under study. That is, a significant association in a GWAS may reflect a functional mutation by LD *if* the (unknown) functional mutation is responsible for a substantial proportion of the illness in the cases surveyed" (McClennan & King, 2010, p. 215).

Last but not least, the authors remind the importance of neutral mutations by explaining how evolutionary forces have likely led most common variants to be neutral.

To summarize, the very question on the line is: how much do we have to zoom into the genome? Quantitative genetics answers by referring to single nucleotides—after all, this is the most rational answer if we start from the QuAdM. However, my brief review about the possible role of rare genetic variants of large effects suggests that SNPs are not the best genomic elements to look at for studying genotype/phenotype correlations. The concerns highlighted above suggest that the validity of the QuAdM is not empirically demonstrated by association studies (as several scholars think). Rather, (missing) empirical findings point to its inadequacy in accounting for complex phenotypes. The assumption that a multitude of small alleles influence phenotypes will be strongly questioned if rare variants of large effects (high-penetrant), de-novo mutations, and structural genomic

changes, will turn out to be relevant for phenotypic variation of complex traits. This would also point to the qualitative approach I suggested in Chapter 2.

Appealing to variants of large effect sizes does not imply the existence of monogenic traits. It is possible that large genetic effects (some alleles or structural genomic variants) are capable of taking down development with an individual appreciable effect, being related to the occurrence of pathological conditions (see Chapter 5). Unfortunately, GWAS seems to be unable in identifying rare variants: the burden of proof shall not lie in this methodology.

4. The Role of Interactions

The MHP does not solely pertain to contemporary molecular research, but rather it arises in relation to classical heritability studies. As I mentioned, several solutions have been proposed to understand the portion of heritability that GWAS cannot detect. Zuk and colleagues (2012) rephrase the problem in the following terms:

"geneticists define the proportion of (narrow-sense) heritability of a trait explained by a set of known genetic variants to be the ratio $\pi_{explained} = h^2_{known}/h^2_{all}$, where (i) the numerator h^2 known is the proportion of the phenotypic variance explained by the additive effects of known variants and (ii) the denominator h^2 all is the proportion of the phenotypic variance attributable to the additive effects of all variants, including those not yet discovered. The numerator can be calculated directly from the measured effects of the variants, but the denominator must be inferred indirectly from population data" (Zuk et al., 2012, p. 1193).

It is not surprising that several scholars have called into question the general reliability of heritability analysis. In fact, almost anybody recognizes how controversial the topic is: it is generally acknowledged that the relationship between heritability, heredity, and genetic causation is problematic (see Chapter 4). In the context of GWA literature, the main problem involves the additive genetic variance. However, additive genetic variance is often evaluated as insensitive to significant biological phenomena usually clustered under the label 'interactions'. Therefore, it is straightforward why many of the attempts to account for the MHP refers to this aspect.

Hirschorn et al. (2002) and Moore (see Eichler et al., 2010) propose that the undetected heritability might be due to gene-gene interactions (e.g., dominance and epistasis) and gene-environment interactions that cannot be the target of a statistical analysis of SNPs. Indeed, epistasis occurs whereby the effect of one locus depends on the genotype at another locus. Gene-environment interactions, instead, concern the differential reaction of genotypes to the environment. Both are classical sources of non-additive variation (see Chapter 4 for more details).

Slatkin (2009) proposes to look at epigenetic inheritance as a source of variability among relatives and, then, as a part of the heritability. Inherited epigenetic changes that

modify genetic expression may contribute to disease risk and similarities among relatives, albeit they would not be detectable in GWAS (see also McCarthy & Hirschhorn, 2008).[63]

Moore and Nadeau (see Eichler et al., 2010) consider the complexity of biological systems and refer to sources of biological variability that are not under the microscope of quantitative genetics. Moore refers to networks of interactions at any level of organization which would make it impossible to explain phenotypic variation by merely adding together independent genetic effects. Nadeau refers also to the context-dependent nature of genetic expression. Since this complexity is attested for simple organisms like yeast, why it should be different for humans?

All these concerns are rapidly dismissed by Yang et al. (2010), who remind us that genetic interactions and gene-environment interactions are not included in the narrow-sense heritability estimations (h^2) (see Chapter 1). Then, they argue, it is not possible to explain the missing heritability by means of those phenomena: what should be explained is only the missing phenotypic variance due to additive genetic variance.

The validity of this argument seems to be undeniable, but we must notice that Yang and colleagues assume that the MVP and the MHP do coincide. However, while it is not possible to appeal to non-additive effects to account for MHP, it is feasible to appeal to them to account for the MVP. I would then suggest reading those hypotheses about the missing heritability as, in fact, hypotheses about the missing variance.

This is, for instance, Gibson's strategy (see Eichler et al., 2010), who casts doubts on the fact that the real issue at stake concerns missing heritability, suggesting, rather, a "missing variance problem". Here, it is possible to appeal to genes-environment interactions, which are not tracked by heritability analyses and GWAS.[64] The fact that the phenotypic variance is not accounted for by the detected SNPs may point to two different solutions: either the genetic effects are smaller than we thought, or the genetic effects are larger and more interactive with each other than we thought.[65]

Zuk et al. (2012) suggest that, in the equation $\pi_{explained} = h^2_{known}/h^2_{all}$, heritability has been overestimated and, then, the portion of the explained heritability has likely been underestimated. They call this gap "phantom heritability". According to them, quantitative genetics has paid little attention to genetic interactions and has wrongly limited itself to the analysis of linear combinations between alleles. Therefore, geneticists have wrongly presumed that h^2_{all} does coincide with what the authors call h^2_{pop}, that is, the heritability estimated via correlational analysis of populations. The problem is that the

[63] However, Slatkin highlights that epigenetic inheritable changes would contribute to missing heritability only if they were more common than genetic mutations. The author concludes that further empirical research is needed to assess the weight of epigenetic inheritance for risk disorders. Bourrat et al. (2017) and Tal et al. (2010) refer to a method for testing whether and to which extent non-DNA factors contribute to correlations among relatives. The authors propose to include epigenetic factors as part of the picture.
[64] This is a compelling reason for distinguishing the MVP from the MHP, as I suggested above.
[65] Gibson, however, admits the co-existence of these two aspects.

two heritability indexes do not coincide unless a trait is solely under additive genetic effects. By considering interactions, h^2_{pop} will be expected to be much larger than $h^2_{all.}$[66]

Similar considerations are made by Hemani et al. (2013). Epistasis, they say, might contribute to narrow-sense heritability in two ways: 1) by generating real additive variation as marginal effects from higher order genetic interactions; or 2) by creating a statistical illusion of additive variance through confounding between non-additive and common environment effects in twin study based estimates.[67] However, broad-sense heritability, involving both additive and non-additive genetic effects, is intractable for non-clonal populations. Then, they argue, quantitative methods are strongly biased in their assumption about additivity, leading to an overestimation of additive effects:

"The additive framework that is used in GWA studies follows Occam's razor, employing the hypothesis that introduces the fewest new assumptions (i.e. non-additive variation cannot be estimated, thus SNPs are not modelled to have non-additive effects)" (Hemani et al., 2013, p. 2).[68]

These recent remarks bypass the counterattack provided by Yang et al. (2010). Furthermore, they suggest that quantitative genetics is unable to account for several biological phenomena: GWAS do not represent good methodology to study the G-P map, being this much more complex than how the quantitative-additive view holds. The MVP is not a genuine problem, in a sense.

Conclusion: Much Ado About Nothing?

Association studies are statistical analyses which do not seek for explanations in a strict sense (e.g., causal, functional, mechanistic, or developmental). Rather, GWAS are explicitly unbiased or agnostic and de-emphasize considerations of biological plausibility, physiology, and etiology. As McClennan and King argue (2010), applying pure statistical analyses leads to inevitable shortcomings:

"To date, [GWAS] have published hundreds of common variants whose allele frequencies are statistically correlated with various illnesses and traits. However, the vast majority of such variants have no established biological relevance to disease or clinical utility for prognosis or treatment. […] Very few published risk

[66] The authors propose a method to address h^2_{all} that is consistent not only with additive effects but for any type of genetic architecture. However, Stringer et al. (2013) criticize some of the model's assumptions as biologically implausible.

[67] The authors analyze epistasis as the core mechanism which allows the maintaining of the genetic variability under the action of natural selection (remind that one of the reasons to prefer the QuAdM over Mendelism, is exactly the need to explain the maintenance of biological variability, see Chapter 2). Moreover, according to the authors, if additive variation is observed (a small amount, however), then there are likely non-additive genetic components that allow it to persist in the population.

[68] Similar reasoning might be traced also in early Modern Synthesis (see Chapter 2). According to the authors, with sample size growing and computational tools, a reexamination of the Occam's principle in this context should be made.

variants lie in coding regions or in promoters. Far fewer have been shown to alter the function of any of these sequences. How did genome-wide association studies come to be populated by risk variants with no known function?" (McClennan & King, 2010, pp. 2010-215).

According to the authors, a major limitation of GWAS is the lack of any functional link between the most risk variants and the disorders they are supposed to influence. This is not surprising if we remind that in the human genome approximately 35% of base pairs lie in introns, and therefore the same proportion of SNPs have no functional role at all. The suitability of statistical analyses is called in question by behavioral geneticists themselves:

"To make progress, we should shift away from the traditional model of epidemiology via statistical significance testing, in which large significant correlations are the standards of success and worthy of newspaper headlines, while negative results are considered a failure and destined for the file drawer" (Chabris et al., 2013, p. 15).

"What is missing from contemporary genomics is not heritability, but a meaningful link between statistical and etiological models of the transmission of complex traits" (Turkheimer, 2011, p. 232).

However, the agnostic approach adopted in GWAS relies on the mathematical foundation of contemporary quantitative genetics. As I have suggested in Chapter 2, many problems can be due to an ontological mistake for which geneticists have come to be committed to the QuAdM as a real description of how genes work and how organisms develop.

As the missing variance, the missing heritability problem has been blamed to represent much ado about nothing. This line of thought is consistent with the radical thesis adumbrated by Lander (2011), who suggests that some of the missing heritability may simply be an illusion:

"Heritability is estimated by applying formulae for inferring additive genetic effects from epidemiological data. The estimates may be inflated because the methods are not very effective at excluding the (nonlinear) contributions of genetic interactions or gene-by-environment interactions, which are likely to be significant" (Lander, 2011, p. 193).

Therefore, a gloomy perspective arises if one matches the MHP with the heritability debate. What if molecular genetics misled about classical studies by starting from fragile theoretical assumptions and empirical data? What if the QuAdM is simply unable in framing biological phenomena? What if heritability analysis is not reliable to provide any biological explanation? Especially this last suspicion arose several times over the past decades. In the next chapter, I shall return to what is likely the most important problem in behavior genetics: assessing the magnitude of genetic influence on behaviors.

Chapter 4.
Thinking About Heritability

Since many decades, a long-running debate engages experts from several fields (e.g., biology, statistics, psychology, and philosophy) about Galton's *nature* and *nurture*. Heritability analysis lies at the core of the dispute as it has been taken as a methodology for clarifying the relationship among the two. The reason is straightforward: most scholars agree in tying heritability, heredity, and genetic causation with each other. This is, in fact, why molecular research takes heritability analysis as a starting point: a high heritability is supposed to be an evidence that a trait is influenced by genes. Moreover, heritability is supposed to be a way for estimating the weight of nature and nurture on phenotypic individual differences.

Several doubts have been raised on the reliability of this methodology: What heritability can really tell us? Is it applicable to human behaviors? Sometimes, this methodology has been judged unable to establish if a trait has genetic bases and the importance of genetic influences on complex human traits. Such criticisms, if well founded, would lead to a radical rethink of both quantitative and behavioral genetics. For instance, the critics say, if the real meaning of heritability was understood, the missing heritability problem would not have come up. In the context of the study of human intelligence, the topic of heritability attracted even more criticisms. Indeed, the majority of what we know about the genetic bases of intelligence relies upon heritability findings. Roughly speaking, one might say, genome-wide association should not be applied to phenotypic traits just because of their heritability

In this chapter, I do not mean to draw any conclusion about this debate, but rather to show that heritability analysis cannot be taken as a compass to investigate human behaviors like intelligence. It is in this respect that the heritability debate has a role to play in my inquiry about human intelligence. The assumption according to which intelligence can be studied as a quantitative phenotypic trait (namely, the IQ) largely relies on some assumed connection between heritability and genetic causation. In order to show that heritability analysis is not suitable to achieve its goals, I shall focus on several oddities arising in heritability research and on the debate about the genes-environment interplay. Both these two aspects suggest that the problem of genetic causation cannot be addressed by means of statistical analyses of phenotypic variation in natural populations. This will pave the way to seek different types of research lines that I discuss in Chapter 5.

1. The Magnitude of Nature and Nurture

During the last century, heritability has been widely used in medical and behavioral genetics even though it was born in quite a different context, that is, the context of artificial selection. Being heritability understood as an index of genetic variability, it was originally a measure of the accuracy of selection (Lush, 1945, 1949). In other words, it indicated, and still indicates in such a context, how quickly a population will react to selection (Pigliucci, 2001, pp. 8-9).

Lush distinguished two types of heritability: a) broad-sense heritability (H^2) refers to the entire genotype, which includes genes which interact non-additively; b) narrow-sense heritability (h^2), instead, only refers to additive genetic effects and, therefore, to those genes which are of interest to plant and animal breeders. Narrow heritability serves as a breeding value coefficient to predict how much a population will change over generations—since the index represents the genetic source of the resemblance between parents and offspring. For instance, if h^2 is 100% for height (i.e., all variation is genetic), then the value of the trait in the offspring will be midway between the parental values (Schaffner, 2016, p. 23).

However, as Schaffner underlies, non-genetic factors (e.g., nutrition or climate differences) may affect resemblances between parents and offspring for phenotypic traits. Lush himself understood that highly heritable traits in animals were influenced by the environment as well. As I discussed in Chapter 2, the complexity of quantitative traits has been typically linked to both their polygenic architecture and environmental influences. As a matter of fact, quantitative genetics focuses on genes and environment (assumed as synonymic terms of *nature* and *nurture*) as the most important (or the only) sources of variation among individuals.

Since it is hard to separate environmental and genetic influences, the study of quantitative traits requires complex statistical analyses, among which heritability has enjoyed a great reputation for being able to disentangle genetic and environmental effects on phenotypic variation—under specific experimental precautions. Given its apparent explicative power, heritability has been applied not only in population genetics but also in behavioral genetics. However, this transition carried remarkable differences in the application of the methodology and some underlying assumptions that deserve to be considered.

1.1. The transition towards behaviors

Focusing on human behaviors did not lead geneticists to embrace broad-sense heritability as the main target. Rather, like breeders, human behavioral geneticists have mainly remained interested in narrow-sense heritability (see Chapter 1). This might strike one as

surprising: indeed, one would say that the whole genetic variance deserves to be addressed in the study of human behaviors—especially because, intuitively, some sort of interaction between genes and environment is likely involved in complex traits. For instance, the relationship between height, light, and nutrition seems to play by these rules. So, it seems reasonable that behaviors, even if strongly influenced by genes, require to meet some sort of environmental stimulus.

Briefly, the additive genetic variance is the only source of genetic variation accounted for by models in human populations such as the ACE model (see Chapter 1). However, this is not for nothing: in non-clonal populations, broad-sense heritability represents an intractable problem since we do not how important non-additive genetic variance is (see Hemani et al., 2013; Visscher et al., 2008). From a theoretical perspective, instead, the additive variance has likely drawn more attention than the non-additive one in relation to the theoretical assumptions underlying the QuAdM.

Another important reason relies on the fact that, since the beginning, heritability analysis has been adopted for behaviors to achieve the ambitious purpose to disentangle *nature* and *nurture*, that is, estimating their *relative importance* in bringing about phenotypes. Fuller and Thompson, in their founding textbook *Behavior Genetics* (1960), appear to have been the first who extended the definition of heritability to include the estimation of "how much" of the observed variation is due to genetic factors (see Joseph, 2004, p. 143). This seems to be a very "quantitative" sort of question, on which I shall return in §2.

Another remarkable aspect of the transition towards human behaviors concerns a methodological adjustment. In human populations, we cannot manipulate variables (genotype or environment) so easy as in experimental contexts. Family studies (e.g., twins and adoption studies) have been hence taken as an alternative way to estimate heritability. In the human species, heritability is estimated using correlations among relatives:

> "Behavioral genetics analysis contrasts phenotypic similarity between related individuals (which is measured) with their genetic similarity (which is known from genetics). If individuals who are more closely related genetically also tend to be more similar on a measured trait, then this tendency is evidence for that trait being heritable—that is, the trait is at least partially influenced by genes. […] The main focus is on the covariance between relatives. […] If we measured height in sibling pairs, we could calculate the covariance between an individual's height and sibling's height. […] If the covariance is greater than zero, this would imply that taller individuals tend to have taller brothers and sisters" (Purcell, 2013, p. 379).

The correlation among certain types of relatives is compared with their (presumed) genetic similarities via a coefficient of genetic relatedness. For instance, DZ twins share, on average, half of the additive genetic variance. MZ, instead, share all their genetic makeup. Narrow-sense heritability is supposed to be twice the difference between the correlations observed for MZ and DZ twin pairs (see Chapter 1).

After the transition towards behaviors, on the one hand, heritability continued to be widely used in agriculture and in laboratory, where various factors can be controlled. On

the other hand, behavioral geneticists applied this methodology to human populations to assess the weight of genes and environment in behavioral variability. High heritability (generally between 50% and 80%) has been estimated for any interesting psychological trait—e.g., IQ, personalities traits, and mental disorders (see for instance Bouchard et al., 1990, 2003; Bouchard, 2004; Eysenck & Kamin, 1981; Wright, 1997).

Such findings have been interpreted by many scholars as evidence for the genetic theory of human behaviors by assuming a link between heritability and heredity.[69] For instance, according to Chabris and colleagues:

"if a trait is heritable in the general population, then in principle it should be possible […] to identify molecular genetic variants that are associated with the trait. General cognitive ability, or *g*, […] is one of the most heritable behavioral traits. Estimates of broad heritability as high as 0.80 have been reported for adult IQ in modern Western populations" (Chabris et al., 2012, pp. 1-2).

Heritability is also understood as an index of the extent to which environmental interventions could be able to modify phenotypic differences:

"high heritability of a trait (in a given population) often signals the worth of digging further in the sense that an important genetic mechanism may thus be uncovered. High heritability indicates that genes are strongly implicated simply by telling us that the other causal factor (environment) is, under the circumstances, relatively unimportant for the phenotypic variation. Manipulating the environments in the environmental range where high heritability prevails is unlikely to have an appreciable effect on the existing phenotypic differences" (Sesardic, 1993, p. 405).

Several authors criticized the application of heritability to human behaviors and the subsequent shift towards genetic influence, arguing that, albeit heritability analysis is a useful tool for artificial population, it is not reliable in the study of natural populations or human behaviors (see Barnes & Dupré, 2013; Block, 1995; Jacquard, 1983; Joseph, 2004; Kempthorne, 1978, 1997; Lewontin 1974, 1982; Northcott, 2006; Rose et al., 1984; Rose, 1997; Sober, 1988; Wahlsten, 1990, 1994).

The most critical period for heritability research arrived during the last two decades of the twentieth century. The well-known IQ controversy is the best example in this respect: detractors of heritability research highlighted several weaknesses of the behavioral geneticists' interpretation of such a technique. They referred to a wide range of issues, which have never been fully acknowledged by the advocates of the heritability research. For instance, doubts have been raised about the improper inference from heritability to probability; the biological meaning of the term heritability; the lack of mechanistic and developmental explanations; the social and philosophical consequences of the related empirical findings.

[69] The meaning of heritability has often been extended far beyond its proper boundaries, leading to discussions motivated by political concerns (see Block, 1995; Herrnstein & Murray, 1994; Jensen, 1969; Eysenck & Kamin, 1981; Turkheimer, 2011).

To broadly summarize, the critics state that from heritability data one cannot legitimately infer anything interesting about genetic causation on a phenotypic trait. Rather, such concepts as 'innate', 'genetic', and 'heritable' have no clear connection with each other. Moreover, to say that genetic variation is suitable in explaining the 80% of the variance of a trait in a population does not mean that this trait in an individual is 80% influenced by genes and 20% by the environment; nor it does mean that an individual has an 80% chance of inheriting it.

Nowadays, the debate is still open. Heritability analysis has been defended by "few dissenting voices" (Sesardic, 1993; Pearson, 2007; Tal, 2009)—as Downes (2015) defined them. For instance, Sesardic has argued that heritability can say something interesting even though it is a very local estimation and it does not say anything strong about causation.

Likely, the struggle has not come to an end because heritability reflects the debate about genes and environment, innate and learned, that remains of vital interest for all the life, psychological and medical sciences. In general, results coming from heritability studies, far from being abandoned, still serve as the basis for more recent genetic analyses like GWAS. Perhaps, there is no other valuable method to establish if a trait is genetically determined in human population but, by proceeding in such a way, any kind of issue related to molecular research (e.g., the missing heritability problem) could have been inherited by (or could be reduced to) some more fundamental problem related to heritability research and to how hard understanding genetic causation is. There is no simple way to tell this story. I shall start by discussing some of the chief issues that have come out from this Pandora's box.

1.2. A tricky methodology

Before approaching hard questions about methods, genes, environment, and causation, it is worth noting that, even from a semantic point of view, the relationship between the two terms of heritability and heredity is not conceptually clear. As Joseph states (2004, p. 146), it is often believed that the words 'heritable' and 'inherited' are synonymous, but it is not legitimate to move from one to the other—at least not without a clear explanation of why and how, which generally lacks in genetics literature (see e.g., Plomin et al., 2013). Hirsch (1997, p. 220) emphasizes that heritability and heredity are two different concepts that have been hopelessly conflated. Because of their assonance, he says, when we hear one of the two words, automatically we think of the other. This is the reason why Stoltenberg (1997, p. 96) proposes that the word 'heritability' should be replaced by 'selectability' in order to not confuse the technical definition of heritability with the folk one (see also Visscher et al., 2008).

The worry makes sense because, in several circumstances, heritability is misunderstood as a dispositional property of individual organisms, e.g., the capacity of an organism of inheriting a trait, or the probability that this will happen. Moreover, heritability does not regard individual development, but is rather the feature of a population.

Behavioral geneticists, however, are unlikely guilty of such a systematic and trivial misinterpretation. Rather, they seem to believe that what is good for artificial contexts is fairly good for natural populations, or that heritability does not represent a local index about a specific population which is hard to be generalized to others.[70]

To show why heritability cannot be reduced to heredity so easily as many geneticists think, I shall delineate some scenarios in which the sloppy link between the two terms may be clearly understood. The following "oddities" are due to some feature of the analyzed population or of the surrounding environment (i.e., the *genes-environment balance*, see Block, 1995).

Let us start with the case in which heritability is estimated as near to zero. Being heritability a measure of the phenotypic variance due to the genotypic variance, this happens whenever the phenotypic variance is almost entirely due to the environmental variance. What is needed is a population in which the variability among individuals for a trait is not due to genetic differences (the population is genetically homogeneous for the trait), but to environmental ones. That is, there is no genetic variation which is relevant to the character in the population. Nonetheless, low heritability does not imply the absence of genetic influence. For example, let us consider the fact that almost every individual has two arms, something strongly influenced by genes. Now suppose that in a population every individual with one arm has lost the second one during the war. Under these circumstances, the variability among individuals would be due entirely to non-genetic factors. Thus, the heritability of the two-arms phenotype would be zero (see Samuels, 2004).[71] How is it that a heritable trait has a heritability close to zero?

Let us now turn to cases in which heritability is equal to 1. Being heritability a measure of the phenotypic variance due to the genotypic variance, this happens when the entire phenotypic variation in the population is due to genetic variation. In Mendelian genetics experiments, or in knock-out studies, heritability might be 100% for the so-called differing traits insofar as the phenotypic variation in the population is due to a single allele which sharply segregates.[72] In quantitative genetics, instead, this just means that the relevant environment to a given trait does not vary at all (it is homogeneous). Nonetheless, even in similar cases, the trait might not be completely determined by genes. Rather, we could estimate $h^2 = 1$ even if the trait is just weakly influenced by genes. Take height in

[70] This is, roughly speaking, Sesardic's (1993) defense of the analysis of variance. See the discussion about the locality of heritability analysis in §2.2.

[71] Of course, the two-arms phenotype is not a quantitative trait, therefore nobody would apply heritability analysis to such a trait (Schaffner, 2016). However, this applies to quantitative traits as well. In a clonal population, every individual is genetically similar for any given trait. Here, h^2 will be zero for any trait.

[72] See Chapter 2 for details about qualitative variation. See also Plomin et al., 2013, p. 93.

plants. It is widely recognized that nutrition and climatic circumstances strongly affect this trait. However, if the environment is completely leveled, genetic variation will account for the whole phenotypic variation in the examined population. Briefly, the more the environment is homogeneous, the more heritability increases. How can a highly heritable trait be said to be strongly influenced by environment? Take now an example about IQ. If education—and only education—has an impact on IQ, and if every individual of the population has access to the same educative experience, then the heritability of IQ will tend to 100%.[73] If the experiences to which scholars of the same school are subject to are similar, heritability will be very high. Indeed, no environmental aspects are responsible for phenotypic differences (IQ levels); therefore, any difference among the scholars would be due to genetic differences (see Plomin et al., 2013, p. 92). This is true, in principle, in both high and low-level education.[74] How is it that, let us say, if we improve education then the genetic influence on IQ increases?

Another interesting thing is that sometimes (e.g., in the case of IQ) heritability changes over time, increasing with aging (Bouchard, 2004; Plomin et al., 2013). Intuitively, one would say that an index of the importance of genes cannot change over the lifetime, but remind that heritability represents the weight of genetic influences to individual differences. However, one would also say that two MZ twins are more similar when they are children: they share the same familiar environment and likely the same school and the same friends. Nevertheless, empirical data attests the contrary: once they start to live distinct life, the MZ twins' similarities seem to arise more and more, until middle age. How could the genetic influence increase over aging? How is it that MZ twins are more alike as long as their environment stop being shared?

One last thing that deserves to be treated concerns heritability and genetic expression (e.g., transcriptome, proteome, epigenome). Plomin et al. (2013, p. 152) consider gene expression part of the phenotype. Therefore, individual differences at this level may be due to both genetic and environmental differences. However, the authors concede that individual differences in gene expression are not highly heritable, since gene expression is responsive to intracellular and extracellular environmental variation.[75] There is something weird in this conclusion, especially if we consider two things: 1) heritability is high on genetic and behavioral levels; 2) heritability increases over the lifetime. In respect of the first concern, heritability is in fact estimated on a behavioral level: it does not analyze the genome in any direct way. In other words, if heritability is a genetic feature of living

[73] It is an empirical question what environmental influences are important for a phenotypic trait's variation. Without a compelling theory of environment, we cannot unambiguously interpret heritability data. Indeed, we cannot assess the degree of environmental variance for the analyzed population.

[74] However, heritability may change depending on environmental variables. For instance, it has been found that heritability of IQ increases with increasing the socio-economic status (see Visscher et al., 2008).

[75] Indeed, heritability is modest for these aspects, which implies that most of the variability in transcript levels is due to environmental factors (see Cheung et al., 2003; McRae et al., 2007; Monks et al., 2004; Sharma et al., 2005). And indeed, members of identical twin pairs become increasingly different in gene expression profiles throughout the lifespan (see Fraga et al., 2005; Petronis, 2006; Zwijnenburg et al., 2010).

beings, then it is evaluated indirectly by assuming a correspondence between heritability on higher levels (i.e., behavior) and the lower level (i.e., genes). How is it possible that heritability is high on behavioral and genetic levels, but low for the gene expression level? As regard to the second concern, how to combine the fact that heritability increases over time for behaviors and genes, and the fact that gene expression becomes increasingly different in MZ twins over time?

The extreme scenarios delineated above (heritability = 100% or 0%) are very infrequent in natural populations. More often, heritability is estimated between 0 and 1. Also in these circumstances, we can figure it out how heritability and heredity are disentangled with each other. For instance, it is quite typical to read about the heritability of political and religious preferences, or about other behaviors which we do not expect to be so much heritable as behavioral genetics attests. As I mentioned in Chapter 1, several strategies have been adopted to make sense of these data.

A plethora of scholars has responded to such findings by asking for a revolution in social and psychological science: environment has a little role, if any, in behavioral phenotypes and psychological traits (a renowned example is Jensen, 1969). Such a demanding conclusion might lead people from any research area (from psychology to education) to wonder what is the very meaning of their (merely social) professional tools and, perhaps, to start thinking about genes instead of environmental interventions to improve intellectual skills or to deal with psychopathologies. Likely, many have already fallen for it. However, it is worth asking: how can genes influence traits which seems to be so linked to social phenomena—and, to a certain extent, accidental?

To summarize, weird things happen with heritability analysis: a) if the genes-environment balance changes, heritability changes as well; and b) there is a mismatch between heritability at different levels of analysis (behavior, genetic expression, and genome). That said, it should be clear how heritability is not so easily connected with heredity, being the latter no doubt stable during the lifetime. These questions are not supposed to find easy answers unless we reevaluate the assumptions underlying heritability research. As I shall show, if one gives up some of these assumptions, several of the oddities sketched above will melt away like snow in the sun.

1.3. Assumptions in heritability research

Inspired by Occam's razor, behavioral genetics embraces several simplifications. Unsurprisingly, they look like the original niceties that facilitated Fisher in elaborating his quantitative model for the study complex traits (see Chapter 2).[76] The most important one concerns additivity. In human heritability research, additivity plays a role through the

[76] I have said quite a lot about the role in contemporary literature of an indefinite number of genes with small and equal effects (see Chapter 3). I shall here deal with the additivity assumption thoroughly.

lenses of the ACE model, the most adopted model in twin studies.[77] Far from being a mere mathematical model, ACE is supposed to reflect real biological aspects (see the discussion about model fitting in Purcell, 2013, p. 384). Purcell himself acknowledges that several assumptions have been made for the sake of simplicity. We can separate them into two groups:

1) **Methodological assumptions**: They concern methodological simplifications in family studies, including: the shift from artificial to human populations (see §1.2); the equal environments assumption (EEA); non-assortative mating; sex and age homogeneity of the samples; no sibling interaction.;

2) **Theoretical assumptions**: They mainly entail assumptions about the genes-environment interplay. The term 'additivity' may summarize this set of assumptions, which deny the importance of genetic interactions (epistasis and dominance), genes-environment correlations, and genes-environment interactions (see Purcell, 2013, pp. 381-383 and 399-401). A derivative assumption leads to thinking that h^2 is what we should really analyze, being genetic effects mostly additive.

I will not deepen assumptions concerning family study designs. In those cases, critical work has been already produced and, however, it is in principle possible to fix heritability studies in those respects.[78] Conversely, the shift towards human behaviors and the additivity assumption must be analyzed in more details. Let us proceed step by step.

Does the analysis of similarities among relatives represent a proper counterpart of the heritability method in artificial populations? It is not my aim to fully address this problem, but a real distinction between the two methods might exist. Jacquard (1983) distinguishes between *biometric heritability* and *genetic heritability*: the former represents the slope of the linear regression line of the measurements of the character among children on the mean of the measurements of the character for their two parents.[79] The only hypothesis, here, is the existence of a linear regression. The heritability defined in such a way is a measure based on observation which allows the prediction of the offspring

[77] The ACE model illustrates the partition of variance into three components: additive genetic variance, shared environmental variance, and non-shared environmental variance. Accordingly, three factors influence a trait's variance: additive genetic effects (A), common environment (C), and non-shared environment (E). See Chapter 1.

[78] As far as I know, only the EEA has been pointed out as a fatal fragility of heritability research (see Eysenck & Kamin, 1981; Joseph, 2001, 2004, 2006).

[79] Galton was the first to note that, in respect of quantitative characters such as stature, pairs of human relatives–for example, son and father–were connected by relations of linear regression and correlation (see Norton, 1975, p. 539).

values from those of the parents. This empirical heritability provides no causal explanation for the observed resemblance, much less a genetic one.[80] According to Jacquard, the situation is different for the other type of genetic heritability (which includes both H^2 and h^2); here, the determination of causes is, in fact, the goal.

Importantly, the author highlights that biometric heritability and genetic heritability are not related: indeed, for connecting them, several assumptions would be necessary but, from Jacquard perspective, they would take us far from reality. Indeed, genetic heritability is meaningful only under very restrictive assumptions (the ones clustered in Purcell's second group of assumptions, see above). Thus, the genetic portion of variance described by genetic heritability has no meaning in the general case. Of course, h^2 can be very important, at least, in guiding the choice of selection techniques used to improve a characteristic, but it cannot shed light on the biological mechanisms at play.

In the next paragraphs, I focus on the theoretical assumptions behind heritability research

2. Additivity *versus* Interactions

I already shown how shaky the relationship between heritability, inheritance, and genetic influence can be. However, for the sake of the argument, I shall concede that heritability is a way to establish, at least, if genes play a role in the individual differences of a trait. But is it possible, in principle, to quantitatively separate the two causal forces in the individual development? Quantitative geneticists have constantly neglected the importance of non-additive phenomena in behavioral traits. Plomin et al. (1988, pp. 228-229) state that non-additive interactions rarely account for a significant portion of variance. I list below some similar positions:

"The data from the twins reveal no interaction (in the technical sense) of heredity and environment" (Herrnstein, 1973, p. 180).

"Nothing like it has yet been found in human mental ability" (Jensen 1981, p. 124).

"The simpler additive model in most cases comes close to fitting the expectancies" (Cattell 1982, p. 66).

"We will proceed, as most workers in the field do, by ignoring the genotype × environment interaction" (Futuyma 1986, p. 197).

"So far, it has certainly been easier to talk about genotype-environment interactions than it has been to find them" (Plomin, 1986, p. 108).

[80] Remind that biometricians rejected any explanation based on theoretical concepts like causation (see Chapter 2).

"There is very little empirical support for [the] existence [of genotype-environment interactions] in the behavioral domain" (McGue, 1989, p. 507).

"If interactions were rampant, evolution (at least in sexual species) would be impossible [and that a] certain amount of additivity is a prerequisite for evolution" (Crow, 1990, p. 127).

It is worth asking whether the absence of non-additive phenomena is a genuine empirical discovery or rather a consequence of the employed methods and assumptions. According to Jencks, for instance, (1980, p. 732), virtually all analysts of variance just *assume* that the effects of genotype and environment are additive. Sesardic himself (1993, p. 407) acknowledges that the relative lack of empirical evidence for interactions may be due to difficulties in designing the tests that could detect them. It is now my aim to show the available arguments in favor of interactions.

2.1. Interactions and methodological aspects

The neglect of interactions is frequently indicated as the most damaging assumption in quantitative genetics, both in methodological and theoretical respects. Let us see why by starting with the methodological concerns.

In quantitative genetics, the problem of interactions seems to be methodologically unmanageable by its advocates own admission. The existence of interactions, indeed, would greatly reduce our ability to study complex traits: roughly speaking, if interactions were important it would be hard to evaluate *quantitatively* the role played by *nature* in phenotypic variation. In other words, interactions come stiflingly combined with the desire to resolve the nature-nurture problem.

Nowadays, it is considered a truism that the causal influence of genes and environment are both necessary for the development of any trait, but the point still remains: the nature-nurture problem is frequently understood as a matter of *how much* genes influence phenotypes. The fact that heritability comes out as a percentile index facilitates the conclusion that it gives us a quantitative appraisal of the magnitude of nature (then renamed 'genotype') on behaviors—from which it could be derived, by subtraction, the magnitude of environmental factors.[81]

Is it possible to measure the importance of genes separately from the environment? Several authors answer negatively. For instance, according to Benjamin et al. (2002, p. 334) heritability does not describe the quantitative contribution of genes to any phenotype

[81] According to Joseph (2004, p. 145), Plomin et al. (1997) have the merit to have described heritability as the statistic that estimates the genetic effect size. They defined "effect size" as the proportion of individual differences for the trait in the population accounted for by a given factor. Thus, according to Plomin et al., 'genetic effect size' and 'heritability' are synonymic terms.

of interests; Schaffner (2016, p. 12), more radically, says that heritability cannot tell us anything about how much genes affect the phenotype of a given individual.

Kempthorne (1978, 1997) submits that, for separating the role of the two causal forces, one needs to employ experimental tests. The author has in mind interventions in which the alteration of a variable brings about changes to another variable. For instance, what happens to an individual's IQ if one alters the environment from birth? As another instance, let us suppose one could transfer 25% of the genes of a Caucasian population into an African population, at gamete formation or at birth, keeping everything else the same. Then, if she finds that IQ has been raised, let us say, of 10 points, then a causal explanation would be possible. Even if data analysis about IQ were correct, what entitles one to infer causality in the sense of intervention results Kempthorne gives?

Kempthorne's account of causation concerns different-making analyses, that is, a change in a variable occurs when one manipulates another variable.[82] The analysis of correlations among relatives seems to be unable to deal with causality conceived in such a sense. The author concludes that the analysis of variance concerns mere observations (see also Jacquard's remarks about biometric heritability above).[83]

The puzzle can be solved by taking into account that heritability regards variance, which is, once again, a characteristic of a population. If heritability analysis says that genes count more than the environment for individual differences, what it really means is that genetic variance has more impact on the population phenotypic variance than environmental variance. Of course, this might sound confusing—as Schaffner noticed (2016), the concept of variance can be very tricky.

In the next paragraph, I turn to the theoretical problems related to additivity. The additivity assumption is philosophically interesting insofar as it concerns genetic causality and the G-P map, which are assumed to be additive in line with the QuAdM. As I shall show, to encompass interactions, a shift towards developmental analysis is necessary.

2.2. Interactions and theoretical aspects

The critical positions delineated above refer to methodological difficulties involved in separating nature from nurture. Other scholars refer, instead, to ontological aspects and empirical evidence. According to the main critical argument, it is not possible to measure the weight of genes because their effects are *entangled* in complex interactions with other genetic and environmental effects. It is the aim of this paragraph to explain why.

[82] This view is close to Woodward (2002) conception of causality, which is widely adopted in mechanistic theories (see Craver & Tabery, 2017; Glennan, 2002).

[83] Remarkably, as Dupré notices (2013), Mendelian genetics is instead the making-differences analysis *par excellence*. However, Woodward's account of causation has been criticized as defective in distinguishing causation from correlations (see Dupré, 2013; Russo & Williamson, 2007).

Various kinds of interactions are ruled out by the QuAdM, ranging from gene-gene interactions (G×G), to gene-environment interactions (G×E), to genotype-environment correlations (rG×E). The existence of these phenomena is generally acknowledged in behavioral genetics. However, as I explained, from a methodological point of view they are generally not implemented in quantitative models for the sake of simplification.

Ontologically speaking, instead, their importance is considered negligible for genetics studies because they are supposedly not frequent in nature. As I mentioned in §1.3, this is, however, an empirical question: quantitative genetic models rely on the simplifying assumption adopted by Fisher in his model. By saying that interactions are not important, an ontological commitment about that model is exerted (see Chapter 2).

Let us start with genetic interactions. These phenomena have been initially discovered by Bateson and Punnett, who demonstrated that two independently assorting genes can affect a trait (e.g., Bateson, 1909). They individuated cases in which different combinations of alleles from the two genes result in different phenotypes. Such cases are due to interactions between genetic products at the biochemical or cellular level (Snustad & Simmons, 2012, p. 72).

Epistasis is likely the most important type genetic interaction. It occurs whereby the effect of one locus depends on the genotype at another locus. As Snustad and Simmons explain,

> "a particular phenotype is often the result of a process controlled by more than one gene. Each gene governs a step in a pathway that is part of the process. When a gene is mutated to a nonfunctional or partially functional state, the process can be disrupted, leading to a mutant phenotype. Much of modern genetic analysis is devoted to the investigation of pathways involved in important biological processes such as metabolism and development. Studying the epistatic relationships among genes can help to sort out the role that each gene plays in these processes" (Snustad & Simmons, 2012, p. 75).

Epistasis is a classical source of non-additive variation for which we have many pieces of evidence (see e.g., Cordell, 2002; Haley & Carlborg, 2004; Nelson et al., 2013; Phillips, 2008; Schaffner, 2016). Both epistasis and dominance cannot be excluded by any model aimed at describing the real G-P map.

Let us now focus on the genes-environment interplay. This generally refers to a multitude of biological phenomena, among which are included epigenetic mechanisms (where environmental influences alter genetic expression), variation in heritability in relation to environmental changes (see §1.2), and rG×E.

Roughly speaking, genes-environment interactions concern the differential reaction of genotypes to the environment. As Oftedal says (2005), one cannot determine from the heritability analysis alone whether high heritability results from a low sensitivity of the trait to changes in environment or from a high similarity of environments in relevant conditions for the trait. For instance,

"[a low heritability] for a trait indicates that almost all phenotypic variance is due to environmental variance and that genetic differences hardly contribute to phenotypic differences in the trait. However, the reason could be that genotypes are very sensitive to environmental influence for the trait in question. It cannot be ruled out, though, that genes influencing the trait are fixed in the population, that is; all individuals in the measured population have identical genotypes influencing the trait in question. To control for this, other methods must be used" (Oftedal, 2005, pp. 701-702).

Quantitative genetics is usually blamed for adopting an unrealistic additive view of the genes-environment relationship, namely G+E. Lewontin (1974) advanced criticisms in this respect in his renowned paper *The Analysis of Variance and the Analysis of Causes*, where he emphasizes that separating causal influences on an individual level is impossible. Over time, Lewontin's paper became a manifesto of a sort of "qualitative view" of the relationship between nature and nurture. More recently, Gottlieb has inaugurated a tradition devoted to pursuing a similar path (see Hood et al., 2010). The additive framework is contrasted by many developmental biologists who emphasizes such a relationship as being G×E (see Tabery & Griffiths, 2010). According to Wahlsten (1994), the variance could be divided meaningfully into separate parts only if an individual score is a *sum* of two components. When we want to learn about the relative contributions of the *combined* effects of two or more distinct factors, partitioning of variance clearly becomes perilous.

The main reason why Lewontin's paper is important concerns the inclusion of an alternative methodology to study the G-P map, that is, the study of reaction norms (NoR).[84] A norm of reaction is a graphic representation of how phenotypes change in different genotype-environment combinations (see Figure 4.1).

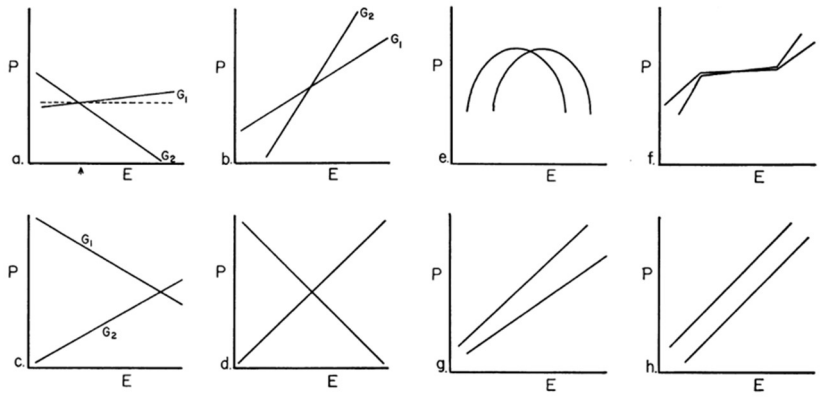


Figure 4.1: Examples of different reaction norms. From Lewontin, 1974, p. 405.

[84] NoRs have been introduced in 1909 by Woltereck to study clonal populations. The Western genetic literature ignored this method for a long time (except for a few commentators). Around the 1960s, Dobzhansky brought NoRs to the West after the collapse of Lysenko genetic program. However, the advocates of quantitative genetics, especially Fisher, had already set the premises for a widely different approach, based on the analysis of variance (see Fuller & Sarkar, 2005).

The x-axis reports the environmental parameter to which the organism is exposed. The y-axis shows the phenotypic value in response to each environmental manipulation. The slope (i.e., the NoR) shows how strong the phenotype responses to environmental changes.

A NoR graph shows phenotypes that would result from the development of chosen genotypes in a large range of environments. Each line in the graph represents data about different genotypes—the phenotype (P) is plotted as a function of environment (E) for different genotypes (G1 and G2).

This analysis clarifies that genotypes tend to react to environmental changes in a way that implies G-E interactions.[85] Indeed, there is only one type of reaction norms which accounts for additivity between genetic and environmental effects, that is, Fig. 1h. If the lines are parallel, then there is an effect of both environmental and genotypic variation but there is no G-E interaction. Here, the differences between different genotypes are constant in every environment and the differences between different environments are constant for every genotype. Thus, the change in the phenotypic outcome within different environments can be predicted from the genotype.

In situations of additivity, heritability estimates would be of general interest because, as Oftedal (2005) notices, they would not be just local analyses: the result from one environment could be generalized to other environments. Empirically speaking, however, this does not happen frequently. Rather, assuming this is the way any case works implies, for Lewontin, a circular argument: 1) additivity is a simplifying assumption; 2) results from analysis of variance are an approximation of the truth; 3) if analysis of variance finds negligible genes-environment interactions, the assumption of additivity is justified.

Lewontin's main argument against ANOVA's causal pretensions is called "the locality objection": accordingly, the effects strongly depend on the actual distribution of environments and genotypes in our sample (Lewontin, 1974, p. 402). Sesardic (1993) tried to defend the meaningfulness of heritability by pointing out that it is possible to generalize data from heritability analysis. At first, the author recognizes that heritability estimations are local:

> "When it is, for instance, discovered that in a given population the difference between organisms with genotype G1 and those with genotype G2 accounts for most of the differences in phenotypic trait P, biologists usually do not assume that a scientific law has been discovered according to which GI causes the increase in P. They are aware that they are still too much in the dark about the way G influences P, and that it would be therefore inappropriate to honor this discovery with the term 'law'. A law requires, at the very least, that something be also known about the conditions under which G causes P. But this requirement is not satisfied here" (Sesardic, 1993, p. 403).

At this point, Sesardic connects heritability and genetic causation in the local experimental situation, but he admits that such a causal relationship is far from being attested:

[85] For an explanation of how to establish interactions in NoR analyses, see Fuller & Sarkar (2005).

"Apart from knowledge that G causally influences P in this very situation, it is unclear what altered conditions would continue to sustain this causal connection. It is known only (a) that in a given situation genetic differences are strongly reflected in phenotypic differences; and (b) that, very probably, the hidden underlying mechanism by which G influences P is so complex that it is uncertain what would happen under changed circumstances. Not knowing which factors are relevant and in what way, we simply cannot be sure whether our locally discovered causal relations would reappear even under just slightly altered conditions" (Sesardic, 1993, pp. 403-404).

Nevertheless, the author suggests reevaluating heritability analysis as informative also for other populations:

"Granting all this, however, by no means justifies the extreme thesis that a heritability estimate obtained for a given population has no informative value for other populations. On the contrary, we can concede the point that heritability extrapolations may be highly sensitive to small environmental perturbations, but still assert that the less a new population differs from the original one, the more reasonable it would be to expect similar heritability values" (Sesardic, 1993, p. 404).

The problem with Sesardic's argument concerns the limits of our knowledge about the environment. As Oftedal (2005) says, it is very hard, if not impossible, to understand whether complex environments influencing complex traits are actually similar to each other:

"Environments influencing for instance IQ could change substantially from family to family even within the same socioeconomic group. Seemingly small differences in environments could turn out to have large effects on the heritability of phenotypic traits. The great difficulties in controlling for environments in complex behavioral traits also make it hard to tell whether the environmental range one is sampling from is small or large" (Oftedal, 2005, p. 706).

Therefore, by claiming that it is only the prevalent environment that is interesting, Sesardic underestimates the complexity of environments and the unpredictability of the effects that seemingly small environmental changes could have. The case would be different if an additive relation between genotypes and environments was established for a trait.[86] However, Sesardic cannot provide any empirical evidence for supporting the idea that additivity is common in complex behavioral traits:

"He rather shifts the burden of proof by citing several authors stating that it is difficult to find evidence for nonadditivity in the behavioral domain. […] However, several of the authors he refers to in order to back up his point do not argue beyond that interactions are ignored due to the traditional and problematic assumption that interactions are ignorable in heritability analysis […] Sesardic also cites several people that have been much criticized for ignoring the possibility of statistical interactions" (Oftedal, 2005, p. 706).

[86] As I mentioned above, in cases of additivity, heritability estimates indicate a general relationship between genotypes and traits which can be extrapolated to other environments.

According to Oftedal, Sesardic is nonetheless right in claiming that, although being a local analysis, heritability analysis can trace genetic causation.

"Canceling out common causes and only take into account those causes that make a difference on a certain background is a common and very important method of singling out causes in science. […] Thus, even if it is difficult to obtain high quality conditions for empirical research on heritability, it is not in principle impossible to obtain causal information from heritability analysis" (Oftedal, 2005, p. 708).

However, Oftedal herself acknowledges that whether additivity is common in nature is an empirical question on which there are few good results, but at least empirical research indicates that several simple traits in plants and animals present a non-additive genetic structure. In other words, one might notice that the NoR analysis does not provide a detailed insight into molecular mechanisms: it is just a general overview about the genes-environment interplay. The general lesson is nonetheless that in a number of species, the NoR method makes a nonsense of the additivity assumption underlying ANOVA.[87]

The ubiquity of interactions in nature is accounted for by several empirical studies. This is the reason why Wahlsten blames behavioral genetics for being a naïve enterprise:

"There are so many instances where the response of an organism to a change in environment depends on its genotype or where the consequences of a genetic defect depend strongly upon the environment, that genuine additivity of the two factors is very likely the rare exception. The truth of this assertion is widely recognized (Bateson, 1987; Cairns et al., 1990; Fentress, 1981; Gollin, 1985; Gottlieb, 1991; Oyama, 1985), but not in human behavior genetics. Abundant evidence at the molecular level now allows little room for doubt that circumstances can determine when and where a gene acts to influence the course of development and neural activity (Rusak et al., 1990; Schoups & Black, 1991)" (Wahlsten, 1994, p. 5).

Genotype-environment correlation represents another important aspect which I believe might shed light on some oddities of heritability research. Indeed, the G-E covariance involves situations which encompass complex causal stories. These cases can yield a high heritability for traits that, in an intuitive sense, seem to not be related to genetic variation (see Block, 1995; Gibbard, 2001; Jencks, 1980; Sober, 2001). The central point is that, as Lynch and Bourrat explain (2017), in the heritability equations it is assumed that G and E are independent of each other. But if they are not independent, variation in genotype may be associated with variation in environment.

Plomin et al. (2013, p. 107) understand rG×E as a matter of genetic determination of environmental aspects: we create our own environment, in terms of experience, in part

[87] Unfortunately, studying reaction norms in humans is not an option, since the same genotype cannot be tested in a variety of environments. For instance, the study of MZ twins reared apart cannot solve this issue because a couple of individuals is not enough to obtain a norm of reactions: several phenotypic values are needed for the analysis of reaction norms—hence, clonal populations are the best candidates.

for genetic reasons.[88] Therefore, geneticists reduce the problem of rG×E to the problem of *explaining nurture by nature*: accordingly, something that seems to be an environmental influence, has in fact a genetic origin. However, by looking at some examples of correlations between genes and environment, one might notice that the situation is more complex than this.

Block's discussion about indirect heritability may be helpful to explain how confounding may be the study of genetic influence via heritability research in the presence of rG×E.

> "Consider a culture in which red-haired children are beaten over the head regularly, but all other children are treated well. This effect will increase the heritability of IQ because red-haired identical twins will tend to resemble one another in IQ (because they will both have low IQs) no matter what the social class of the family in which they are raised. The effect of a red-hair gene on red hair is a "direct" genetic effect because the gene affects the color via an internal biochemical process. By contrast, a gene affects a characteristic *indirectly* by producing a direct effect which interacts with the environment so as to affect the characteristic" (Block, 1995, pp. 116-117).

As Block highlights, in this hypothetical example the red-hair genes affect IQ indirectly. The problem with behavioral traits is that no one has a full understanding of how to separate out direct from indirect genetic effects because no one has much of an idea how genes and environment affect IQ. For this reason, we do not know whether or to what extent the heritability of IQ is due to direct or indirect genetic effects. Heritability analysis counts differences in traits as caused by genetic differences. But what are these genetic differences "for"? In Block's example, the heritability methodology focuses on the difference between the red-hair genes and genes for other hair colors, not on IQ differences.[89] As the author suggests (1995, p. 120), there may be a large component of heritability due to indirect genetic effects, including (but not limited to) gene-environment correlation, that is outside the boundaries of what quantitative methods can measure.

One might wonder whether such counterfactual scenarios may help one in understanding human behavior and whether they are good reasons to account for a non-additive view in an empirical sense. The fact is that studying the genes-environment interplay in human species is far from simple. However, an important index of the genes-environment interactions is uncovered by epigenetics research. As Champagne and Mashoodh (2009) explain, the control of gene expression is ultimately determined by how accessible DNA is to factors within the cell that are involved in transcription. Influences that determine the expression of DNA without altering the sequence of DNA are referred to as 'epigenetic'. They may include a great variety of factors, ranging from transcribed elements

[88] This allows them to make sense of the high heritability estimated for life events over which we have some control, such as problems in relationships and financial disruptions. This is also consistent with the fact that heritability increases with aging.

[89] See Chapter 5 for a distinction between direct and indirect genetic influence.

(RNA molecules) to environmental aspects that can bring about changes to the functioning of the whole cell. For instance,

"a recent breakthrough in our understanding of gene–environment interplay comes from studies exploring the epigenetic processes that are altered by an individual's experiences during development. Based primarily on studies in rodents, these paradigms address the question raised by G × E research: ''What are environments doing to genes to alter their impact?'' […] Studies of monozygotic (MZ) twins also provide important insights into epigenetic effects in humans. Comparison of the gene expression of 3-year-old and 50-year-old MZ twins indicates a higher level of discordance in patterns of gene expression among older twins that is associated with increasing differences in DNA methylation in older compared to younger twins (Fraga et al., 2005)" (Champagne and Mashoodh, 2009, pp. 128-129).

Can heritability analysis account for the genes-environment interplay? No matter what is the answer, mature behavioral genetics cannot ignore the problem. As things stay at present, it seems reasonable to think that the genetic expression is environment-dependent—not only for physical traits like stature, but also (and perhaps especially) for behaviors. Therefore, it does not make sense to assess the role of inheritance regardless the interaction between hereditary factors and environment. Additivity really seems to be an unjustified ontological assumption.

3. The Limits of Population Analyses (Once Again)

All the topics discussed above require dealing with development. Indeed, quantitative genetics, heritability, and association studies seem to be unable to detect and appreciate those aspects which go beyond the boundaries of the strict statistical and simplified model originally elaborated by Fisher.

One of the oddities discussed in §1.3 about heritability is instructive in this respect. In that scenario, a highly heritable trait may still be strongly sensitive to environmental circumstances. For instance, human stature is a highly heritable trait (heritability around 80%). This would make it natural to think that the environment has a little influence on height. However, this conclusion would be wrong. Indeed: 1) high heritability does not imply that the trait is not influenced by the environment so much; 2) it does not imply the ineffectiveness of some sort of intervention strategy (e.g., for pathologies); 3) environmental interventions are able to modify the genotype-environment balance and, in turn, the heritability within a population.

Does it make sense to ask whether genes are causally prior to environment—or *vice versa*? The so-called "developmentalists" would say no. In his book *Behaving* (2016), Kenneth Schaffner explains that behavioral genetics is faced with the so-called 'developmentalist challenge' (pp. 70-72). The author detects *seven deadly sins* about causation and *four major mistakes* about nature and nurture that classical approaches to the study of behavior have made (see Bronfenbrenner & Ceci, 1994; Griffiths & Gray, 1994;

Gottlieb, 1995; Gray, 1992; Johnston, 1988; Layzer, 1974; Lehrman, 1953, 1970; Lewontin, 1982; Oyama, 1985; Stent, 1981; Wahlsten, 1990).

In sum, eleven bones of contention are considered by Schaffner. For the present moment, I shall focus on those about the nature-nurture problem:

1) Nature and nurture are inseparable causes of development;
2) Genes are not the principal actors in traits' production; rather, they are part of complex systems in which top-down causation can influences genes (e.g., from cytoplasm, hormones, external sensory stimulation, external environment);
3) Genes do not produce clear and specific phenotypes: they react to environmental influences (see norms of reactions);
4) Behaviors are not divided neatly into innate and learned, as Lorenz originally proposed in 1965; this approach should be replaced by an interactionist, epigenetic, ecological, or 'life-cycle' approach;
5) Heredity and environmental effects cannot be disentangled into specific percentages;
6) Analysis of variance (ANOVA) cannot say anything about the causes of individual development;
7) Heritability is a useless and misleading concept.[90]

Lately, Schaffner (pp. 95-98) condenses the developmentalists' framework as being about five theoretical principles: 1) parity; 2) neopreformationism; 3) contextualism; 4) indivisibility; 5) unpredictability. The author, then, compares these stances with developmental research about the worm *Caenorhabditis elegans*. For the present discussion, the interesting aspects are parity, contextualism, and indivisibility.

According to the parity thesis, genes are not "more special" than other developmental elements: DNA, other molecules, and environmental stimuli, are similarly causally relevant for development. This idea has turned out to be correct, even though Schaffner maintains the priority of genes from an epistemic and heuristic perspective because of its relative simplicity: being a linear molecule, DNA is conceptually simple and relatively suitable for being experimentally manipulated.

According to contextualism, different environments and developmental stories lead to different behaviors (learning, plasticity, and environmental influences on gene expression). The *C. elegans* community accepts this point.

According to the indivisibility thesis, causal effects in development cannot be disentangled from each other. This idea is tackled by Schaffner: the causal schema between

[90] The other four mistakes regard the relationship between genotype and phenotype: 1) there is no linearity between genes and phenotypes; 2) Genes do not "contain" biological information; 3) DNA sequences have no fixed meaning; 4) Genes do not make behaviors and neural structures in any direct way. I focus on these four aspects in Chapter 5.

genotype and phenotype is very complex but not inestricable in principle. By means of specific molecular analyses, biologists can, in fact, understand the causal contribution of the parts. As far as I can see, Schaffner argument against indivisibility misses the mark: developmentalists do not say that causes cannot be separated from each other. What they say, rather, is that: a) population methods, e.g., heritability analysis, are not suited for this aim; and b) if the causal contribution of the genome depends on the environment and other biological aspects, then genetic effects have no meaning if taken in isolation.

Take the case of phenylketonuria (PKU).[91] This pathological condition is often regarded as strongly genetic, and not for no reason. But if we ask what is the cause of the related symptoms (e.g., cognitive disability), we must conclude that both genetic and environmental aspects contribute to causing them. Both an allelic variation and an environmental contribution are necessary causes of the metabolic trouble which is responsible for the intellectual impairment. According to Kempthorne (1978, p. 7), we have a clear case of joint causality by two "forces". Can we assess the weight of genes for symptoms? Can we partition the causality in any meaningful way? The answer cannot be positive.[92]

The problem with behavioral genetics seems to be, again, its inability to account for development. Individual organisms are the product of a continuous dialogue between genes and environment. The environment (e.g., nutrition and exposition to light in plants) plays a major role in individuals' development. Looking at individuals, one cannot disentangle the magnitude of genetic influence on a trait's development, because there is no genetic causality to be separated by other causal influences: development processes entail the orchestral interaction of countless multilevel causal pathways over time.

To summarize, one could say that partitioning variance is not the same as partitioning causality. As Wahlsten explains (1994), heritability analysis has persuaded many academics that the importance of genes for behavior can be understood without knowing anything about the cells or physiology which connect genes and behaviors. Partitioning variance with ANOVA represents a second nature to psychologists. This may help to explain why heritability has found a receptive audience in psychological research rather than in developmental biology where interactions are ubiquitous.

This would be fine if scholars were aware of this as they declare. The problem is that the weaknesses of statistical methods match harmfully with the geneticists' aims: that is, saying something interesting about the genotype-phenotype relationship, about the genes people have, and about intervention strategies to improve cognitive abilities or to prevent

[91] The PKU is a metabolic disorder for which the organism cannot metabolize phenylalanine properly. This amino acid is located in many types of food. Phenylalanine is stockpiled in PKU carriers' blood and brain, leading to neural developmental issues and cognitive disability. See Chapter 5 for more details.

[92] Sesardic himself recognizes that genes-environment interactions are necessary to make sense of PKU, albeit it is evidently subject to strong genetic causation: "after the discovery that its pathogenic manifestations can be avoided by taking an appropriate diet, it is now better regarded as the result of gene-environment interaction" (Sesardic, 1993, p. 405).

diseases. Philosophically speaking, behavioral geneticists would like to understand underlying mechanisms, how things work.[93]

Heritability has been clearly pushed beyond the domain of applicability of quantitative methods and beyond the domain of any rational inquiry. As Turkheimer summarizes,

"people conducted twin studies of [doubtful] candidates, and to a troubling degree these all came out to be heritable as well. How much television children watch is heritable (Plomin, Corley, DeFries & Fulker, 1990). Political attitudes are heritable (Alford, Funk, & Hibbing, 2005). Divorce is heritable (McGue & Lykken, 1992). There has been no end to it, especially as twin studies have expanded into the yet-to-be exhausted areas of economics (risk preference: Zhong et al., 2009) and political science (voting: Fowler, Baker, & Dawes, 2008)" (Turkheimer, 2011, p. 230).

The universality of heritability muddied the relationship between the statistical and causal aspects of Fisher's model. It is unlikely that a genetic etiology of divorce is about to emerge, notwithstanding its heritability. As Turkheimer says, having failed to learn the lesson from a century of twin studies, many still speculate about undiscovered biological etiologies of voting behavior. However, Turkheimer does not go for most rational conclusion that heritability analysis is irreparably flawed. He just clarifies that high heritability does not imply determinism, reductionism and the reference to "the gene for X":

"the possibility that so troubles psychopathologists, that an outcome is simply an event that becomes more likely with the accumulation of the tiny effects of a very large number of undifferentiated genes, makes perfect sense for divorce, or at least a lot more sense than the possibility of discovering meaningful 'divorce genes' with attendant 'divorce pathways' leading to 'divorce circuits'. The universality of heritability teaches us that heritability is not incompatible with what we have always thought of as the psychological, and conversely that there is no reason to infer from heritability that reduction of a complex characteristic to genetic or neurological structures has become more likely" (Turkheimer, 2011, p. 230).[94]

Is heritability research, after all, worth pursuing? The behavior geneticists McGuire and Hirsch (1977) stated that, while heritability is "of little interest to geneticists," it has been used and misused in psychology in the belief that it is their nature-nurture ratio or index to the causes of a trait. However, it does not describe the average influence of heredity in determining the level of trait expression in a population. Similarly, Crusio (1990) wrote that heritability studies had been used in the past to convince psychologists and ethologists of the causal role of heredity in psychological trait differences, while it is by now clear that this approach is basically sterile and that these efforts should be abandoned.

[93] A statistical approach is not a mechanistic approach. Even molecular methods such as GWAS do not look for mechanisms: they look for correlations between variables. In Chapter 5, however, I notice that mechanistic explanations in biology are far from being uncontroversial.
[94] See also Plomin et al., 2013, p. 94.

Conclusion: What About Genetic Causation?

In his paper, Lewontin (1974, p. 410) concludes that analysis of variance is useless for the search for causes and indeed it has no use at all. He suggests that we should stop the endless search for better methods of estimating useless quantities. This suggestion has been very influent in genetics literature. Kempthorne (1990) and Wahlsten (1994), indeed, agree on the fact that most of the literature on heritability in species that cannot be experimentally manipulated should be ignored. Behavior geneticists Rutter et al. (1993), agreed with the environmentalists' criticism that knowledge of the level of heritability of any given trait is of very little interest with respect to policy and practice because quite a high heritability does not rule out effective environmental intervention.

Nevertheless, some authors tried to defend heritability as a valuable methodology in understanding the role of genes in complex phenotypes. Therefore, behavioral geneticists still massively employ the concept of heritability for several purposes. After his (defective) arguments pro-heritability, Sesardic (1993, p. 416) concludes by saying that the use of ANOVA in behavioral genetics (and elsewhere) requires caution. Applying the technique uncritically can lead to inferences that distort the picture of actual causal connections. But, Sesardic says, pointing to ANOVA's limitations cannot support the strong claim that we should abandon the goal of partitioning variance among mutually exclusive causes and of calculating heritability coefficients, or that the ANOVA approach is somehow irremediably and intrinsically defective as an instrument in the search for causes. The struggle really seems to never end.

In any case, Rutter et al. (1993) hold that it is a serious mistake to suppose that behavioral genetics is mainly involved with quantifying heritability. According to them, there is indeed very little interest in calculating the precise level of heritability as such. A similar concern is expressed by Wahlsten (2010), who has conducted a bibliographic research that points out how heritability studies are of little interest in genetics research: less than 1% of research employs the term 'heritability'. This does not mean, however, that heritability has no role in genetics research. Rather, this tells us that several other methodologies (e.g., inbreeding and knock-out) are employed where the focus is not on human complex traits. Just looking at textbooks (e.g., Plomin et al., 2013), it is straightforward that heritability has been the cornerstone of human behavioral genetics.

My brief review of the heritability debate may suggest the necessity of other types of research programs to explore the link between genotype and phenotype. Thankfully, quantitative genetics is not the only available framework in biology. At most, it seems to be the most attractive one in the study of complex behavioral traits.

What is the very reason why so many authors believe that heritability is not worth pursuing? The answer lies in the problem of genetic causation. In the next chapter, I deal with this topic which seems to play, surprisingly, a secondary role in behavior genetics. I

shall show how a different understanding of behavioral phenotypes, not involving mere quantitative analyses, is possible and necessary.

Chapter 5.
Genetic Causation: Guidelines for Intelligence Research

In the previous chapters, I retraced the history of the statistical methodologies employed by behavioral genetics and highlighted their theoretical issues. Criticisms have long been raised against statistical methods such as genome-wide association and heritability studies, often accused to be unsuited for clarifying the relationship between genes and behaviors. The aim of this chapter is to show in more details how focusing on statistical methodologies has irreparably worked against the very purpose of behavior genetics, that is, explaining the causal relationship between genes, cognitive phenotypes, and behaviors. However, setting aside the statistical approaches in genetics does not rule out the possibility of explaining the biological, causal background of human behaviors.

The problem of genetic causation is often related to the Biological Information Theory and to the conditions for which a trait can be said to be "genetic". In this chapter, I look at these problems and I argue that, at the very least, genetics explanations of complex behaviors may be neither epistemologically useful nor ontologically meaningful.

I am aware of how complex is the problem of genetic causation: indeed, both philosophical and empirical research have not yet achieved a full understanding of how genotypes bring about complex phenotypes. Therefore, in this chapter, I compare some available positions and propose some guidelines to proceed in the analysis of complex behavioral traits, with a special focus on human intelligence.

I also argue that the two concepts of quantitative (complex) and qualitative (simple) traits should be ruled out by genetics' vocabulary. Not only because, as I showed in Chapter 2, these two terms hide the reference to quite different aspects (e.g., continuous/discontinuous population variation, oligogenic/polygenic inheritance, monogenic/polygenic traits), but also because the employment of such terms misleads about genetic causation.

I also provide some reasons for which intelligence should not be understood as a phenotypic trait in the narrow sense. IQ represents an individual's intellectual level assessed by an IQ test, in a population for which the test has been standardized. Conversely, intelligence, whenever understood as a biological characteristic (e.g., the *g* factor), should be conceived as a set of distinct cognitive processes. Therefore, it is misleading to seek the genes "for" the IQ or "for" intelligence, because they represent general quantifications over different aspects of human cognition.

1. Genes and Behaviors: Open Problems

Worst-case scenario, heritability analysis is not capable to detect the presence of genetic influence on a trait and is not well-suited for assessing the magnitude of genetic variability on phenotype variability.[95] Any molecular method which is bases on heritability research, such as GWAS, is doomed to fail. Conversely, best case scenario, heritability can detect something concerning genetic variance and phenotypic variance. No one is exactly sure about what the association between the two stands for.

As many scholars point out (e.g., Chabris et al., 2013; Lee & Chow, 2014; Ramus, 2006), heritability is a macroscopic parameter of a genetic system: beyond establishing that genes matter, it says little about the genetic architecture of a trait, i.e., how many alleles affect the trait or how they interact. Molecular statistical methods, like GWAS, try to address these questions, but recent failures (i.e., low replicability of results and the missing heritability problem) suggest that the statistical approach fails to hit the mark.

I would say that this is not enough to completely dismiss genetics research in the behavioral domain: no matter how flawed statistical analyses are, looking for the genetic bases of phenotypes is still a valuable research topic. If heritability and GWAS are not the right way to achieve a better understanding in this respect, it does not follow that behavioral traits are not influenced by genes, or that individual differences could not be related to genetic differences. Briefly, the problem of genetic causation on behaviors remains (see Ramus, 2006, for similar remarks).

In the next paragraphs, I ask what requirements successful behavioral genetics should satisfy to account for such an important topic. I suggest that what lacks in behavioral genetics is, ironically enough, biology.

1.1. Becoming truly biological: the focus on causation

As the reader may have guessed, I agree with those who think that quantitative genetics is irremediably afflicted by insurmountable and constitutive fragilities. In Chapter 2, I argued that the quantitative-additive model (QuAdM) was not originally committed to ontological aspects.[96] Any contemporary model which adopts the QuAdM for explaining developmental and molecular phenomena cannot resist a careful philosophical and biological scrutiny. In Chapter 3 and 4, I evaluated statistical methods as unsuited to

[95] I use the term 'variability' instead of 'variance' because the two concepts are quite different from each other: while variance is a statistical concept, variability regards the biological realm. The distinction between the two has been overlooked by quantitative geneticists, whose assumed that variance is meaningful also as a biological concept.

[96] As far as I can see, the QuAd model represents an abstraction. However, it is possible that the architects of the QuAdM were in fact driven by ontological assumptions. This topic will be further analyzed in a paper written with Dr. Flavia Fabris.

achieve their goal, that is, explaining the relationship between genes, cognitive processes, and behaviors. That said, it is not my aim to reject behavioral genetics as a general scientific enterprise. Rather, I will submit that major issues followed by its marriage with statistics.

It is often tacitly assumed that quantitative genetics is the best we have got: being a synthesis of different genetics traditions, there is nothing outside it (see e.g., Plomin et al., 2013). However, contemporary behavioral genetics and quantitative genetics represent, at most, the natural evolution of biometrics. What is there on the other side? I woild say *biology*. This conclusion might strike one, but there are good reasons to think that biology has never really been at the core of behavioral genetics.

So far, I looked at the history of genetics as broken down into two separate sides: biometrics and Mendelism. This is, roughly speaking, the distinction perpetuated by behavioral and quantitative geneticists. Conversely, in *Genetics and Philosophy: An Introduction* (2013) Paul Griffiths and Karola Stotz focused their attention on the differences between Mendelian and molecular genetics. The authors highlight that, on one side, Mendelism and molecular genetics developed two different gene concepts; but, on the other side, they have had similar aims. In a sense, the early molecular genetics was supposed to be the natural prosecution of classical genetics, aimed at identifying the material bases of the Mendelian gene, which was previously just a difference-maker theoretical concept (see Chapter 2).

This led, at first, to the chromosomal theory of heredity and to the discovery that genes live inside the DNA. Crick's dogma and the de-coding of the genetic code followed later on. The relationship between genes and proteins, in that early context, was supposed to be linear. But, at some point, molecular geneticists turned their attention to eukaryotic cells and realized that things were much more complex than initially thought. This led molecular geneticists to take into account gene regulatory mechanisms, metabolic pathways, and their interactions, that is, the whole systems to which DNA, RNA, and proteins belong—involving the other levels of the phenotype and the environment. This tied molecular research to developmental biology.

We can now realize that two different traditions have been involved in the study of phenotypes. The watershed is, at the very least, the *focus on causation*. On the one side, quantitative genetics—including both population genetics and human behavioral genetics—represents a unified framework of the early Mendelism and biometrics, under the aegis of statistical analyses.[97] On the other side, molecular biology represents the bearer

[97] The advocates of this approach (or at least some of them) would say that their approach is in fact engaged in the analysis of causes, e.g., GWAS are aimed at identifying the genes for behavioral traits. However, even though they are frequently described as "molecular", they seek statistical associations rather than mechanistic explanations. Kempthorne has convincingly argued that statistical analyses are not comparable to any proper analysis of causes, where he defined the latter as a difference-making intervention over variables. According to him, "much of the IQ controversy is based on a total semantic error; the data can only give us "due to" in the statistical (and incorrect) sense, and a single worker having satisfied himself on "due

of the research about causality in genetics, including not only molecular genetics but also its links with systems and developmental biology research programs.[98]

To address the problem of genetic causation, behavioral genetics should become *truly biological*. This change could presumably happen by means of two methodological and theoretical shifts: a) from statistical analyses to molecular biology; b) from genetics to development. Let us see them one by one.

Quantitative genetics makes use of statistics for several reasons which sounds nowadays harmful or, at best, obsolete: the statistical approach is a remnant of biometrics, where biological explanations were unachievable. Of course, statistical validity (i.e., generalizability) is a necessary requirement for a biological model, but statistics cannot be an expedient to avoid ontological questions, especially if genetics is aimed at understanding facts about nature (e.g., about organisms and causality). Statistical analyses are not committed to biology, and hence their aim is not to find plausible biological explanations. We might condense the problem by saying that quantitative genetics does not seek causal or mechanistic explanations (see Tabery & Griffiths, 2010).

Concerning causality, statistical analyses are undoubtedly quite distant from causal explanations. As Pearl says (2009, p. 97), causal questions require some knowledge of the data-generating process; they cannot be computed from the data alone. Quantitative genetics employs association concepts (e.g., correlation, regression, dependence, likelihood, ratio), while causal explanations entail quite different terms (e.g., randomization, influence, effect, confounding, 'holding constant', spurious correlation, intervention, explanation).

> "The slogan 'correlation does not imply causation' can be translated into a useful principle: one cannot substantiate causal claims from associations alone, even at the population level—behind every causal conclusion there must lie some causal assumption that is not testable in observational studies. […] Associational assumptions, even untested, are testable in principle, given sufficiently large sample and sufficiently fine measurements. Causal assumptions, in contrast, cannot be verified even in principle, unless one resorts to experimental control" (Pearl, 2009, pp. 99-101).

About mechanistic explanation, instead, several scholars hold that a biological explanation requires an investigation into the underlying mechanisms of a given phenomenon, the parts composing the mechanisms, and the relationship among them (see Bechtel & Abrahamsen, 2005; Craver & Tabery, 2015; Glennan, 2002; Machamer et al., 2000).[99]

---

to" in the statistical sense, in the same writing and almost in the same breath uses "due to" in the intervention sense" (Kempthorne, 1978, p. 6).

[98] Developmental biology has a quite independent history but, in a sense, it is nowadays strictly related to (or at least consistent with) molecular biology and systems biology.

[99] The problem of what a good explanation in biology looks like is, however, controversial in philosophy of science; it is not my aim to argue that mechanistic explanations represent the best explanatory strategy for biology. Indeed, mechanistic philosophy has been tackled for both its epistemological and ontological fragilities (e.g., Baetu, 2015a, 2015b, 2016; Campbell, 2006; Dupré, 2013; Nicholson, 2012; Dupré & Nicholson, *forthcoming*; Russo & Williamson, 2007; Weiskopf, 2016). This is the reason why, in this chapter,

By the way, even if behavioral genetics would achieve a satisfying level of understanding about the molecular bases of behaviors, it would still not be enough: it is very likely that a complete knowledge of the genotype-phenotype relationship requires far more than an inquiry on molecular mechanisms, no matter how detailed. Behaviors have not only genetic determinants; rather, they are part of huge networks of multileveled interactions.

This consideration leads us to the second required shift for behavior genetics becoming truly biological, that is, from genetics to development. Behavioral genetics seems to be so "junky" whenever it does not appreciate, to use a figure of speech, the complexity of life. Even though behavioral geneticists have lately come to appreciate the role of the environment (especially for individual differences), most of them still adopt a form of "outdated biology". As I mentioned in Chapter 4, Schaffner (2016) has summarized decades of criticisms against quantitative behavioral genetics as "the developmental challenge". This story entails numerous protagonists whose, over the past decades, have shown how behavioral genetics misleads about biology and, arguably, about genetics itself. It is straightforward to many scholars that only a developmental framework would be able to offset such a fragility.

As it is now clear, the story began at least with the foundation of the Modern Synthesis: after the Mendel wars, a statistical theory of continuous phenotype (biometrics) has been superimposed on a theory of some underlying discontinuous elements, the Mendelians factors (see Chapter 2). Several issues in contemporary quantitative and behavioral genetics seem to derive from such a shift, which bears an exaggerated focus on genotypes and the exclusion of the other levels of organization from the equation—including the organisms themselves. Population geneticists set apart development, phenotypes, and organisms in favor of genes (later identified with stretches of DNA). And the rest is history.[100]

As Griffiths and Stotz have pointed out (2013, pp. 102-107), biological explanations entail two steps: reduction and integration. The former concerns molecular research (bottom-up) and it is, everything aside, a necessary step—even though sometimes slippery

---

I focus on the general problem of causality instead of more specific accounts such as the mechanistic framework. However, explaining what are the mechanisms involved in the genes-behaviors relationship would be at least a result. Unfortunately, even molecular-reductionist explanations represent a distant target to behavioral genetics.

[100] Ever since the backbone of the synthetic quantitative view has been established, Conrad Waddington strongly criticized the focus on a-temporal and a-contextual atomic elements as genes came to be conceived in population genetics (see Peterson, 2011; Waddington, 1968). He advanced what is probably the more articulated view antagonist to the Synthesis by highlighting the importance of developmental trajectories with his epigenetics. Dobzhansky brought unorthodox elements in the Synthesis, among which the study of norms of reaction. Several authors managed to frame the issues of quantitative genetics from a psychological and philosophical viewpoint, contrasting reductionist and statistical analysis in favor of development—tackling in this way traditional behavioral genetics as well. Gottlieb (2003) called behavioral genetics to become "truly developmental". Several criticisms by philosophers and scientists followed similar lines (see Chapter 4).

(ontological reductionism is right around the corner and reductionist methods do not always return a genuine picture of the reality). The latter concerns systems biology, where knowledge about things are integrated in accordance with their "very nature" (top-down).

> "There is a profound difference between the molecular function of a molecule, which it owes to its molecular structure, and the realized cellular function which it owes to its context and its interaction with other entities […]. This distinction poses something of a challenge to the very idea of specificity. It suggests that specificity may not be an intrinsic property of a sequence or a structure but a contextual property" (Griffiths & Stotz, 2013, p. 106).

While it is true that quantitative genetics does not seek for causal explanations, this does not have to necessarily apply to behavioral genetics. Nowadays, this discipline should still consider embracing biology and to be not merely genetically oriented.

Two questions are central in the following paragraphs: 1) is it epistemologically or practically useful to focus on individual differences? and 2) are genes really important for explaining behaviors in a biological sense? I first submit that behavioral genetics should not only focus on individual differences, but rather in how behaviors develop in the normal range of value (§1.2). Then I submit that, while it is legitimate to seek for genes involved in behavioral phenotypes, this can be understood in different manners which should be clearly distinguished one from another (§1.3 and §1.4). The importance of genes in explaining human behaviors may be drastically smaller than generally thought. This intuition will play a central role in the problem of defining intelligence.

1.2. The problem with individual differences

With the distinctions between quantitative and molecular genetics delineated above, two possible strategies to analyze the genetic bases of behavioral traits might be adumbrated. The first one has mainly attracted the attention of behavioral geneticists and concerns the relationship between phenotypic differences and genetic differences. This route involves the study of populations rather than individual development (e.g., heritability analysis and GWAS). The second one, instead, concerns understanding genetic causation as a matter of developmental analysis: is a given behavior influenced, caused, determined, by genes? Unfortunately, the two problems have been hopelessly conflated with each other and linking genetic causation to population analyses (see Chapter 4). Even though understanding the genetic influence on behaviors seems to be the very problem for geneticists, behavioral genetics has almost never been involved in it.

It is important to separate these two different appraisals of the putative framework for studying genetic causation. Northcott (2006) has persuasively stressed that, to confront the impossibility to separate genetic and environmental causes, biologists developed a misleading understanding of causal efficacy (of which ANOVA is the main bearer),

which applies solely to a whole population. Some behavioral geneticists are aware of this and do strongly advocate for individual differences as the distinctive mark of their discipline. For instance,

> "Scarr, following in the Fisherian tradition of focusing on the relative contributions of various sources of variation, argued that developmental behavioral genetics should seek the causes of *phenotypic variation*, rather than the causes of phenotypes, and ask *how much* phenotypes depend on certain causes, rather than how they depend on them (see also Plomin, 1983, p. 254; Scarr, 1992) […]. These methodological stipulations have been used for over 50 years to defend traditional behavioral genetics against the accusations that (1) it does not yield causal explanations and (2) it cannot explain phenomena at the individual, as opposed to the population, level" (Tabery & Griffiths, 2010, p. 44; emphasis added).

Are individual differences relevant to achieve a better understanding of genetic causation? Several authors have answered negatively by pointing out that the attention should be directed elsewhere (see Noble, 2008; Tabery & Griffiths, 2010). When we ask what is the genetic basis of cognitive processes and behaviors, we are not only interested in the genetic bases of individual differences. We would like to know, rather, whether an individual level of intelligence *is caused by genes*, how genes influence human brains, and which genes make the humans intelligent the way they are and how.

One might notice that the focus on individual differences in behavior genetics derives from methodological necessities: there is no genetic analysis without individual differences: How could have Mendel found his well-known pattern of inheritance if not by working with phenotypic variation? How could molecular genetics track down the genetic patterns without knocking-out genes in different animal strains (e.g., in rodents)? How could genome-wide association work without different groups of people (i.e., case-control experimental design)? This is by no means anything outstanding, but simply to recognize that it is hard, if not impossible, to disentangle patterns of inheritance in face of genetic or phenotypic homogeneity. As Noble (2008, p. 3007) notices, the assignment of functions to genes depends on observing differences in phenotype consequent upon changes (mutations, knockouts, etc.) in genotype. This is likely one of the reasons why individual differences have attracted the researchers' attention within genetics analyses.

Other reasons pertain to historical, social, educational, and clinical aspects. As I explained in Chapter 1, the history of quantitative genetics is intertwined with Galton's eugenics. Here the goal was not just understanding the bases of human intelligence, but rather understanding why people are different from each other. Therefore, no matter how behaviors develop, heritability analyses are enough for the prefixed aim of understanding how to improve the quality of the human species.[101] Over time, in educational and clinical contexts, individual differences have been maintained as the central aim of behavioral

[101] In a sense, eugenics was supposed to be the equivalent of animal inbreeding in the human population. This explains why heritability was not related, at the beginning, to the study of the causes of individual differences.

genetics to find intervention strategies capable to bridge the intellectual gap between people. However, it is nowadays clear that it is not possible to do so without a comprehensive understanding of how behaviors develop and of what is the role of genes and environment in this process.

If we ask how intelligence develops in humans, we are not dealing with individual differences, but rather, with individual *resemblances*. Humans are similar in respect of intelligence, even though such trait may widely vary within and between populations. A very important biological process must be at stake in every human being.[102] A first compelling reason for reevaluating the role of individual differences in the quest for genetic causation is explained by Noble:

> "Differences cannot reveal the totality of functions that a gene may be involved in, since they cannot reveal all the effects that are common to the wild and mutated types. We may be looking at the tip of an iceberg. And we may even be looking at the wrong tip since we may be identifying a gene through the pathological effects of just one of its mutations rather than by what it does for which it must have been selected. This must be true of most so-called oncogenes, since causing cancer is unlikely to be a function for which the genes were selected. This is why the Gene Ontology (GO) Consortium (http://geneontology.org/) excludes oncogenesis: 'oncogenesis is not a valid GO term because causing cancer is not the normal function of any gene'. Actually, causing cancer could be a function if the gene concerned has other overwhelming beneficial effects" (Noble, 2008, p. 3007).

A second reason pertains to systemic features that entail a mismatch between genotypic and phenotypic aspects, i.e., robustness and plasticity. Since biological systems are robust and plastic, it may follow that genetic differences are not important as geneticists tend to think for the aim of explaining phenotypic traits. Take phenotypic plasticity. It might be the case that individuals belonging to the same species, albeit phenotypically different from each other, do not differ genetically, in any interesting sense, for species-specific traits like intelligence. Rather, it might be that the members of a species are genetically homogeneous, at least from a functional point of view. In other words, alleles (read: single-nucleotide polymorphisms) might not have an interesting role in phenotypic variation within a population because the population is largely homogeneous on a genetic level for most of its traits.[103]

---

[102] One strategy to analyze the genetic origins of human intelligence would consist in analyzing the intelligent behavior in animal models. Indeed, humans are likely intelligent the way they are because they have specific biological characteristics not shared by, e.g., warms and rodents. Unfortunately, comparative analyses received far less attention than the quantitative study of variation within human populations. However, one may wonder whether one could really understand human intelligence by analyzing animal models like, e.g., rats. Even if something like intelligence does exist in rodents, maze performances seem to be quite distant from the abilities we generally associate to intelligence (for similar concerns, see Garlick, 2002, p. 129). For a review about the evolution of general intelligence, see Burkart et al., 2017.

[103] Thinking about evolution, this seems to be a promising hypothesis for traits which play an adaptive role (see Richardson, 2000, and Darwin's work for similar hypotheses). Likely, genetic variation is not a good thing in evolutionary terms, because it could lower the fitness; hence natural selection might have acted to make most mutations neutral. Variants of large effect size represent a different matter: they affect the phenotype because they can break the normal development in a way against which the biological system cannot

If my hypothesis is true, individual behavioral differences should be framed as *phenotypic variation despite genetic invariance*. This is, of course, just a speculation that should be investigated from an empirical viewpoint; however, it follows as a plausible hypothesis from what we know about biological systems. Noble identify a similar issue in relation to robustness:

"Identifying genes by differences in phenotype correlated with those in genotype is therefore hazardous. Many, probably most, genetic modifications are buffered. Organisms are robust. They have to be to have succeeded in the evolutionary process. Even when the function of the gene is known to be significant, a knockout or mutation may not reveal that significance" (Noble, 2008, p. 3007).

To summarize, intra-specific differences have attracted all the attention because they represent a valuable compass for genetics analyses. Unfortunately, focusing on differences instead of on resemblances may risk harming genetics research rather than being helpful.

As I show in next paragraphs, even philosophers have sometimes been cheated by an excessive focus on individual differences. This is, by the way, what Noble (2008) called "the differential effect problem". Indeed, the focus on individual differences might confound the role of variables in causal models and, in turn, it may hide how behaviors develop regardless individual differences—which is likely the very question on the line. How a zygote can develop in such a way to becoming a multicellular organism that shows high-level cognition? Of course, this cannot be the place to fully address such an important question. My aim in the next paragraph is to point out that the problem of genetic causation on complex phenotypes assumes different appearances if one focuses on individual differences or not.

1.3. Genetic causation in relation to (and regardless) individual differences

What is genetic causation? Is it somehow related to the problem of the genes "for" X? What is a genetic trait? As I shall show, all these problems are related to each other, and understanding human intelligence in a biological sense cannot overlook those questions.

Schaffner (2016) deals with the problem of genetic determinism by comparing several positions within the philosophical debate. Here, Schaffner delineates the definitions included in the works of Gifford (1990), Sarkar (1998), Kitcher (2001), Kendler (2005, 2012) and Waters (2007). As it will be clear soon, Schaffner effort is a very useful benchmark. Gifford (1990, p. 333) defines a genetic trait in such a way: a trait can be said to be

deal with. For such pervasive problems, robustness is never enough. Conversely, normal development could be related to genes that do not vary so much within and between populations of the same species.

"genetic" in a population if genetic factors "make the difference" between those individuals with the trait and the others.[104] Kitcher (2001, p. 348), on its own, holds that we can refer to the "gene for X" if substitutions on a chromosome would lead, in the relevant environments, to a difference in the "X-ishness" of the phenotype. Waters (2007) adopts the Woodward's theory of causation and claims that genes have a privileged ontological status, because they are "actual difference makers".[105] Waters concludes that genetic alleles are the real causes of traits.

These criteria for addressing the causal role of genes, and for defining genetic traits, are misleading for two reasons related to the excessive attention to individual differences I treated above. First, as Schaffner explains, those criteria do not account for traits which do not vary in the population; in other words, they do not include, for instance, the fact that humans have two legs, which is reasonably intended as a phenotypic trait. Second and consequently, they concern Mendelian traits only, or as I shall explain, what I call *differential genetic causation*.

If one looks at the problem of genetic causation by focusing on individual differences, one looks for genes which involve phenotypic differences within populations. Here, we have two analytic frameworks: the difference-making analyses (i.e., Mendelism) and the analysis of variance (i.e., biometrics). Two kinds of genetic causality should be distinguished from each other:

- Differential genetic influence narrow-sense: Phenotypes might vary within population following Mendelian patterns (i.e., qualitative variation) just because a single allele represents the whole genetic population variability. This type of genetic influence concerns experimental isolation of differing genes, where an allele segregates sharply. In medical genetics, the so-called monogenic pathologies entail the fact that a single allele is responsible for the discontinuous phenotypic variation in a population (see Chapter 2).

- Differential genetic influence broad-sense: Complex traits vary continuously within populations. Individual differences are due to different genetic makeups characterizing individuals (the allelic toolkit pivotal in the QuAdM). This type of genetic influence concerns the small influence of several alleles on a trait variance in the population, as ANOVA and GWAS address (see Chapter 3).

[104] Gifford proposes another definition that I analyze in §1.4.

[105] Accordingly, X is the actual difference maker with respect to Y in a population p if and only if: a) X causes Y (in the sense of Woodward's manipulability theory); 2) the value of Y actually varies among individuals in p; 3) the relationship expressed by "X causes Y" is invariant with respect to the variables that actually vary in p; 4) Actual variation in the value of X fully accounts for the actual variation of Y values in p (via the relationship X causes Y).

Most of the literature in quantitative genetics regards, essentially, these two kinds of genetic influence. However, it is not granted that differential genetic influence is worth pursuing neither to achieve a better understanding of the genotype-behavior relationship nor to define what are genetic traits.

Mendelian genes are generally understood as sufficient causes for developing a trait, where "sufficient" is defined as "when the cause is present, the effect must occur" (see e.g., Page et al., 2003). However, as I anticipated in Chapter 2, monogenic pathologies are not subject to the control of single genes: a single gene is capable of causing an abrupt interruption of the normal development of the trait, but any phenotypic trait develops under the influence of several genes (see also §2.1). For clarifying this aspect, let us consider Dawkins' definition of the "gene for X":

"When a geneticist speaks of a gene 'for' red eyes in Drosophila [sh]e is implicitly saying: there is variation in eye color in this population; other things being equal, a fly with this gene is more likely to have red eyes than [is] a fly without this gene' (Dawkins, 1982, p. 21).

As Dupré argues (2013, p. 23), this characterization entails many problems. While a gene "for" X will presumably initiate some causal pathways that lead to the trait X, it may very well also be involved in mechanisms that tend to prevent the appearance of trait X (see also Gifford, 2002). Let us explain. Mendelian traits are generally related to single genes; however, Mendelian genes do not bring about traits; rather, they cause a discontinuous variation in a given population. The trait's development (e.g., the flies' eye color) is nevertheless due to several genes. Pigmentation, for example, is related to a metabolic cascade where every enzyme brings about a specific pigment. The gene "for" a trait in Mendelian genetics ("for" monogenic pathologies and "for" simple traits) is a gene which is actually *"not for"* a trait. It is generally a non-functional allele. Perhaps, in the case of eye color, that gene normally codes for an enzyme related to cells' pigmentation; but, in its mutant form, it does not code for anything, thus causing a phenotypic variation in the population. In sum, Mendelian genes are not, by definition, genes "for", because their expression is compromised (i.e., they do not code for efficient genetic products). In other words, Mendelian genes are not sufficient causes for development: they are, rather, sufficient causes for a discontinuous variation.

As regard as the development of phenotypic traits, I would say that there is nothing *less causally efficient* than a Mendelian gene: when we talk about the genetic influence we seek for genes that are *part of the causal network* by which a trait develops. In the case of Mendelian genes, it is precisely *the absence* of those genes in the causal network that is responsible for the onset of the trait or of the pathology.

So, what if one studies genetic causation regardless individual differences? I would say that one can focus on the proper relationship between genes and behaviors. In accordance with this, I propose two definitions:

- Genetic causation: A gene causes/is "for" a phenotypic trait if the gene plays a role in the causal network that leads a zygote to develop into a multicellular organism which carries that trait.

- Genetic trait: A phenotypic trait can be said to be genetic if one or more genes are involved in its development.

Two remarks: 1) the definition of genetic causation might be revised in several manners, but we can safely say that a gene is causally related to a trait if and only if that gene is involved in that trait's development; 2) the definition I proposed for a genetic trait is practically useless. Nobody would deny that every phenotypic trait is somehow related to genes: DNA is involved, even in a week sense, in the production of any characteristic of the organism. Therefore, asking whether a trait is genetic is not so useful. The question on which I would like to focus concerns, rather, what kind of genetic involvement is required for saying that genes play an important role in a trait's development. Does DNA play *a relevant role* in the causal network that leads a zygote to develop into an intelligent multicellular organism? How can DNA stretches bring about cognitive systems and intelligent behavior?[106]

These queries uncover puzzling issues that have yet not been solved. A good answer should involve not only genomics but also transcriptomics, proteomics, epigenetics, cellular biology, neurobiology…you name it. For instance, such an explanation would explain how genetic products (i.e., RNAs and proteins) influence a particular cell type (i.e., neurons) and how genetic products influence, to a greater or lesser extent, neural networks.[107] The activities of the human brain are the products of a complex interplay between factors at multiple levels; hence, the gap between genes and cognition can only be bridged by a systems biology account of brain development and functioning (see e.g., Fisher, 2006; Grant, 2003; Marcus, 2004).

For the present goal, let us remain in the genetic domain. The Biological Information Theory represents a promising framework to address some issues on this level—indeed, the problem of genetic causation is frequently related to encoding, that is, how genetic information passes through to phenotypes. Allow me to first say that genetic information

[106] It is worth noticing that these questions do not entail the causal role of DNA on behaviors, but rather on the neural structures that make individuals behaving in a certain way. Indeed, for paraphrasing Brenner (1974, p. 72), behavior is the result of a complex set of processes performed by nervous systems; it seems essential to decompose the problem into two: the question of the genetic specification of the nervous systems, and the way nervous systems work to produce behavior.

[107] A complete explanation would also involve how genetic products, neurons, and neural networks, interact with the environment and with each other over time. By the way, psychology and cognitive sciences cannot be excluded by this huge explanatory effort: how could we study a behavioral trait without defining it ? I introduce psychological sciences in my discussion from the next chapter.

is limited to genetic products: genes do not code for phenotypic traits; they code for molecules (see Godfrey-Smith, 1999; Godfrey-Smith & Sterelny, 2016).[108] That is, genes code for things at the molecular level and do not code for a definite part of a test score at the psychological level (see Wahlsten, 1994). This is relevant insofar as there is an important relationship between the problem of genetic causation, the problem of encoding, and questions about the genes "for" phenotypic traits. I shall argue that, by assuming the Biological Information Theory, we should recognize that there is no gene "for" behaviors, and that genetic causation does not play a special role in explaining them.

1.4. Genes "for" X, encoding and genetic influence

As I discussed in Chapter 2, one might be tempted to adopt the QuAd model because, among other reasons, it allows one to avoid simplifying phrasing such as the "gene for X" and the G-P linearity. However, one could agree on the fact that there are genes "for" behaviors *once assumed* that they are quite a lot and they have small effects. In other words, one might want to dismiss the simplifying idea of the gene "for" intelligence and other complex behaviors (e.g., for ruling out G-P linearity) but, at the same time, one could agree on the existence of genes for intelligence in a loose sense, that is, in accordance with the QuAdM.

Schaffner (2016, p. 220), for instance, accepts GWAS findings and buys the idea that behavioral traits are related to thousands of genes of tiny effect-size—then, he likely would not agree on the criticisms on GWAS I summarized in Chapter 3. At the same time, like many scholars, he believes that there are no genes "for" X. This idea is based on X being very complex and the G-P map being not linear. In this respect, Schaffner accepts Kendler's proposal (2005) according to which five criteria for X being a gene for Y: 1) strength of association of X with Y; 2) specificity of relationship of X with Y; 3) non-contingency of the effect of X on Y; 4) causal proximity of X to Y; 5) the degree to which X is the appropriate level of explanation for Y. As Schaffner notices,

> "In general, in behavioral and psychiatric genetics, the strength of association I modest or weak […]; the specificity of relationship is not one-to-one but typically many-to-many. Genes work in complex contexts, and gene-gene and gene-environment interactions are frequent. Furthermore, biochemical chains of influence are long and complex […]. Finally, the results that seem well supported are not at the level of clinical descriptions of such disorders, but are rather phrased in terms of endophenotypes that may eventually be identified with higher-level disorders" (Schaffner, 2016, p. 222).

[108] To be precise, while it is often said that DNA encodes RNA, according to Griffiths and Stotz (2013, p. 46) this is not strictly right: the relationship between DNA and RNA is dictated by chemical complementarity; genetic code is a relationship between RNA and amino acids only. For the sake of brevity, I talk about encoding both for RNA and proteins.

A reasonable conclusion is that there are no genes "for" behaviors. Nevertheless, another consequence might follow, according to which one can maintain the wording "the gene for X" to denote any gene which is somehow related to a given behavioral trait. For instance, Godfrey-Smith (1999) argues that one or more concepts of "gene for X" are tenable, such as the fact that a gene codes for a protein which causes a trait. To make another case, when Plomin and colleagues say that a trait is heritable or genetic, they mean that at least one gene as a measurable effect on that trait (Plomin et al., 2013, p. 373). As far as I understand their positions, these authors would assign the label "genes for intelligence" to thousands of alleles insofar as GWAS attest that thousands of alleles are related to (or have an effect on) intelligence. Schaffner, on his own, would likely say that thousands of genes influence intelligence. I believe that this is not a valuable epistemological principle.

When I say that there are no genes "for" behaviors, I do not just mean that there is no a one-to-one relationship between genes and behaviors (as it is often understood the classical phrasing "the gene for X"). Rather, I meant that there are no genes which bring *specific information* to behaviors: the genes' causal contribution to behaviors is rather very broad and, in a sense, not interesting.

Let me explain. Let us say that genes represent just one part of the causal network that materially realizes a behavior—it seems to be empirically true that phenotypic properties (e.g., neurobiological properties) and the environment are very relevant to behaviors. Being intelligence a complex trait and the IQ a behavioral outcome, it stands to reasons that huge networks of biological causes are involved into them (both in the development of intelligence and in the related behavioral individual differences). This point is well explained by Ratner, who assigns to genes just a necessary (but not sufficient) role for developing behavioral traits:

"A gene generates a broad physiological substratum upon which psychological phenomena can be constructed. […] The construction process itself is not directed by genes. It is organized by cultural and mental processes. Genes are necessary to generate the general capacity for psychological phenomena […] Genes only have codes for physiological matters such as 'association neurons'; genes do not have codes for psychological phenomena such as grammar, love, problem solving, or syllogistic reasoning. Mutant genes may produce deformed physical substrata which cannot support psychological functions […] There is a key difference between being necessary and being a cause. […] Normal biology is a necessary foundation for our cognitive skills; however, it does not cause them. […] One cannot extrapolate from the causal power of a defective element to the causal power of a normal element. The fact that a defective element prevents some act does not imply that a normal element causes the act" (Ratner, 2004, pp. 30-31).

We could then distinguish between two types of genetic influences—this time regardless any concern about individual differences:

- Specific genetic influence: This is the proper *encoding*. It does not act on high-level phenotypic traits, but rather it acts on specific entities located on the molecular level

(e.g., RNA and proteins). This type of genetic influence accounts for the phrasing "the gene for X", where X is a molecular genetic product.

- Non-specific genetic influence: It is the consequence of the fact that genes code for proteins and proteins have, in turn, causal effects on higher-level traits. So, this genetic influence does not directly act upon higher-level phenotypic traits like cognitive phenotypes and behaviors: it is, let us say, indirect. Roughly speaking, genetic information tends to *dilute* once arrived in proteins, then becoming spread along complex interactive causal networks.

Genes serve as blueprints for genetic products. Hence, DNA represents, like other parts of a biological system, an important aspect of the material realization of behaviors. Likely, every gene involved in the "maintenance" of the whole system is somehow related to traits like intelligence, but no one would say that all the genes which are necessary for a normal neurocognitive development are genes "for" intelligence![109]

We could now distinguish between three meanings of the "genes for X": 1) genes "for" molecules (i.e., encoding); 2) genes "for" population variation (that regard differential genetic influence and the so-called monogenic traits)[110]; 3) genes "for" complex traits. The latter meaning is the one I am moslty interested in. As I mentioned above, it is not epistemologically useful to assign the name "gene for intelligence" to a multitude of genetic factors involved in the complex biological network which leads to that normal development of an individual.

At this point, we can finally return to the remaining four mistakes characterizing behavioral genetics, highlighted by developmentalists (see Schaffner, 2016):

1) There is no linearity between genes and phenotypes (e.g., behaviors), that is, the relationship is many-many and incredibly complex;
2) Genes do not contain information. Biological information is, in fact, the product of ontogeny;
3) DNA sequences have no fixed meaning, but rather their causal role is exerted in the context of broader causal networks of non-genetic interpreting molecules;
4) Genes do not make behaviors and neural structures in any direct way: they produce

[109] The open question is whether it is reasonable to say that the genes that cause some characteristic of the neural cells cause also intelligence. In other words, one might say that, if a gene codes for a protein which is causally related to the development of a trait, then the gene *causes* the trait. This implies that it is meaningful to talk about genetic causality on complex traits, even if genetic causation is small and indirect. This counterargument relies on the assumption that causality is transitive. From a metaphysical viewpoint, Hibberd (2014) has convincingly argued against this position. In §3, I shall show that, from both an epistemological and a metaphysical perspective, the distinction between direct and indirect genetic influence may play an important role.

[110] As I shown in §1.3, this usage of the wording "the gene for X" is tricky and likely meaningless, because it attests the absence of genetic causation (what I called the genes "not for" X).

> molecules that affect cell differentiation to yield neurons that become specific types of neurons in specific places with connections with other neurons.

Developmental research about *C. elegans* do account for these principles. The first and the fourth points are attested by the fact that there is a many-to-many relationship both between genes and neurons and between neural networks and behaviors (overlapping circuits). The second point recalls the criticisms against neopreformationism: no DNA sequence "represents" a behavioral trait. The third point concerns contextualism (see Chapter 4, and Schaffner, 2016, pp. 93-97).

For the present discussion, the most important point is the fourth one. Talking about genetic causation on behaviors overlooks what genes really do, that is, producing molecules for which the targets are some cellular feature. As far as I can see, non-specific genetic influence is what characterizes the relationship between genotype, cognitive processes, and behaviors. As Turkheimer says,

> "Complex human behavior emerges out of a hyper-complex developmental network into which individual genes and individual environmental events are inputs. The systematic causal effects of any of those inputs are lost in the developmental complexity of the network. Causal explanations of complex differences among humans are therefore not going to be found in individual genes or environments any more than explanations of plate tectonics can be found in the chemical composition of individual rocks" (Turkheimer, 2011, p. 600).

My argument fits in some aspects of Gifford's and Sarkar's definition of genetic traits. According to Gifford (1990, p. 343), a trait is genetic if it is individuated in such a way that it matches what some genetic factors specifically cause. According to Sarkar (1998, p. 182), instead, a trait is genetic if the immediate products of the alleles at these loci do form part of the biochemical characterization of the trait.[111] As Schaffner (2016, p. 219) underlies, the notion of specificity is here invoked to account for the fact that traits are often defined too broadly (e.g., high cholesterol and intelligence) or too narrowly (e.g., speaking French).

In the second part of this chapter, I interpret both quantitative and qualitative traits in the light of the distinctions I provided above: differential *versus* non-differential, and specific *versus* non-specific, genetic influence. I shall argue that identifying high-level characteristics (e.g., cognitive processes and intelligence) and behaviors (e.g., IQ and schizophrenia) as phenotypic traits in the narrow sense is a practice that deserves far more care

[111] Sarkar's complete definition is more restrictive; a trait is genetic if and only if: a) the trait is under the control of a few loci; b) the trait shows high expressivity in all populations; c) the immediate products of the alleles at these loci form part of the biochemical characterization of the trait. The condition (b) concerns strong genetic determinism (I bypass this problem here). Schaffner criticizes (a) insofar as we now know that traits are related to several alleles. I believe, however, that there are two different ways to think so: the first one, adopted by Schaffner, concerns GWAS findings; the second one, instead, accepts that every trait is polygenic but rejects the quantitative-additive framework. I defend the second position in §2.

than often assumed. Whether an observable feature of an organism (e.g., intelligence measured by IQ tests) is a phenotypic trait depends on how one defines genetic causation, on how important genes are for its development, and on whether genes are important for defining it. The concept of phenotype cannot be so general to include any observable feature of an organism.

2. Rethinking Phenotypes

Is there genetic influence for any observable feature? Is epistemologically fruitful to consider as "genetic" a trait which is influenced by genes indirectly, by a sort of non-specific genetic influence? How should we conceptualize qualitative and quantitative traits in accordance with the previous remarks about genetic causation? These are the questions that I aim to address in the following paragraphs.

My first move is to recover the classical distinction between *traits* and *characters* drawn, for instance, by Spuhler:

> "What we observe is phenotypic variation. After analysis, phenotypic variations may be divided into two sorts: 1) characters, and 2) traits. Here 'characters' mean phenotypic attributes or attribute sets whose variation (for a defined environment) has been demonstrably associated with a defined set of genes. The definition of 'character' presupposes certain specific genetic information. Since there are a limited number of genes in man, there are a limited number of characters. […] 'Trait', as used here, means all phenotypic attributes or attribute sets that are not 'characters.' No specific genetic information […] is presupposed in the definition of traits. The variation of traits is often (but not necessarily) continuous. Traits may be associated with 'factors', that is, with unidentified genes. Statements about factors (as in the sentence 'stature is controlled by multiple factors') presuppose different prior information than statements about genes. By genetic analysis with positive results traits may become characters" (Spuhler, 1954, p. 131).

This distinction reminds us that taking a phenotypic trait is always an abstraction and implies the possibility that an observable characteristic is not influenced by genes. Let us take IQ as an example. It is for sure a *trait*, an observable and measurable feature of an organism. Whether it is a *character* (influenced by genes) is a matter of empirical facts. However, a character may be subject to several types of biological influences and focusing on genetics might be misleading. As I shall suggest, the importance of genetic factors for complex behaviors might be less than generally thought.[112] As I anticipated in Chapter 2, both qualitative traits (e.g., monogenic pathologies) and (at least some) quantitative traits (e.g., IQ) should not be understood as phenotypic traits in a strict sense.

[112] Moreover, if the only way to assess whether IQ is influenced by genes relies on heritability analysis, there might be a serious trouble. However, for the sake of the argument, I concede that any behavioral trait is causally related to some genes, even indirectly.

2.1. Qualitative traits

There are thousands of traits that are thought to be monogenic. They are mostly pathological conditions, among which more than 250, many of them extremely rare, include cognitive disability among their symptoms (see Plomin et al., 2013, p. 183). Who accepts the broad distinction between quantitative and qualitative traits tends to treat pathologies as monogenic traits, influenced by single genes. Conversely, who adopts the strict quantitative model, where common disorders are reduced to quantitative traits, tries to account for them in a purely quantitative framework (see Chapter 2).

My argument is against both these approaches. The broad model fails in acknowledging that single genes are not "for" traits, but rather "not for" traits (see §1.3). This misleading perspective is not assumed with the strict model. However, a major issue is shared by both the two models, that is, they assume that single-gene pathologies (e.g., PKU) and yes/no traits (e.g., schizophrenia) should be understood as phenotypic traits in the narrow sense. I submit that they are not traits, but rather *particular variants* of much complex traits, which I call, instead, *proper traits*. I shall first take phenylketonuria as an example suitable to explain why we should avoid the reference to monogenic disorders as traits. Schizophrenia will be, instead, the central example of a yes/no trait which has recently been framed quantitatively (see Plomin et al., 2009) that should not be conceived as a narrow-sense trait.

PKU is a metabolic disorder for which the organism cannot metabolize phenylalanine properly. This amino acid is located in many types of food. Phenylalanine is stockpiled in PKU carriers' blood and brain, leading to neural developmental issues and cognitive disability. From the inheritance viewpoint, PKU is a monogenic autosomal-recessive condition. Therefore, an individual must carry two recessive alleles to manifest the clinical conditions—those people with only one recessive allele are health carriers.[113]

First, it is misleading to think of clinical pictures like PKU for which a genetic defect is a necessary and sufficient condition to develop it. On the one hand, a genetic defect may not be the sufficient condition to develop a monogenic pathology. In the case of PKU, an environmental contribution is necessary as well. Decades ago geneticists thought that a genetic defect was the only cause of PKU. Subsequently, geneticists understood that by watching a specific diet, poor of phenylalanine, during the early infancy, symptoms can be prevented to occur. Therefore, a normal diet including phenylalanine is a necessary factor to develop the disorder as well. In other words, the genetic defect and the phenylalanine intake are both necessary conditions, but not individually sufficient, to exhibit the disease (see Chapter 4; Kempthorne, 1978; Sesardic, 1993).[114] On the other

[113] It should be noted that even single-gene disorders could be due to different mutations with different outcomes. In the case of PKU, the *PAH* gene, which is located on the chromosome 12, is related to different mutations (up to 500) with different symptomatic severity (see Scriver, 2007; Plomin et al., 2013).

[114] The overtaken assumption that a genetic defect was a necessary and sufficient condition to develop PKU stems likely from the fact that, in a normal range of environmental conditions—comprehensive of a diet

hand, it is a simplification to think that the genetic defect and the phenylalanine intake are *jointly sufficient conditions* for PKU: being organisms complex networks of developmental processes taking place over time, the conjunction of sufficient conditions to develop any particular phenotype is almost countless, involving the entire causal chain between the zygote and the adult phenotype which exhibits a given trait.[115]

A second remark concerns the clinical reasoning. Are pathologies *located inside* the genome? It is often said that a PKU heterozygous is a healthy carrier. This implies that PKU is a condition which goes *beyond* its clinical manifestations, something that is *inside* the genome. This leads to a sort of preformationism. Moreover, by assuming this view, one would say that *PAH* is "the gene for PKU" in a developmental sense instead of a population variation sense. I think we can legitimately say neither that PKU carriers are as such regardless their symptoms nor that people carry PKU in their genome regardless their environment (e.g., regardless their alimentary habits). PKU is primarily a metabolic disorder and secondarily a clinical picture. Without a phenylalanine intake, no disorder emerges. Without symptoms, there is no clinical picture to be identified.[116] The only thing we could legitimately say is, rather, that some individuals carry a particular allele. In other words, we can say that an individual is affected by PKU if and only if: a) she carries the genetic mutation; b) she does not watch a specific diet; and c) she develops symptoms. Briefly, only a specific organism-environment interaction leads an individual to be liable to a diagnosis. A genetic defect, by definition, is not a disease. Saying the contrary implies neopreformationism. Pathologies are not *there*, inside the organism, before development: they arise during development through the interaction between genes, their products, and the environment.

Finally, the most important remark. Are monogenic pathologies traits in the narrow sense? Are they comparable to height, weight, skin color, bones structure, organs, tissues, and metabolic aspects? What is, at the very end, a qualitative trait? The tension between qualitative and quantitative approach returns here: quantitative genetics concerns traits that are shared by every individual in different degree. In other words, a quantitative trait varies within the population but there is no individual which does not carry that trait. Every human being 'has' a height, a weight, a blood pressure, a liver metabolism, etc. Conversely, it is generally assumed that qualitative genetics concerns yes/no traits—e.g., an individual might have or not a pathology.

I suggest appealing one last time to Mather's account of qualitative variation (see Chapter 2), this time for the sake of reading monogenic pathologies. Let us take height

---

rich of phenylalanine—, every carrier of the defective gene tend to show the clinical symptoms, albeit in different degrees.

[115] Localizing causality in the conjunction of a DNA portion and a dietetic habit appeals to a *ceteris paribus* clause. We must keep that in mind.

[116] One might notice that some diseases are asymptomatic. However, this means that there are no clinical symptoms even if the disease is diagnosable. The case of healthy carriers is slightly different, because, apart from genomic scans, there is no clue that might be useful to diagnose the disease.

and dwarfism[117]. Normal height varies continuously within populations, so it is a polygenic trait. By contrast, dwarfism is neither a trait nor a monogenic one: it is rather a condition which appears to be either present or not because a single allele segregates sharply within the population. However, once again, the normal development of height is related to several genes. Then, we might say that height, but not dwarfism, is a phenotypic trait: dwarfism represents a *variant form* of the general trait height. I call traits like height the *proper phenotypic trait* and variations like dwarfism *variations on the theme,* that is, different forms of the proper trait.

Accordingly, PKU should not be understood as a trait, but rather as a variation on a theme. What is the proper phenotype, the *theme*? Two hypotheses. First, the phenotypic trait is limited to the enzyme phenylalanine hydroxylase, which is spread in the liver; hence, the theme on which PKU varies, is the functioning of an enzyme. Generally, this enzyme is capable to transform phenylalanine in tyrosine. In those people who carry two mutant alleles, the genetic product—the enzyme—is different than in the others. The presence of a defective *PAH* gene causes qualitative variation within the population. A second hypothesis does not take the enzyme as the proper phenotype. Rather, the proper phenotype is the set of metabolic reactions related to phenylalanine, the phenylalanine metabolism. In this sense, the theme on which PKU varies is a general metabolic aspect.[118]

If pathologies are not traits, but variations on a theme, we cannot say that there are monogenic, yes/no, qualitative, simple traits. Every phenotype unfolds under uncountable influences. On a genetic level, several genes are involved in every trait (at least if we do not think of enzymes as phenotypes). This is empirically true if we think of proper phenotypes like height, skin color, liver metabolism, and so on. Several genes are related to the normal development of liver metabolism and therefore this trait is polygenic. PKU represents a condition in which an abrupt interruption of normal development occurs, diverting it to a variant form.

According to the reasoning above, it is relatively easy to see why there are no single genes "for" complex traits: roughly speaking, *there are no simple traits at all*. What is "simple" is the population variation. In this respect, monogenic pathologies do not contradict the theoretical statement according to which there are no genes "for" pathologies, because pathologies are not traits in a strict sense. At most, a single gene might be a necessary condition to cause a population difference.[119]

---

[117] I take dwarfism not as a complex syndrome but rather as a variation over a polygenic trait (i.e., height), due to oligogenic inheritance—we might think about this type of dwarfism both in plants and animals.

[118] This proposal does not say anything different than what Mendelian genetics says. Indeed, Mendelian genetics is interested in traits variation. Therefore, qualitative traits and Mendelian traits concern, by definition, a variation on a theme. Nonetheless, this awareness seems to have been lost in several contexts, where speaking about qualitative and monogenic traits as proper traits, has come to be the norm.

[119] As I argued in §1.3, when we look for genetic causation on phenotypic traits, we would like to find genes which are included in the causal network that brings about a trait. The case of monogenic conditions is quite different: the mutant form of the gene is not involved in the causal network, but rather it is exactly outside of it (it does not code for functional enzymes).

Let us now turn to the attempt to explain yes/no traits in a quantitative manner (e.g., Plomin et al., 2009; see Chapter 2). The problem with the contemporary interpretations of the QuAdM—e.g., the multilevel quantitative model—is that diseases are conceived as proper phenotypes. For instance, schizophrenia has been taken as a proper phenotype, on which individuals' clinical severity differs quantitatively. Conversely, if I am right, phenotypic traits are intraspecific—shared by every member of the species but in different ways (not necessarily in different degrees!); therefore, schizophrenia should be intended as a variation on a theme. Which is the proper trait, is a matter of empirical studies.[120]

We can now understand what is the problem with assuming the QuAdM as a biologically-committed model: by thinking of developmental networks of complex traits—instead of simple monogenic traits—we could recognize why the effect of every element is individually important (the same applies on genes, of course) and strongly interactive at every level. Taking seriously interactions is likely the only way to address the problem of complex traits.

2.2. Quantitative traits

Is it possible to maintain the wording "quantitative trait"? If so, for which traits? In this paragraph, I shall argue that, while some classical quantitative traits (e.g., pigmentation) are related to specific genetic causation, some other quantitative traits (e.g., IQ) are doomed to be related to non-specific genetic influence. Therefore, I submit to rule out the term "quantitative traits" from genetics' vocabulary by stopping the widespread practice to quantify over phenotypes, and then expecting to find reliable associations between genes and behaviors. But let us proceed step by step.

As I extensively discussed, quantitative traits vary continuously within populations. Individuals are different to each other for these traits insofar as they manifest the trait in different degree (e.g., higher or lower IQ). However, the word "quantitative" can refer to two different phenomena that must be distinguished one from the other. In Chapter 2, I referred to pea color as a case of quantitative traits. However, pigmentation, if compared to height and IQ, should be considered quantitative—or, better, quantifiable—in a remarkably different manner:

[120] For instance, schizophrenia's proper phenotype might concern dopamine metabolism. However, it is controversial whether schizophrenia, with its heterogeneous symptomatic manifestations, could be intended as a general biological phenomenon. Several authors (e.g., Bearden & Freimer, 2006; Jablensky, 2006; Owen et al., 2007; Tabb, 2015; Tsou, 2016; Tsuang et al., 1990) argue that schizophrenia should be subtyped into different lower-level phenotypes which bring about specific symptomatic manifestations. The failure of genetics research in finding reliable associations between psychopathological traits and genetic variation is here imputed to such a confusion. If this worry is reasonable, the model proposed by Plomin et al. (2009) would be questionable from its first steps, i.e., the problem of defining the disorder and, hence, the phenotypic trait to be investigated. See Chapter 6 for a detailed analysis of the subtyping problem.

- Quantifiable-in-individuals: Some traits are quantitative, or some aspects of these traits are quantifiable, in the light of an *individual* developmental feature. Pigmentation (e.g., pea, human skin and eye color) concerns something quantifiable in the metabolic cascade which produces pigment. The genetic influence on these traits could be quantified because every locus controls enzymes which have an additive influence on the final phenotype—although, as I argued in Chapter 2, we should accept that the individual effects make a qualitative difference.

- Quantifiable-in-population: Some traits are quantitative, or are quantifiable, because of a *population* feature. In the case of height, weight, and IQ, what is quantitative is the variation within a population: we can order individuals on a scale for a trait, and we can obtain a bell curve (e.g., by standardizing IQ tests). But this does not mean that the examined trait *is* quantitative from an individual development point of view, nor that it is influenced additively by several alleles.

In other words, we must distinguish between *order* and *quantity*. As Hibberd (2014) argues, some phenomena are ontologically quantitative while others are qualitative. Qualitative phenomena, however, can be analyzed quantitatively by "ordering" over them:

> "An attribute is only quantitative if it is ordered *and* has additive structure, and there is no evidence that psychological attributes have additive structure (Michell, 2011). If the attribute of interest is not quantitative, it cannot be measured and any claim to be measuring is false. The crucial difference between order and additivity is that if an attribute can be ordered it displays degrees but the differences between degrees are *qualitatively* different from one another [Michell, 2009]. If, on the other hand, the attribute has additive structure (e.g., length, force, weight, and temperature), differences between degrees do not differ qualitatively but are homogeneous. […] Estimating the magnitude of a quantitative attribute is what is meant by measurement; the question asked is 'How much?' But if the attribute is without additive (quantitative) structure, this is obviously not the right question. The researcher could not possibly be measuring anything, and their use of quantitative methods is pseudoscientific. The attribute is, at best, ordinal and using a qualitative method is the only defensible practice [for example, ability tests, at best, identify a heterogeneous order of cognitive states (Michell, 2011)]" (Hibberd, 2014, p. 175).

According to the distinction proposed above, quantifiable-in-individual traits, e.g., pigmentation, are due to the additive effects of several pigments, which is brought about by several enzymes originated by the related loci. Here, the G-P map is relatively simple and the genetic influence *specific*: there is a clear causal chain between genes, their products, and the phenotypic trait. Hence, it is relatively simple to understand this G-P map in a reverse genetics perspective—that is, starting from DNA and going towards proteins.

The case of the quantifiable-in-population traits is quite different. The IQ concerns forward genetics, that is, going from the phenotype to the genotype. From an epistemic point of view, one cannot be sure about the right *grain of zooming*: Should we look for genes related to general intelligence, assessed by IQ tests, or should we look for genes related to the cognitive subcomponents of intelligence? What is the target of the genetic

influence? What is the best way to carve organisms in parts? What empirical problems arise if we adopt a bad definition of a phenotypic trait? It seems that there is not a single right answer to these questions. However, I believe that there are good and bad ways to identify phenotypes or to "cut out" organisms.

Allow me to say that defining a phenotypic trait is not just an empirical problem. Rather, the effort of achieving definitions and theoretical models is, at first, up to psychologists, psychiatrists, and cognitive scientists. As Ramus states,

> "cognitive scientists have an important part to play. For one thing, genetic analyses can only be as good as the characterization of the phenotype, and cognitive phenotyping is (or should be) in the hands of cognitive scientists. […] Cognitive scientists are needed in genetics precisely because good cognitive models are needed to design behavioral genetic studies. […] Ideally, we should put more brain in genetic studies themselves, i.e., by defining *neural phenotypes* that are related to the cognitive phenotypes of interest, and running genetic analyses on the basis of the former" (Ramus, 2006, p. 249).

What if behavioral phenotypes are not well-defined as geneticists think? Something will likely go wrong in empirical research. Psychometrics defines intelligence as a general cognitive phenomenon, and behavioral genetics takes it as a polygenic trait. However, the IQ might not be a general and uniform phenotypic trait, but rather a cluster of different traits (e.g., cognitive components). If this is the case, then analyzing statistical associations between IQ and SNPs could be problematic. For instance, such associations might not be strong as expected: if one seeks the genes for IQ, one will inevitably find non-specific genetic influence.

This might explain why GWAS have not been conclusive in seeking genes for behaviors: if one applies such a systematic method to a behavioral trait, one will find for sure small correlations between the trait and several genes. This is not, however, due to the fact that hundreds of genes code for that trait, but rather because several genetic products are (weakly) correlated with higher-level properties (behaviors) of the biological system. Conversely, if intelligence is decomposed into different specific phenotypes, also the identifiable genetic influence will be specific. Consequently, by analyzing the association between specific cognitive processes and genotype one might find stronger regularities.[121]

To summarize, I believe IQ is nothing more than a general quantification of several lower-level cognitive aspects of human brain. That is, it does not represent a phenotypic trait in the narrow sense. This implies, by the way, that the G-P map for complex traits is supposedly far more complex than in pigmentation-like cases. Understanding this point rules out the necessity to appeal to the multilevel quantitative analysis for phenotypic

[121] Take the case of mental disorders. If one analyzes correlations between the presence of a neurotransmitter and a disorder, and if the disorder is not well-defined, one will likely obtain confounding data (e.g., spurious correlations and low replicability rates). By contrast, by subtyping schizophrenia into more fundamental processes and mechanisms, genetics research could be more successful in finding reliable associations between behaviors, etiological factors, and individual differences.

traits like IQ, because we do not need to explain how genotypes bring about intellectual phenotypes. IQ is not a polygenic trait in the same sense in which a pea color is polygenic: there is no G-P linearity (as in the alleles-units model), nor any quantitative property to be explained over the various levels of the organisms (as in the multilevel quantitative analysis).

3. Non-Specific Genetic Causation

For concluding my discussion about behavioral traits and genetic causation, one last question must be addressed: Is non-specific genetic influence meaningful in any sense? Biometricians thought that it is useful to analyze IQ as a character because, by doing so, it is possible to explain differences and similarities among relatives. Behavioral geneticists think that the IQ represents an important target for genetics research, and this justifies the quest for its genes. Conversely, molecular biologists may think that IQ represents a too general trait for conducting interesting analyses. This last hypothesis seems to me the more compelling one.

Let us take political attitudes as an example. As far as geneticists say (e.g., Alford et al., 2005), it is a heritable trait. Is it worth thinking of it as a genetic trait? Perhaps, the political preferences are genetically influenced in a weak sense, although not in any way.[122] Nonetheless, the genetic influence for this trait might be too weak and long-lost to talk about the genes "for" political preferences. If this is the case, it would be epistemologically useless to consider such trait as a *character* (see above).

In §1.4, I distinguished between specific (direct) and non-specific (indirect) genetic causation. Accordingly, genes directly cause proteins, but they cause complex traits only indirectly. As I mentioned, this distinction might seem to be far-fetched. Indeed, one might say that, if a gene codes for a protein which causes the development of a trait, then one can also say that the gene causes the trait. This reasoning would lead to conclude that it is meaningful to talk about genetic causality upon complex traits, even if genetic causation is indirect and "diluted" in complex causal networks.

This counterargument relies on the assumption that causality is transitive. From a metaphysical viewpoint, Hibberd (2014) has convincingly argued against this position. According to his *field model of causation* (see also Anderson, 1938; Mackie, 1974), causation involves a context, or field of causally relevant conditions, instead of a simple two-term (cause-effect) sequence. Causes act upon *fields* (local and fine-grained conditions) and the effects are produced from the field itself. It does follow that causation is not a

[122] However, I do not have any idea about how genes and their products may determine this trait. As Charney (2008, p. 311) notices, the assumption that there could be such a thing as liberal and conservative "phenotypes" seems to be quite misleading: "rather than explaining the phenomena better than, say, traditional historical and cultural and sociological explanations, [it] render them mysterious, if not incomprehensible".

succession of transitive links in a chain: if A causes B, and B causes C, A does not cause C. Indeed, at each stage a different field is involved (see Figure 5.1).

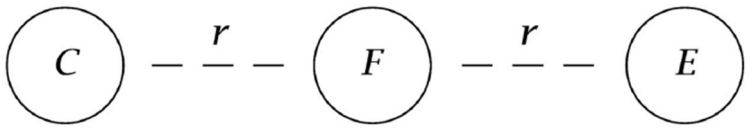


Figure 5.1: The logic of the field model of causation. From Hibberd, 2014, p.180.

It is worth noting that this account of causation fits with several aspects characterizing complex systems. In particular, it is possible to think of the field model of causation as regarding different steps in the developmental trajectory of a biological system.

Let us take fields as the condition of a biological system (S_n) at a specific time (t_n). S_0 represents the status of the system at t_0, before the entry of any causal effect. Let us suppose that at t_1 the system is subject to the set of causes [A_1, A_2, A_3, …A_n] that, for simplicity, we can consider solely genetic. Once the genetic causes affect the system, *the system changes* and the outcome (the newer conditions S_2) represent the *effect* at t_2. Then, the system is subject to the causes [B_1, B_2, B_3, …B_n], and a newer condition is reached at t_3 (that is, S_3). The same applies to the third set of causes [C_1, C_2, C_3, …C_n] at t_3 and so on (see Figure 5.2).

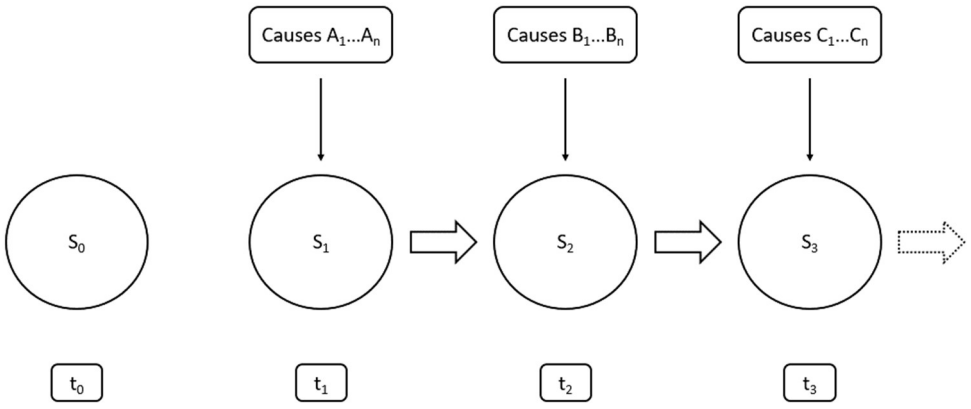


Figure 5.2: An application of the field model of causation to complex systems.

The important thing is that the second set of causes [B_1, B_2, B_3, …B_n] is different from the set [A_1, A_2, A_3, …A_n] because, among other differences, it includes the newer emerging features of the system at t_1 (that is, S_1) that were not present at t_0. Therefore, one cannot say that the set of causes [A_1, A_2, A_3, …A_n] has caused the conditions of the

system at t_3 (that is, S_3), because S_3 causally depends on [B_1, B_2, B_3, …B_n], a set of causes which includes the conditions of the system at t_1 (that is, S_1). Briefly, causality cannot be understood as transitive because it is not true that A_1, for instance, has a direct causal role in the features of S_3.

Pearl (2009) accounts for similar intuitions by looking for causal inference within the context of statistical analyses. In its discussion, Pearl distinguishes between direct and indirect effects:

"The term 'direct effect' is meant to quantify an effect that is not mediated by other variables in the model or, more accurately, the sensitivity of Y to changes in X while all other factors in the analysis are held fixed" (Pearl, 2009, p. 133).

Conversely, indirect effects characterize non-linear systems, where complex interactions and feedback loops make impossible to accept the transitivity of causal effects. Take us a model with four variables (X, Y, U, and Z), where causality follows the following pattern: X → Z ← U → Y. Here, X has no direct effect on Y.

"Holding Z constant would sustain the independence between X and Y, as can be seen by deleting all arrows entering Z. But if we were to condition on Z, a spurious association would be created through U (unobserved) that might be construed as a direct effect of X on Y. […] In nonlinear systems, the values at which we hold Z would, in general, modify the effect of X on Y and thus should be chosen carefully to represent the target policy under analysis" (Pearl, 2009, pp. 133-134).

Moreover, the issue at stake does not solely pertain to spurious correlations and to our inability to disentangle confounding effects. The problem is also that genes do not directly cause behaviors, and in many cases, they do not cause complex traits. Rather, their influence, once arrived at proteins, interacts with a multitude of causal effects coming from the other levels of the system and from the environment.

To illustrate this point, let us take playing basketball. Genetically based height is likely associated with playing basketball. However, as Ratner (2004) notices, the association is due to the rules of the game leading coaches to select tall players, and the rules of the game attract tall people to play because of financial and social benefits (read: genes-environment correlations, see Chapter 4). Of course, several genes are involved in these situations. But are they genes "for" tallness or "for" playing basketball? Is their causal influence directed towards those general phenotypic traits? I would say no.

As far as I can see, several genes are involved in the length of the bones; after all, height is a quantification over the whole dimension of the skeleton. Moreover, they act on bones' length indirectly: indeed, proteins affect cellular features, not the dimensions of the skeleton. Several further steps are required to reach the "basketball-trait": genes related to the development of the cells in the bones do not themselves produce basketball

players. Indeed, there are both tall people whose do not play basketball and short people whose do play it.[123]

Intelligence is likely even more complex than this. Genes code for proteins that act on cellular features, e.g., proteins that are expressed in neural cells and affect cellular properties related to the neural functioning. However, I would not say that genetic causation directly reaches higher-level phenomena, such as intelligence. A "normal" intelligence require everything being alright with genes and a "normal" development. There is likely nothing but this in genetic causation on cognitive processes and behaviors.

To conclude, when we consider a character—that is, something attested to be genetically influenced—we are not necessarily dealing with something that it is interesting or helpful to consider as "genetic".

Conclusion: Is Intelligence a Phenotypic Trait?

In this chapter, I argued that the relationship between genes and behaviors has not been properly investigated by behavioral genetics. An excessive focus on the statistical analysis of individual differences has hidden the very question on the line, which regards genetic causation. The problem of causality in biological systems is an open-ended theoretical problem which will likely attract the attention more and more in the next few years. As Noble notices,

> "There are different forms of causality, ranging from proximal causes (one billiard ball hitting another) to ultimate causes of the kind that evolutionary biologists seek in accounting for the survival value of biological functions and features. Genetic causality is a particularly vexed question partly not only because the concept of a gene has become problematic […] but also because it is not usually a proximal cause. Genes, as we now define them in molecular biological terms, lie a long way from their phenotypic effects, which are exerted through many levels of biological organization and subject to many influences from both those levels and the environment" (Noble, 2008, p. 3012).

I cannot draw any definitive conclusion about the relationship between genes, cognitive processes, and behaviors, if not some distinctions: 1) the distinction between differential and non-differential influence; 2) the distinction between specific and non-specific causation; and 3) the distinction between the three ways of being a gene "for". It might be useful to summarize the definitions we achieved in this chapter:

- Genetic causation: A gene causes/is "for" a phenotypic trait if the gene plays a role in the causal network that leads a zygote to develop into a multicellular organism which carries that trait.

[123] Consider that this could be a confounding factor for GWAS: if one looks for genes for playing basketball, one will find several statistical associations but none of them seem to be interesting for understanding genetic causation on the basketball behavior.

- Trait: An observable or measurable feature of an organism (e.g., pea color and IQ). Referring to a trait does not involve any commitment to the involvement of genes for the development. Traits can be both complex or simple.

- Character: An observable and measurable feature of an organism for which there is a direct genetic information (i.e., at least a gene "for" the character). In most cases, this definition can be applied to genetic products (e.g., RNA and proteins) but not to complex traits. In other words, it is very likely that this definition can be used for genetic products and cellular features only. Indeed, only traits for which the genetic causal chain is very simple (e.g., pea color) can be said to be "characters" with direct and specific genetic influence.

- Proper traits: Polygenic and complex features of an organism for which genetic causation can be both direct and indirect. When causation is indirect, however, epistemological, empirical and metaphysical concerns may arise about the usefulness of talking about genetic influence which is spread in complex interactive and multilevel networks.

- Qualitative traits (or monogenic pathologies): They are no proper traits. Phenotypic variation within a population is due to the existence of a gene which does not code for anything (a gene "not for" X).

- Quantifiable-in-individual traits: They are proper traits that are characterized by additivity on an individual level. Pigmentation is quantifiable in the metabolic cascade which produces pigment. The genetic influence on these traits could be quantified because every locus brings about enzymes which have an additive influence on the final phenotype.

- Quantifiable-in-population traits: They are not proper traits, but rather the compound of lower-level characters (with attested genetic influence) and traits (without attested genetic influence). The genetic influence on these traits is generally indirect.

I submit to avoid any reference to qualitative and quantitative traits. On the qualitative side, every trait is polygenic—that is, there are no monogenic traits. All those apparent exceptions, e.g., pathologies, are not really traits, but rather variations on what I defined as proper phenotypic traits. On the quantitative side, it is not convincing to name polygenic traits as "quantitative", because such phrasing does not recognize different ways in which a trait can be quantitative. To say that a trait is polygenic, and to say that

a trait is quantitative, are not the same things at all: variation within a population may be quantitative, but traits must be understood as solely polygenic.

If what I said applies to intelligence, then intelligence cannot be considered inherently quantitative. My hypothesis is that, even if the IQ varies quantitatively within populations, it does not follow that the QuAdM is well-suited to explain intelligence from a biological point of view. Indeed, the IQ is a general quantification over different cognitive processes. If there are genes related to intelligence, their targets are likely transcribed elements which act on cognitive processes, not on such a general phenotype.

However, this conclusion requires more arguments. It is the aim of the next chapter to provide them. I shall make use of the natural kinds theory to argue in more details why: a) general intelligence cannot be conceived as a worthwhile phenotype for genetics research; and b) intelligence cannot be conceived as a quantitative trait.

Chapter 6.
Natural Kinds to the Aid

Although the concept of intelligence is shrouded in a very controversial aura, it is nonetheless widely used both in genetics, in psychological, and in folk settings. In the previous chapters, I analyzed contemporary genetics research by assuming an overlap between the two concepts of IQ and general intelligence. However, as I mentioned in Chapter 1, several controversies characterize such an overlap. After a century of research, there still is, indeed, an extensive debate going on about the status of intelligence. It is now time to take into consideration these aspects by exploring the relationship between the intelligent quotient, the *g* factor, and intelligence. As I discussed, questions have typically involved empirical problems related to genetics or psychometrics. In this chapter, I adopt a more philosophically-oriented viewpoint to offer a conceptual analysis of the subject matter, involving natural kinds theory.

In §1, I recall the core features of the psychometric-genetic model of intelligence (PSY-GEN) and the controversies around it. In that section, I also discuss issues related to the natural kinds theory, explaining why it represents a useful tool to achieve a better understanding of the concept of intelligence in respect to both epistemological and ontological concerns. In §2, I clarify the essentialist assumptions underlying those controversies by appealing to heritability and molecular research as two methodologies that bear two different types of essentialist thought. In §3, I propose a reconstruction of general intelligence as a homeostatic property cluster kind (HPC). That analysis will serve three main purposes: 1) making sense of the PSY-GEN conceptualization of general intelligence within an ontologically-committed framework; 2) ruling out essentialism; 3) accounting for some common intuitions about cognition. Finally, in §4, I submit that it is unnecessary to conceive intelligence as a unified cognitive phenomenon or as a kind.[124]

1. Biological Intelligence: Ontological Issues

Over the last century, psychometrics has studied intelligence in the light of two theoretical constructs: IQ and *g* factor. IQ represents the individual intelligence level assessed by tests and is, in a sense, a behavioral outcome. Instead, *g* can represent two different things: on the one hand, the outcome of a correlation matrix of cognitive test scores; on the other hand, the IQ's psychological explanation. In this second sense, *g* is

[124] The contents of this chapter have been published by *Philosophical Psychology* (November 2017) in a paper entitled "What Kind of Kind is Intelligence?".

conceived as a general cognitive ability that underlies individual test performances: whatever intelligence is, it is measurable by IQ tests, while *g* explains individual intellectual differences. Therefore, I shall distinguish between the psychometric *g* and the psychobiological *g*.

At some point, psychometrics met behavioral genetics and its quantitative analysis of intelligence. The two fields have benefited from a mutual and flourishing influence, at least apparently. As I analyzed, for many decades the main goal of behavioral geneticists has been to understand how relevant inheritance is in the explanation of individual differences and similarities, by the adoption of the psychometric IQ as a good 'index' of individual intelligence. The two scientific enterprises have over time converged into a unified model, that I called the PSY-GEN model of intelligence. After the adoption of the *g* factor by genetic research, a consensus has been reached around a conception of intelligence as a highly heritable general cognitive ability. Despite the widespread disagreement about *g*'s psychobiological meaning, the PSY-GEN model takes *g* as a prominent psychological variable (Detterman, 2002; Jensen, 2002): "*g* is one of the most reliable and valid measures in the behavioral domain" (Plomin et al., 2013, p. 187). However, neither intelligence nor the *g* factor has been allocated a position in the agenda of cognitive sciences.[125]

Historically speaking, the existence of a general factor of intelligence has been hypothesized by Charles Spearman (1904, 1923). The reason is straightforward: intelligence measurements are positively intercorrelated. Though to varying degrees, if one shows good performance on a given task, one tends to show good performance also in other tasks. This empirical phenomenon is called 'positive manifold'. Thus, *g* is a summary index of a correlation matrix, representing what cognitive tests have in common. In this respect, the *g* factor is relatively uncontroversial (see Chapter 1 for some disagreements).

The subject of the controversy lies in the psychobiological nature of *g*. The PSY-GEN approach does not attempt to understand *g* in any strong ontological sense. As is often the case in psychometrics, a clear-cut distinction between methodological purposes and the reality of a psychological construct is endorsed. In other words, it does not pertain to psychometrics to explore *g* ontologically; it is sufficient to ensure that IQ tests can evaluate intelligence—whatever it is. Nonetheless, in the light of the positive manifold, several psychologists have accepted the existence of an underlying general mental ability (see Garlick, 2002; Van der Maas et al., 2006). Hence, some ontologically driven hypotheses were adumbrated, especially when psychometrics came face to face with biological sciences. The advocates of *g* conceive this factor as a cognitive phenomenon

[125] Since the IQ test skills clearly belong to the cognitive domain, this divergence might strike one as surprising. Nevertheless, the attempt to bridge psychometrics and cognitive sciences is relatively recent—especially considering that the psychometric approach to intelligence dates back to the early twentieth century (see e.g., Pretz & Sternberg, 2005). As I show in Chapter 7, the two traditions of psychometrics and cognitive psychology landed to quite different views of intelligence.

responsible for individual differences in test performances.

The degree of the realist commitment is different depending on the authors' standpoint, but most of them admit *g* as a psychobiological characteristic which influences intelligent behavior (see Chapter 1 for some contrasting positions). For instance, Spearman described *g* as a form of mental energy. More recently, to guarantee that *g* is a valid measure of intelligence, some scholars have tried to relate it to more reliable constructs, like the ones coming from neurocognitive science—e.g., working memory, processing speed, neural efficiency and brain size (see Chapter 7). Briefly, *g* must exist somehow: we can look for its biological correlates to ground it in other cognitive phenomena which seem to exist, insofar as they do not arise from mere statistical research.

By contrast, several authors cast doubt on a strong interpretation of *g*. According to Kray and Frensch (2002), there is no convincing empirical evidence that supports the existence of *g*. For Stankov,

> "there is no single cognitive process that can explain the presence of *g*. […] It is a mixture of many different processes (including non-cognitive influences) that are known to change in the course of development" (Stankov, 2002, p. 35).

Although the advocates of the PSY-GEN approach rarely engage in the philosophical debate, it is fruitful to discuss the theory of biological intelligence and the natural kinds theory conjointly. Within the PSY-GEN model, general intelligence has been reified as a phenotypic trait in the narrow sense. This conceptualization is supposed to reflect something about how cognition works on a biological level. The sort of instrumentalism which often characterizes psychometrics is, in this respect, set aside in favor of a realist view (see Chapter 1).

The link with natural kinds is straightforward. In recent years several efforts have been made to establish whether psychological constructs are natural kinds.[126] These efforts are aimed at exploring the extent and degree within which such concepts meet ontological and epistemological requirements imposed by the natural kinds theory.

Within this theoretical challenge, a link is assumed between scientific realism and natural kinds. Introducing natural kinds allows us to offer a framework within which we can evaluate the ontological status of psychological constructs apart from our scientific theories and categorizations (i.e., as mind-independent). Proponents of the realist view on natural kinds assume that science is able "to carve nature at its joints": entities identified and classified by science correspond to real kinds in nature, tracing natural properties and relations (see Bird & Tobin, 2015; Campbell et al., 2011; Franklin-Hall, 2015; Slater & Borghini, 2011). For instance, discovering biological pathologies correlated to mental disorders (lesions and genetic issues) is often considered analogous to discovering the

[126] Renowned examples are emotions (Barrett, 2006; Griffiths, 2004) and psychiatric disorders (Kincaid & Sullivan, 2014).

atomic number of a chemical element and provide an important foundation for psychiatric nosology and research (Stein, 2014; Zachar, 2014).[127]

If the PSY-GEN approach to intelligence was involved in philosophical debate, it would be a prototypical theory in this respect: thinking of the notion of intelligence as a natural kind would be a way (likely, not the only one) to say that general intelligence deserves a place in our ontology, or a way to exert a clear ontological commitment to general intelligence or to *g*. I will explore this hypothesis.

If intelligence is a kind, one may ask what kind of kind it would be. For three main reasons, the most promising hypothesis points to the HPC theory, introduced by Richard Boyd (1991): the first one concerns traditional essentialism; the second one depends on the fact that Boyd's theory seems suitable for tackling anti-realism; the last one pertains to multilevel analysis. I extensively discuss these points in §2.3.

Can the PSY-GEN model of intelligence be accounted for by a realist viewpoint? If the answer were positive, then this model would reflect the way in which cognition works on a biological level—briefly, it would be suitable to carve the psychological nature at its joints. In particular, the model would denote something in the world that should be admitted along with other neurocognitive processes or architectures into the realm of trustworthy concepts.

1.1. How do behavioral geneticists conceptualize intelligence?

Before proceeding in evaluating general intelligence in the light of the natural kinds theory, it is important to ask whether, and to what extent, behavioral geneticists would be sympathetic with such an approach.[128] As it has been said in the previous chapters, since the nineteenth-century studies of Galton, intelligence has been considered a quantitative trait. This has led to the quantification of intellectual manifestations and to the categorization of people.[129]

---

[127] It is not my aim to evaluate which is the best way to understand natural kinds, that have been invoked in semantics, metaphysics, and epistemology. Here, I mainly assume a metaphysical standpoint—although not excluding some epistemological concerns.

[128] As I said in Chapter 1, we should bear in mind that both psychometrics and behavioral genetics are quite heterogeneous scientific fields; it will be hence necessary to simplify some aspects of the PSY-GEN theory of intelligence to embrace as many positions at stake as possible. Here, explaining the genetics conception of intelligence, I mainly refer to Plomin et al. (2013).

[129] Albeit practical applications have changed largely depending on social circumstances (see Chapter 1), the chief aim of testing is to *measure*, while IQ stands for that number which is useful to *sort* individuals according to their intellectual features. In the clinical context, for instance, intellectual disability is considered a clinical picture related to various diagnostic criteria, among which the first pertains to low IQ level. Other criteria address adaptive functioning for social standards and intellectual and adaptive deficits during development (see American Psychiatric Association, 2000). So, the relevance of IQ testing especially arises for behavioral genetics research. In fact, genetics is not interested in the clinical picture itself, but rather in the so-called "general cognitive disability", which concerns only low IQ (Plomin et al.,

Qualitative traits, like those generally analyzed in Mendelian genetics, fit into discrete distinctions, allowing categorical reasoning. This is not the case for quantitative traits. Thus, the quantitative conception of intelligence would hardly find room in the natural kinds theory. Indeed, it is not immediately clear what the role of quantitative traits in a kinds framework can be. Nick Haslam (2014) states that natural kinds involve categories rather than dimensional aspects—the latter are not kinds in the narrow sense. Somehow, Haslam's intuition seems to be convincing.[130] Often, when one speaks about kinds, one has in mind a set of properties. Classically, natural kinds have been involved in the relationship between properties and classes (see Hacking, 1991). Of course, this is not the only issue related to natural kinds, but focusing on properties became prominent in many theoretical contexts after Boyd introduced his HPC theory, readily adopted by many authors across several fields, from psychology to natural sciences.

Thus, natural kinds theories have been taken to evaluate the ontological status of these properties, how they combine, and whether they are necessary and/or sufficient to define a kind. Let us consider, for example, psychiatry: when we ask whether a mental disorder should be considered as a natural kind, we pick up a collection of properties (behaviors and symptoms), analyzing the relationship between cluster and properties; whether one of them is necessary, for an individual, to join the kind; whether one of them is necessary to define it.

The situation with intelligence appears to be different for at least two reasons. First, intelligence is supposed to be a dimensional phenomenon, that thing measured by intelligence tests that changes among individuals according to a bell curve. However, intelligence is not solely something dimensional because of this statistical feature: it is theorized as a quantitative trait *in itself*—as I said before, the *way* in which genes influence intelligence *is* quantitative. Second, from a genetics point of view, *g* is not a property cluster—that is, a collection of properties in a Boydian sense. Rather, it is *a single property*, a unified psychological trait shared by every human being in varying degrees. For instance, this variable could figure among other symptoms related to a mental disorder: in fact, intellectual disability is a medical picture in which low IQ figures as a symptom, or a diagnostic criterion, among others. Generally, the proper targets of a natural kinds inquiry are not the symptoms (the properties), but rather the disorder itself (the cluster). For instance, the low mood could be intended as something dimensional, too. However, it does not make sense to ask whether it is a natural kind: at most, it would be a necessary (but not sufficient) property to define mood disorders such as major depression.

2013, p. 163). So, for the general disability, four degrees of severity are generally assumed: profound (IQ < 20), severe (20 to 35), moderate (35 to 50) and mild (50 to 70).

[130] John Dupré (personal communication, April 2016) cast similar doubts: could it be possible to account for variables in a natural kinds context? To satisfactorily answer this question would go beyond my purposes. I shall show that we need an argument to arrange general intelligence within a kind framework.

Although I am arguing for a kind view of intelligence, it is important to notice that geneticists might not accept this analysis. Intelligence can look like height, another quantitative trait which would represent a single property unlikely useful for the natural kinds theory. Where would the property cluster be? What about the membership criteria for such a kind?

In the next paragraphs, I shall show that this theory of intelligence, related to quantitative trait analysis, carries as an implication a sort of essentialism that Paul Griffiths (2002) called "folk essentialism". By arguing in this direction, I will pave the way for a kind theory of intelligence. However, not every kind theory is suitable for accounting for intelligence: what one might want to avoid is the so-called traditional essentialism, which is frequently related to natural kinds inquiries. I shall, therefore, ask what should be taken as the best natural kind theory on the market.

2. What Kind of Kind Theory?

Not every type of kind theory is suitable to describe complex psychological phenomena like intelligence. For instance, it is quite uncontroversial that the classical concept of natural kind fails both in psychological and biological sciences. Indeed, for a few decades, essentialism has been typically related to the natural kinds debate as the bigger evil that philosophers had to face with. According to Haslam, the traditional conception of natural kinds rarely applies to psychology since it carries an essentialist luggage:

"Only some causal stories can produce categories that might qualify as natural kinds in the classic, essentialist sense. [The basis of a natural kind] is a single cause that is common to all category members and that directly gives rise to the kind's properties. In the psychiatric domain, for example, a discrete disorder whose clinical features ultimately derived from a specific neural or genetic dysfunction that was shared by all afflicted individuals would qualify as a natural kind in the sense intended here" (Haslam, 2014, p. 16).

In the following paragraphs, I discuss two essentialist relapses frequently detectable in genetics research: folk essentialism and traditional essentialism. As I mentioned, traditional essentialism often involves discourses about kinds and properties:

"Traditional essentialists hold that natural kinds must possess definitional *essences* that define them in terms of necessary and sufficient, intrinsic, unchanging, ahistorical properties" (Boyd, 1999, p. 146).

Conversely, folk essentialism represents a tacit assumption that underlies several scientific debates. According to Griffiths (2002), folk essentialism is a distinctive feature of pre-scientific thought about animate things. It understands biological species as the manifestation of underlying 'natures' shared by all members of a species (Griffiths, 2002,

p. 72).[131] Like traditional essentialism, folk essentialism explains similarities by referring to underlying properties shared by the members of a kind. However, the latter does not appeal to any theoretical reasoning: it is simply a psychological tendency.

Behavioral genetics frequently seems to think in essentialist terms. With respect to folk essentialism, I shall consider heritability analysis. With respect to traditional essentialism, instead, I shall consider molecular research, especially the candidate-gene approach.

2.1. Folk essentialism in heritability research

Folk essentialism generally resides below controversies about innateness. Indeed, Griffiths describes the innateness concept as an expression of folk essentialism. It may involve the idea that the development of an innate trait is established in advance and is inflexible, grounding behavioral differences in genetic differences (see also Samuels, 2004). Innateness and folk essentialism are counterparts of two widespread tendencies of behavioral genetics: the first one pertains to neopreformationism; the second one, pertains to the tendency of partitioning causes adopted within heritability research.

As it has been mentioned, geneticists estimated high heritability for the IQ—generally between 50% and 80%. However, detractors of heritability research state that from these data one cannot legitimately infer anything about genetic influence on a phenotypic trait. Rather, such concepts as 'innate', 'genetic' and 'heritable' have no clear connection with each other. Nonetheless, jumping from heritability to genetic causation is very frequent. Often, this leads even to strong genetic determinism, falling into very popular simplifications ("this behavior is innate", "it is part of human nature", "the gene for…has been found", etc.). As I shall show, the improper use of heritability data hides folk-essentialist assumptions.

As I said, folk essentialism is strictly related to the idea that the development of an innate trait is established in advance or, at least, is hard to change and insensible to external influences. Such a thought is pervasive in the psychometric-genetic literature. The assumption that an IQ test could tell us something about the intellectual destiny of anyone is an effective element that is commonly taken as a criterion for public policies and their ethical and societal consequences. For instance, developmental fixity results in specific policy implementation about education, where children are directed towards a specific educational path according to their intellectual attitudes. This is by no means something new. In a classical publication, Arthur Jensen (1969, p. 8) stated that currently

[131] Folk essentialism looks like psychological essentialism (Gelman, 2004), which is defined as the propensity to think that similar individuals must share an underlying nature responsible for their resemblance. As Medin explains, "people act as if things (e.g., objects) have essences or underlying natures that make them the things that they are" (Medin, 1989, p. 1476).

used IQ tests do indeed reflect innate, genetically determined aspects of intellectual ability in persons from the population on which the tests were standardized and validated.

Let us now consider data from the longitudinal project MISTRA (*Minnesota Study of Twins Reared Apart*). The team headed by Thomas Bouchard analyzed in many different aspects several MZ twins. According to Write's report (1997), meeting again as adults, these twins revealed very interesting resemblances: not only similar IQs or personality traits, but also shared hobbies, political attitude, religious preferences, similar partners, similar pets, and so on.

Lastly, the magnitude of genetic factors seems to increase during development: the IQs of MZ twins correlate to each other more and more with aging. It seems to suggest that environmental factors become almost irrelevant during adulthood:

"early in life, shared environmental factors are the dominant influence on IQ, but gradually genetic influence increases, with the effects of shared environment dropping to near zero" (Bouchard, 2004, p. 149).

In these examples, we may easily recognize the idea that genes work prior to environment to canalize the organism's development: the core idea is that genes can determine complex traits like intelligence. Often, this is assumed regardless of any reference to developmental mechanisms and without invoking any non-genetic influence as *really* relevant. This is related, in a sense, both to neopreformationism and to the tendency of separating developmental causes. While Griffiths enlightened a link between the innateness concept and folk essentialism, I shall explore the relationship between folk essentialism and the quantitative view of intelligence.

Heritability analysis allows us to speak about genetic and environmental causal pathways as separate things—merely additive and quantifiable in percentile terms (see Chapter 4). Trivially, for thinking that inheritance plays a greater role than environment (e.g., 80 *versus* 20), one must assume that the genetic causal power can be separated from the environmental one. This is closely reminiscent of folk essentialism: if, from a causal point of view, genes and environment can act separately, then the genome could be that underlying property capable of explaining individuals' similarities.

One may notice, in the light of the previous discussion, that this assumption originates from a methodological requirement. Heritability research separates genes and environment for analyzing phenomena which are very complex in a natural context. Nonetheless, this "carving perspective", separating genes and environment in development, originates from an ontological assumption which is *prior* to heritability research and arguably related to folk essentialism. This assumption was born, at least, when Galton distinguished conceptually *nature* and *nurture*, remaining hidden through decades of research under methodological purposes. Then, with behavioral genetics, the carving perspective became visible as an explicit ontological conception. Moving from an artificial selection context—e.g., about plants inbreeding—towards complex human

traits as intelligence, a methodological artefact has been taken as a biological principle: the effect of genes and environment is merely additive and there is no relevant interaction between the two. Briefly, findings deriving from heritability research, based on an "artificial" distinction, led several scholars to think that genes-environment interactions are negligible.

For a few decades, developmental biology has brought convincing reasons for thinking that genes-environment interactions play a chief role in development (see Chapters 2-4). In this respect, folk essentialism seems to arise more easily in a quantitative view than in an interactionist one, because, taking interactionism seriously, one cannot make a relevant distinction between the causal power of genes and of the environment on phenotypes. Following such a line of reasoning, we need to give up on a quantitative view of intelligence to set aside this sort of essentialism with respect to intelligence.

In §2.3, I evaluate a kind theory of intelligence to analyze whether it would be more convincing than a quantitative one. However, since a link between natural kinds and tradition essentialism is frequently highlighted, I now show what form essentialism tends to assume in molecular research.

2.2. Traditional essentialism in molecular research

After many decades of studies committed to the estimation of heritability, behavioral genetics has adopted methods to analyze genetic resemblances among individuals finding associations between phenotypes and genotypes. Presumably, if two people show the same features (like a mental disorder or an analog IQ level), they must share some genetic factor.[132] The main target of these methods is to find specific alleles involved in the heritability of a given phenotypic trait—in our case, involved in the heritability of the IQ (see Plomin et al., 2013, pp. 206-209).[133]

Such attempts might appear in contrast with a quantitative conception of intelligence, conceiving intelligence as being weakly influenced by many genes. In other words, since genes have small individual effects on intelligence, it seems hard to identify single alleles related to individual intellectual differences (see §1.1). To shed light on this point—and to make sense of the geneticists' view—I suggest that we are dealing with a categorical reasoning applied atop a dimensional one. Accordingly, we should presume that a

[132] This assumption would be true if heritability research were suitable to ensure that a highly heritable trait is influenced by genes. Unfortunately, molecular research considered as reliable findings deriving from heritability research without further checks (Joseph, 2004).

[133] In this paragraph, I mainly focus on the candidate-gene approach, which is more related to traditional essentialism than genome-wide association studies (GWAS). Indeed, the candidate-gene approach seeks genes that are supposed to be individually important for a trait; conversely, GWAS scan the genome more systematically and look for genes that are not supposed to be individually important. Hence, GWAS concern quantitative genetics and folk essentialism (see §2).

quantitative conception of intelligence admits some *qualitative* distinction: on a quantitative view, no single gene plays a significant role for an individual's IQ level; however, it is possible to find specific genetic variants that are sufficient (and sometimes necessary) conditions linked to a specific IQ level.

In order to clarify this point, let us consider two types of mental phenomena that are often taken to be equivalent: a general cognitive disability due to monogenic conditions, and a general cognitive disability characterized by a threshold on a dimensional scale of values. Let us examine them in turn. Some cognitive disabilities depend directly on inheritance, e.g., in phenylketonuria (PKU). Broadly speaking, one could say that the PKU-related cognitive disability is due to a specific monogenic issue.[134] In this case, we could perhaps legitimately frame PKU inside a categorical-essentialist perspective: every PKU-related behavior (like low IQ) is linked to a specific biological feature. The category members share the same "causal story"—a specific genetic variant and a "normal" diet—which is a necessary and sufficient condition to develop the disease. However, in most cases, cognitive disability is the "negative" pole on a dimensional scale, the outcome of a complex individual story that involves several causal pathways, e.g., birth problems, nutritional deficiencies, head injuries, social or educational issues (see Plomin et al., 2013, p. 164). As a core difference, none of them would be individually necessary to obtain a low score on IQ tests. Thus, with respect to "dimensional disabilities", one cannot readily adopt a categorical-essentialist approach.

Behavioral genetics rarely makes this distinction explicit, treating dimensional variation in a categorical way, like monogenic conditions.[135] The process by which molecular genetics superimposes a categorical reasoning upon a dimensional one is now clear: generally speaking, intelligence is conceived as a quantitative trait; however, the recognition that some genes could have, even if taken in isolation, an appreciable effect on intelligence—as in PKU—leads geneticists to think that some genes are more important than others in explaining the IQ's heritability. Molecular research—the candidate gene approach especially—tries to find them.

Accordingly, one might think of the two types of cognitive disability differently in relation to essentialism. On the one hand, there is the dimensional phenomenon, that is related to quantitative genetics and heritability research, and is therefore tied to folk essentialism (§2.1); on the other hand, there is the monogenic phenomenon. In this case, we could identify the traditional form of essentialism, insofar as carrying some genes

[134] This is, however, a simplification: what is genetically determined in PKU is the metabolic issue, not the cognitive disability. See Chapter 5 for more details.

[135] It has been often assumed that, if cognitive disabilities due to monogenic inheritance exist, then genes involved directly in IQ level should exist as well. This argument is, at best, questionable. There are conditions, like PKU, in which a single genetic variant plays a role in IQ level: specific alleles can be involved in dysfunctional developmental mechanisms which lead to an abrupt break in normal development—and, consequently, to low IQ. However, it does not follow that any IQ level (both within and outside the normal range of values) is always determined by the effect of specific genes.

would be a sufficient (and, in some cases, a necessary) condition to develop a cognitive disability. Thus, traditional essentialism is related to molecular research *desiderata*, according to which it might be ideally possible to detect every gene involved in the development of intelligence.

By emphasizing the role of genes as underlying elements, molecular research leaves the door open to an essentialist framework. Such a framework seems to be unsuited to account for biological and psychological kinds because, as Zachar (2014) notices, they tend to be the outcome of the interaction of several internal and external causes over time.

2.3. Why homeostatic property clusters?

As I mentioned, the classical natural kinds theory is often evaluated as unsuccessful to describe psychological phenomena. For instance, psychopathological symptoms may depend on many causal mechanisms. As far as we know, there are few clinical conditions due to single lesions or genetic defects: more often, several factors are involved in many outcomes, while the same outcome may have several causal stories.

Among several theoretical proposals devoted to replacing the classical view, one of the most accepted was introduced by Boyd. Before discussing the theory, let me highlight two relevant purposes of the author that have influenced my subsequent choices. The first addresses Hacking's objection (1991), according to which natural kinds are, to some extent, mind-dependent. Hacking holds that our scientific taxonomies cannot trace the real structure of the world because boundaries depend on epistemic purposes. The second aim of Boyd's theory is to avoid essentialism about kinds and properties. Frequently, in the biological domain, one cannot identify necessary and sufficient conditions for a kind membership:

"[Biological kinds] are more heterogeneous than elements in a periodic table. Unlike all atoms of gold, individual members of a species need not share all their properties" (Kendler et al., 2011, p. 1147).[136]

A theoretical solution to both these issues consists of postulating the existence of a homeostatic mechanism, a causal pathway that explains why properties are statistically clustered together. Unpredictability and difficulties in categorization represent the main problems with biological entities: in virtue of their variability, they do not fit into narrow categories. The homeostatic mechanism grants enough flexibility to admit even huge variations among a kind's members. Furthermore, it grants stability to the kind, allowing

[136] Let us suppose a lack of any "deep" biological knowledge about species membership. For instance, we may refer to different tigers as belonging to the same kind because of their surface properties. Nonetheless, we may also name as 'tiger' a tiger which lacks many of those surface properties (e.g., a tiger without stripes, without claws, and with just three legs). In fact, we tend to think that (at least some) surface properties are not relevant to define the *nature* of an individual and its species membership.

us to ground prediction and inductive reasoning. Hence, good taxonomies might support successful scientific practices insofar as they are capable of tracing (some aspects of) the causal structure of the world, i.e., its causal mechanisms.

So, concerning essentialism, Boyd's theory admits individual cases in which not every property is shown. Indeed, the essence of a kind does not involve properties. Rather, it involves, at most, the mechanism.[137] At the same time, Boyd tried to save the notion of natural kind as mind-independent—accounting thus for realism—by highlighting the difference between a mere set of properties and a natural kind. Such a difference consists in a *non-arbitrary association of properties*, based on the existence of a mechanism. This causal link is very important: without it, properties would be unrelated with each other, forming a mere *property set*; but, if causality is established, properties form a *cluster* (Khalidi, 2014; Wilson et al., 2007). In this second case, a kind can satisgy both ontological and epistemological requirements.[138]

Anti-essentialism and realism are two reasons why an HPC theory would adequately account for a theory of intelligence. Another important reason concerns multilevel analysis. Psychological phenomena are characterized by properties spread across different levels of organization—ranging from behaviors to lower-level mechanisms. Cognitive systems, in turn, are frequently described as hierarchically organized (see Bechtel, 1994; Craver, 2002, 2015). This applies to psychometric intelligence as well: several models try to explain how different datasets, concerning different variables (e.g., neurobiological, cognitive, and psychometric ones), could match with each other, maintaining the validity of the related scientific enterprises. In particular, these models seek for a plausible organization of different variables in relation to the *g* factor (see Chapter 7 for a detailed discussion). The *g* factor may be understood as a causal link that holds variables together. In order to make sense of this relationship among variables, the HPC theory is especially well suited. Indeed, this theory acknowledges a hierarchical relationship among a phenomenon (a property cluster) and its parts (e.g., properties and mechanisms), understood as causally related to each other.

Albeit attractive, Boyd's theoretical solution has been evaluated as explanatorily weak: indeed, the HPC theory is, on its own, inadequate in identifying those mechanisms which matter to pinpoint a property cluster (Boyd, 1991; Craver, 2009; Wilson et al., 2007). This vagueness will be relevant later in the discussion. In the next paragraph, I delineate an HPC model of intelligence to account for the intuition encompassed in the PSY-GEN model of intelligence.

[137] See Samuels (2007) and Khalidi (2015) for different positions on this point.
[138] This is also one of the reasons why HPC have taken root in the philosophy of psychiatry: behaviors and clinical symptoms are conceived as observable properties occurring over more fundamental phenomena, i.e., underlying causal patterns.

3. A Kind Theory of Intelligence

Consistently with what has been said above, an HPC theory of intelligence might have advantages over other views of natural kinds concerning traditional essentialism and multilevel analysis. Another potential merit is that it might allow the PSY-GEN view of intelligence to avoid folk-essentialism. In this paragraph, I investigate whether an HPC theory of intelligence can avoid folk essentialism and whether it fits with some widespread intuitions about human cognition, concerning the relationship between IQ and cognitive processes.

Let us consider general cognitive disability as the "negative pole" on a scale of values. Psychometrically, it is described as a variable conceptually comparable to normal IQ level—an IQ below the average, roughly placed under 70 points. In this sense, cognitive disability is a quantitative feature. From a genetic point of view, this is a phenotypic trait related to an additive genetic influence. In this respect, the only difference between a normal and a low IQ consists of carrying different alleles—and, of course, being subject to different environmental influences which, as we have seen, are generally less considered than the genetic ones. According to the reasoning in §1.1, it is hard to apply an analysis in terms of natural kinds.

Consider now two individuals obtaining the same low IQ score. Intuitively, even though they have an identical IQ, we rarely tend to claim that their cognitive profiles are identical or that they could derive from similar causal mechanisms. Thus, if we try to explain *why* an individual shows a low IQ, there are several suitable explanations that refer to different properties. On a cognitive level, two similar individuals could have widely different abilities useful in solving subtests belonging to different categories—mathematical, logical, linguistic, and so on.[139] On an etiological level, two similar IQs could depend on many factors combined in several ways—inheritance, trauma, education, and social aspects.

Briefly, two people affected by cognitive disability would presumably obtain a low IQ score for different reasons. Similarly, two people that are "successful" in IQ tests are similar with respect to IQ, so one could put them in the same category—but their similarities might derive from widely different causal mechanisms, related to different cognitive processes. If this is sound, then all the conditions are in place for providing an HPC theory of intelligence.

The HPC model I propose has five main characteristics. First, IQ depends on various

[139] Tests generally contain different subtests requiring different skills (mathematical, linguistic, logical, etc.). As is said in §1, test performances are statistically intercorrelated. However, the degree of these correlations varies largely (Stankov, 2012). This is consistent with the idea that different cognitive profiles reside under similar IQ scores in different individuals. Anyway, the "positive manifolds do not automatically reveal their meanings. For example, it is quite possible to obtain a positive manifold due to an overlap of task demands, rather than due to the influence of a general ability" (Kray & Frensch, 2012, p. 186). See Chapter 1.

cognitive processes combined in different ways in distinct individuals. Second, IQ is a behavioral measure assessed in an experimental context, rather than a cognitive phenomenon itself. Third, such cognitive processes are not individually necessary and jointly sufficient, for a given individual, to get a specific IQ level. This saves us from traditional essentialism. Fourth, the HPC model maintains the core ideas of PSY-GEN, according to which intelligence is a general cognitive ability and *g* plays a role. Fifth, the model is hierarchical in the manner proposed by psychometricians: domain-specific cognitive abilities (broad factors) are distinguished by *g*. Let us look at some details shown graphically in Figure 6.1.

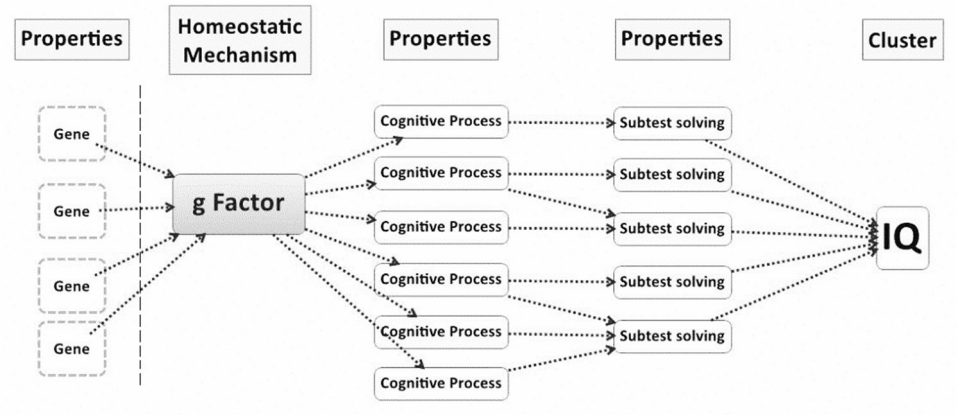


Figure 6.1: The homeostatic property cluster model of intelligence. IQ represents the cluster. The property set entails subtest solving, cognitive processes and biological correlates. The *g* factor is conceived as the homeostatic mechanism.

- Cluster: This is represented by the IQ, which is useful in categorizing people. The IQ score is conceived as a behavioral output, and it represents the cluster. It may change or not among individuals, but every behavior related to IQ tests belongs to the same phenomenon. This is consistent with the intuition that intelligence is a unified thing.

- Properties: For a start, the property set includes performances related to subtests. I call them 'surface properties' because they are useful in categorizing individuals and in analyzing their cognitive profile starting from subtest solving. Other properties reside on deeper levels—e.g., those cognitive processes required to solve specific tasks. These processes presumably work differently in distinct individuals. Moving down a few steps, it is possible to suppose that biological factors, too, could represent the cluster's properties—e.g., some specific genes could be stably associated with an individual IQ. At once, one can suppose that the property set is

joined by other biological correlates.[140] Which properties are relevant in categorizing people depends on the scientific target: surface properties would be more relevant in the psychometric context; cognitive processes would satisfy a psychological inquiry; biological properties are useful for biological analyses.

- Homeostatic mechanism: This is the *g* factor, conceived as the causal source of the property cluster. For an HPC theory, it is necessary to individuate a homeostatic mechanism in order to be sure that a cluster is really a cluster and not a mere set of properties. This assumes that the properties have not been associated with each other on artificial grounds, merely as a result of our predilections to lump certain properties together. The model assumes that IQ differences among individuals reflect differences in the functioning of *g*. So, it accounts for both the flexibility and the stability of the intelligent behavior. The *g* factor is a mechanism flexible enough to allow for even considerable variations among individuals without appealing to different mechanisms for different categories (e.g., cognitive disability, normality, and genius). Nonetheless, the mechanism is stable enough to serve methodological purposes—e.g., generalization, prediction, and identifying categories of similar individuals *even if they do not exhibit the same surface properties*.[141] In sum, *g* allows us to explain why some properties are co-instantiated both in similar individuals and in widely different individuals, even analyzing the latter cases under the same phenomenon. In this sense, people with low and high IQ scores are dissimilar to each other (at least with respect to the surface and the cognitive properties), but their differences do not reflect two (or more) distinct psychobiological mechanisms, but rather variations within a range of *g* functioning.

What is the role of the term 'intelligence'? It is *the kind itself*, as distinct from the cluster. The cluster is a set of properties that are related for non-conventional reasons. Conversely, the kind represents intelligent behavior as a very broad phenomenon. Consistent with the geneticists' conception of intelligence as a phenotypical trait, all human beings—and not only them—participate in the kind.[142]

To summarize, if we assume that intelligence is a natural kind, and if we assume that *g* exists, the HPC model sounds more promising than the quantitative view insofar as it accounts for the intuitions sketched above. Those intuitions remind us of a general trend

[140] In Figure 6.1, there is a dotted line between deeper properties and the homeostatic mechanism because we still do not know what there is between the two—i.e., which biological correlates matter.
[141] This is a central point. We might assume a standpoint which is less metaphysically loaded than the one I assumed so far by thinking of natural kinds from an epistemological perspective (for a discussion, see Magnus, 2012). In this respect, thinking of intelligence as a natural kind means thinking of it as an answer to methodological issues.
[142] One may try to extend the question to other species. It worth noticing that some authors (e.g., Jensen, 1980; Burkart et al., 2017) argue that *g* is not solely limited to the human species.

in cognitive science, according to which cognition does not involve a general mental ability, but rather it consists of different domain-specific abilities (see Chapter 7).

If the model sounds convincing, then intelligence is not a quantitative trait itself; what is quantitative is the IQ variation within populations, as I discussed in the previous chapters. Ultimately, what is missing in the quantitative approach is a clear distinction between intelligence, IQ, and *g*, which frequently collapse upon each other. IQ is a variable: one can measure it because it changes quantitatively within populations. This does not imply that intelligence is a unitary phenotypic trait describable by quantitative genetics (see Chapter 5).

Furthermore, if the model is sound, then a traditional essentialist interpretation of intelligence seems untenable, insofar as one assumes that: 1) HPC theories describe psychological phenomena better than classical kind conceptions; and 2) HPC theories can avoid essentialism about properties—that is, none of them is individually necessary nor jointly sufficient to ascribing an individual to a kind.

In the next paragraph, I discuss the other side of the coin. I see something wrong in the temptation to lump several mental skills into a single comprehensive phenomenon. Thinking of intelligence as a set of many cognitive processes leads us to identify a set of widely heterogeneous behaviors without any empirical commitment to the neurocognitive mechanisms involved. So, I shall highlight a concern about the HPC model and suggest that it is unnecessary to consider intelligence as a kind in any sense.

4. Is Intelligence Really a Kind?

The HPC model seems to be suitable to account for properties according to an anti-essentialist perspective, especially for subtest solving and cognitive processes. But is the model good enough to say something about deeper, biological properties and about the homeostatic mechanism? Is the genetic influence on *g* quantitative or not? Is there any gene that is necessary and/or sufficient for a specific *g* functioning? These are empirical questions that we need to address.

By adopting Boyd's theory, such questions are likely doomed to remain unsolved because of the vagueness of the theory itself. As is said in §2.3, this theory is inadequate in identifying those mechanisms which matter to pinpoint a property cluster. According to Boyd, the mechanism may or may not be underlying; it might derive from a single cause or not, involving several phenomena; it might be internal or external.[143] About psychopathology, Kendler et al. (2011, p. 1149) ask: which of the diversity of possible causal processes should we emphasize when we construct our nosology? This is not solely

[143] For instance, phenotypic variability in biological species might be related to both internal (e.g., genetic variability and developmental factors) and external mechanisms (e.g., natural selection and environmental influences). However, Boyd's theory does neither explain what the relevant mechanisms are nor how different mechanisms interact with each other to constitute the property cluster.

a conceptual issue. Rather, it should be assigned to a strictly empirical research: it depends on the case under examination and on which is the best scientific discipline to solve the puzzle.

So, it is an empirical question whether *g* might be a good candidate for the role of the homeostatic mechanism. Psychometrics cannot deal with this alone. As is said in §1, it is generally assumed that *g* arises as a stable phenomenon. Nonetheless, we need external validators to meet neurobiology: correlation matrices are not causally self-explanatory (see Chapter 1). If cognitive sciences deserve to play a role in this inquiry, then we should consider the disagreement about *g* as a genuine cognitive phenomenon. Indeed, if the opponents of *g* are right, then the homeostatic mechanisms needed to ground an HPC theory of intelligence would be lacking: thus, general intelligence would not be a cluster, but merely a *set of properties*.

If this is the case, general intelligence would lose any epistemic advantage over other theoretical constructs in terms of our understanding of human cognition, not being a rewarding posit to support epistemically successful science. Ontologically speaking, instead, the PSY-GEN model would be unable to carve nature at its joints—unless HPC theory is an inadequate manner to model intelligence. In both the epistemic and the ontological respects, intelligence would not be a natural kind. This problem looks like the one that Griffiths (2004) detected for emotions:

> "The question about emotion […] is not whether we can give a single 'account' of the category in the sense of a philosophical analysis of the emotion concept, but whether the category thus singled out is a productive object of scientific enquiry" (Griffiths, 2004, p. 904).

If the various instances of the intelligent behavior are different one to another in a relevant sense, then intelligence will not bring any epistemological advantage if compared with specific cognitive abilities. As I show in Chapter 7, empirical data about these questions are still under debate. For the present moment, we might advance a conceptual perspective, figuring a *splitting strategy*: subtyping to different kinds (e.g., specific cognitive processes) something that seems *prima facie* a unified kind (intelligence) (see Figure 6.2).

Test solving seems to be due to many cognitive abilities and it is pointless to categorize them as a single thing. As I discussed in Chapter 5, if intelligence is a heterogeneous bundle of distinct phenotypic traits, any statistical analysis which tries to relate IQ and genetic variation will be inevitably doomed to find spurious, weak, and unreliable correlations. Accordingly, it seems empirically more productive to search for genes that act on specific cognitive components rather than genes that act on general intelligence; indeed, the targets of genetic and environmental influences are likely those cognitive abilities. Some efforts have been made in this direction, but the overall trend is to maintain the hierarchical model of general intelligence, including *g* as the main target of genetic influences (see Plomin et al., 2013, p. 217).

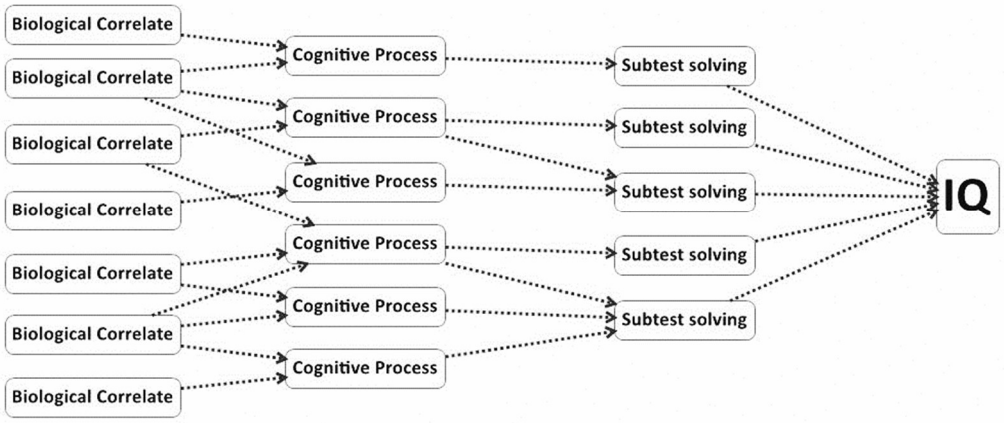


Figure 6.2: A splitting strategy to account for intelligence. The figure shows the relationship between biological correlates, cognitive processes, tests abilities, and IQ.

By analogy, it would be as if we were looking for genes for the "escape velocity from predators" instead of genes for other phenotypical traits that matter somehow for speed. Of course, one can compare quantitatively speed (some individuals are faster than others), but it does not follow that speed is a quantitative trait. Van der Maas and colleagues (2014, p. 14) assume a similar position, noting that

"if *g* is not a causal source of the positive manifold, the search for a gene or brain area 'for *g*' will be fruitless. […] The comparison with health is instructive. There are no specific genes 'for health', and health has no specific location in the body. Note that this line of reasoning does not apply to genetic and brain research on *components* of intelligence (for instance, working memory) as these components often do have a realistic reflective interpretation. Working memory capacity may very well be based on specific and independently identifiable brain processes, even if *g* is not".

The attempt to make sense of specific cognitive components as natural kinds is a matter of future research.

Conclusion: To Lump or to Split?

In this chapter, I adopted the natural kind theory for analyzing the ontological status of general intelligence. The HPC theory has been taken as the most promising conceptualization for making sense of several intuitions about the natural-kindness of psychological and biological phenomena related to the human intelligent behavior. If my analysis sounds, a quantitative view of intelligence seems to be unconvincing in several respects.

However, the HPC model is just a conceptual analysis of the problem: by focusing on empirical issues as assessed by neurobiological sciences, one can find further reasons

to think that a quantitate view of intelligence is untenable. For instance, environmental and multilevel biological influences seem not separable from one another. These data, that are less controversial than the ones related to the *g* factor, point in the direction of an interactive model rather than a purely additive one. In other words, the theory of biological intelligence meets in a slippery way both biological and cognitive data.

It is important to notice that the proposed HPC model of intelligence leaves several questions open about the nature of the psychobiological *g*, especially concerning its causal aspects. In any case, the HPC model has the merit of making clearer the relationship between the IQ, cognitive processes, and the *g* factor. The advocates of the PSY-GEN model rarely engage in ontological analyses of this sort. In such a way, they often bounce from instrumentalist to realist positions about general intelligence.

An alternative framework that subtypes intelligence into cognitive processes satisfies intuitions concerning surface properties. But it does more than this: it rejects explanations based on *g*. Moreover, it seems capable of avoiding that sort of essentialism that follows easily from the quantitative perspective.

The recent endeavor to analyze psychological phenomena as natural kinds—and, then, as Boydian kinds—relies on the expectation that widespread concepts like intelligence must reflect some feature of the outside world. If we need a naturalistic theory of natural kinds, we should admit property clusters suitable for tracing the causal structure of the world. General intelligence does not seem to fulfil this requirement.

In Chapter 7, I consider ontological issues about general intelligence by considering the debate going on in cognitive sciences. Is intelligence a general cognitive phenomenon or a bundle of distinct cognitive processes? What cognitive processes, if any, would better explain the nature of *g*?

Chapter 7.
Intelligence and Cognition

In contrast with what the PSY-GEN theory of intelligence states, IQ and general intelligence cannot be understood as quantitative traits. In the previous chapters, I argued for this conclusion from two points of view. The first one concerns genetic causation and the definition of proper phenotypic traits in genetics research: if the target of the genetic effects are not the IQ, or general intelligence, but rather lower-level cognitive and biological phenotypes, one cannot say that IQ and general intelligence are phenotypic traits in a strict sense. Accordingly, IQ should be understood as a behavioral outcome of several cognitive phenotypes plus a quantification over individual differences in respect with those cognitive traits (that is, the IQ is a quantifiable-in-population trait). The second reason concerns the natural kinds theory: general intelligence, as it is conceptualized by the PSY-GEN approach, rather than being a quantitative trait is better describable as a homeostatic property cluster related to several biological, cognitive, and behavioral properties plus an underlying causal mechanism (namely, the *g* factor) that connects all these multi-leveled properties.

However, I raised doubts about this kind-like interpretation of intelligence: the possibility that general intelligence consists of a cluster of causally-related properties (an HPC) depends on whether a supporting homeostatic mechanism does exist—and this is an empirical question. If no such thing as a homeostatic mechanism exists for a given complex phenomenon, then the associated properties have no causal connection with each other. Hence, the phenomenon is just a property set and, then, there is no real reason to be a realist about the phenomenon itself—not in the way natural kinds theory prescribes, however.

In the case of general intelligence, the question is whether the required causal mechanism (i.e., the *g* factor) does exist. As I highlighted, a heated debate about *g* has taken place in the last decades. In this chapter, I analyze how and why scholars have questioned the reliability of *g* as a genuine neurobiological phenomenon. Is there any empirical evidence of the existence of *g*? What cognitive process, if any, would better explain the nature of *g*, or would better fit with its description generally provided by the PSY-GEN approach? Is there any single biological mechanism which accounts for the intelligent behavior?

By answering these questions, one may clarify whether intelligence is a general cognitive phenomenon or not. If not, the only feasible manner to look at intelligence would be as an arising phenomenon, or at most an instrumental concept, to denote human cognition in a broad sense. No ontological commitment would then be required for that thing

called 'human intelligence'. In order to argue for this conclusion, I provide a developmental explanation of the psychometric *g* that does not include any general ability like the psychobiological *g*.

1. How General Is It?

In 2002 Robert Sternberg and Elena Grigorenko, two leading scholars in intelligence research, edited a companion entitled *The General Factor of Intelligence: How General Is It?*. Numerous scholars, from as many scientific fields, confronted with each other about the generality of human intelligence. In some cases, such question implied asking whether cognitive sciences have confirmed or disconfirmed the existence of the *g* factor, finally making room for neuroscientific methods and conceptions within this secular debate. Unfortunately, no definitive conclusion had been reached. In fact, the authors did not really discuss with each other with the aim of achieving a shared conclusion about the subject manner, but rather they brought grist to their respective mill. For instance, Sternberg himself, far from being an impartial editor, wrote in the conclusion:

> "The time has come to move beyond conventional theories of intelligence. In this chapter I have provided data suggesting that conventional theories and tests of intelligence are incomplete. The general factor is an artifact of limitations in populations of individuals tested, types of materials with which they are tested, and types of methods used in testing. […] I have proposed a theory of successful intelligence and its development that fares well in construct validations, whether one tests in the laboratory, in schools, or in the workplace. The greatest obstacle to our moving on is in vested interests, both in academia and in the world of tests, where testing companies are doing well financially with existing tests. We now have ways to move beyond conventional notions of intelligence; we need only the will" (Sternberg, 2002, p. 472).

.

Unsurprisingly, the debate over the generality of intelligence has continued across both psychological and biological journals.

Another companion, entitled *Cognition and Intelligence*, has been published in 2005 (edited by Sternberg and Pretz) to bridge the gap between psychometrics and cognitive sciences. Unifying models have been proposed for explaining human intelligence from the perspective of cognitive psychology, now welcome (at least for some authors) in the psychometric domain. Also thanks to this type of enterprises, something seems to have changed in the last decades. At first sight, one can see that the psychometric approach to intelligence is no longer the only one available: cognitive approaches tend to be better represented in intelligence research.

The question is: What cognitive scientists really do within intelligence research? In few exceptional cases, they gave birth to innovative theories of intelligence, although their relationship with the psychometric models and purposes is still unclear (see §2.2). More often, cognitive methodologies and conceptions have been placed at the service of the classic problem of the generality of intelligence: What cognitive science can say about

that? Is there any neuroscientific evidence of the existence of the *g* factor? So, cognitive sciences have been employed for making sense of the PSY-GEN model of intelligence. This is especially true in those cases in which correlations are sought between neural or cognitive variables and *g* or the IQ (see §3). The problem of general intelligence remains open to such an extent that Sternberg writes:

> "If there is one finding in psychology that has been replicated more than any other, it may be the general (g) factor that results from factor analyses of large numbers of psychometric tests of intelligence. […] Many researchers accept some version of Carroll's (1993) taxonomy of abilities as representing the relationship between the g factor and more specific abilities. The remaining question is just how general the general factor is (Sternberg & Grigorenko, 2002). Some theorists, like Jensen (1998), believe that g and the various subfactors under it account for most or all of intelligence; others, like Sternberg (1985) […] and Gardner (1983), believe that general ability only scratches the surface of the range of human intellectual abilities" (Sternberg, 2013, p. 177).

In the following paragraphs, I delineate the debate about the generality of intelligence and its contemporary interpretations. I also list contemporary theories of intelligence and I highlight their ontological commitment to the *g* factor. Before proceeding, however, some remarks should be made.

In Chapter 1, I briefly explained the basic lines of the historical disagreement opposing two viewpoints about intelligence: on one side, the single-factor (or, let us say, generalist) theories, according to which intelligence is a general cognitive ability; on the other side, the multiple-factor (or, let us say, anti-generalist) theories, where intelligence is composed of several cognitive abilities, from which general intelligence arises as an abstract entity. It might seem easy to understand what distinguish these positions, but a careful examination reveals that it is not so easy. What does it really mean that intelligence *is* a single thing, i.e., a general cognitive ability? And what does it mean that intelligence *is composed of* cognitive abilities? There are two ways for framing these questions: the first one concerns individual differences, while the second one concerns causality and neurocognitive mechanisms. Let us see them one by one.

The first way to understand these questions pertains to individual differences within populations. The single-factor view states that individual differences in IQ tests (read: the population variance) are better accounted for by a general factor of intelligence (i.e., the psychometric *g*). By contrast, the multiple-factor view holds that individual differences in IQ tests can be interpreted in several ways, among which appealing to the *g* factor is not by principle privileged.

The second way concerns the causal relationship between intelligent behaviors and some other process. The generalist view holds that intelligent behavior, with its heterogeneity, is ultimately related to a single general underlying mechanism (i.e., the psychobiological *g*), which causally affects a cascade of other aspects (i.e., cognitive

processes and behavioral properties such as IQ).[144] By contrast, the anti-generalist view holds that the intelligent behavior is so heterogeneous because it is related to different aspects of humans' neurocognitive systems. These aspects are not necessarily independent of each other—rather, they likely interact—but they are autonomous to extent. The term 'general intelligence', accordingly, is nothing but an arising phenomenon, or an instrumental concept, to denote human cognition.

The differences between these two ways of interpreting the debate about the *g* factor rest upon the historical separation between psychometrics (or, differential psychology) and cognitive psychology. As Anderson (2005, p. 276) notices, cognitive psychologists and cognitive scientists use the word 'intelligence' to talk about the property of the entire human cognitive system. So, it could be argued that cognition and intelligence are synonymous and that all the work in cognitive psychology is about the psychology of intelligence. Nevertheless, as Jensen (1998) has stated, this broad meaning of the term intelligence misses the focus that has interested the traditional researchers of intelligence, namely, the basis of individual intellectual differences.

Do the terms cognition and intelligence refer to the same *thing*, then? In a sense, they do not. On the one hand, there is the study of intelligence that could in principle be conducted by studying single individuals—that is, discovering the universal structure of an idealized cognitive mind. This is, as Jensen and Anderson argue, the proper focus of cognitive psychology, a nomothetic enterprise. On the other hand, there is the study of individual differences and of the *g* factor. Thus, *g* is conceptually and methodologically quite a different thing than human cognition.

The attempt to bridge the gap between intelligence and cognition seems to be, at present, inevitable. However, several precautions must be taken. Something that makes things tricky within this enterprise is the theoretical distance between differential psychology and cognitive psychology. As Naglieri and Das (2002) notice, contemporary thinking tends to focus on a sort of functional segmentation of the neurocognitive architecture. This applies, for instance, to modular and dual-process theories (see e.g., Evans & Frankish, 2008), but also to cognitive-oriented theories of intelligence (see §2.2). The basic idea in cognitive psychology is, roughly speaking, that the brain consists of many modules which process information independently of each other.[145] According to Naglieri and Das (2002, p. 57), although the brain can be seen as working as a whole, it cannot be conceived to have one general function that is identified with intelligence.

---

[144] As Kray & Frensch (2002, pp. 184-186) say, one can think of that mechanism as *the only* source of intelligent behavior or as a source that *affects* all intelligent behavior. See §3.1 for more details.

[145] I use here a broad definition of module, provided by Ramus (2006). A module is a specific information-processing function (cognitive level), together with its neural substrate (a specialized brain structure on the anatomical level). Properties like innateness, and the modules' evolutionary history, should be determined empirically. For instance, the visual word-form area is a module (that processes sequences of letters as part of the reading system), even though it has not evolved to read, and even if it turns out to process other stimuli than sequences of letters. I do not assume that the neural substrate must be a single localized area, rather than a distributed network.

Then, psychometrics and cognitive psychology can be understood as two scientific inquiries separated by a profound theoretical difference. It is likely because of this mismatch that the initial hopes of a successful integration of information processing theories and intelligence theories have not yet been realized (see Hunt, 2011).

"Many researchers have proposed various models by which individual differences in cognitive abilities can be seen as parameters of information processing [see §3]. As yet, these models are mostly untested hypotheses and are not yet ready to be applied in everyday practice (Floyd & Kranzler, 2012). However, ultimately, we hope to have a consistent account of the philosophy of mind, neuroscience, universal cognitive processes, and individual differences in intelligence" (Schneider & Flanagan, 2015, p. 337).

Perhaps, it is because of this mismatch that questions about (individual differences in) intelligence and cognition do frequently overlap, leading to a constant shift from psychometric to psychobiological statements, from statistical to causal analyses, from methodological to ontological problems, from realist to instrumentalist viewpoints. As I mentioned, I am mostly interested in the ontological side of the debate about intelligence. The reason is two-folded: first, one cannot address the nature of intelligence by solely relying upon psychometrics; second, psychobiological *g* is something more than psychometric *g* (the former concerns how cognition works, the latter pertains to factor analysis).

The aim of this chapter is to compare findings from cognitive sciences with the PSYGEN model of intelligence. Right now, no one really knows what the *g* factor is. Several attempts have been made to reveal the nature of intelligence and several authors hold that we cannot discard the *g* factor as a valuable scientific posit. By contrast, one might think that psychometricians need to ask whether their operational definition of *g* is tightly articulated and defensible. As Plucker and Shelton (2015) notice,

"this necessitates exploring additional questions: Are there confounds that could be driving observed results? Is *g* a distinct underlying property of mental abilities, something that would allow one to search for the genetic correlates of a common source, or is it an emergent property stemming from those mental abilities such that one would be establishing the correlates of many different sources that together comprise a *g* factor? And perhaps most importantly, is the construct of *g* itself too complex for establishing robust results?" (Plucker & Shelton, 2015, p. S23).

1.1. Two types of theories and their ontological commitment

The original psychometric *g* was a summary index of a correlation matrix, representing what cognitive tests have in common and explaining a certain portion of their variance (see Chapter 1). As Van der Maas et al. (2014) explain, factor analysis reflects a latent variable in psychometric datasets. Often, a general factor of intelligence is hypothesized as representing the common cause "out there" that we "detect" using

factor analysis, and that should have an independently ascertainable identity in the form of, say, a variable defined on some biological substrate (see §3). This is, however, a misleading employment of factor analysis:

"Behavioral factor analyses do not provide an unambiguous model of the underlying cognitive architecture, as the factors themselves are inaccessible, being measured indirectly by estimating linear components from correlations between the performance measures of different tests. Thus, for a given set of behavioral correlations, there are many factor solutions of varying degrees of complexity, all of which are equally able to account for the data. This ambiguity is typically resolved by selecting a simple and interpretable factor solution. However, interpretability does not necessarily equate to biological reality" (Hampshire et al., 2012, p. 1225).

"[Factor analysis] is indeterminate (the same data can generate literally an infinite number of 'solutions') and how any of them generated by behavioral data relate to any of them generated by biological data is a wide-open question. […] There is [no] one best factor solution waiting to be discovered. Rather all solutions can be appropriate in some circumstances and not in others, and evaluating any solution is a matter of judgment […]. Because this is true of factor analysis in both brain and behavioral data, the fact of finding associations between one kind of solution in brain and behavioral data doesn't necessarily say anything about whether or not there may be similar associations between another kind of solution in the same brain and behavioral data. In fact, because you're working with the same two covariance matrices either way, such similarity of associations is effectively inevitable" (Haier et al., 2014, pp. 327-328).[146]

Again, the temptation to infer substantial ontological claims from statistical analyses seems to represent the very problem afflicting the PSY-GEN approach to intelligence. In fact, in the light of the positive manifold originally detected by Charles Spearman, several psychologists have accepted the existence of an underlying general mental ability capable of explaining IQ individual differences and, somehow, the nature of human intelligence. This is the reason why, at some point, the debate about intelligence has landed in biological sciences.

We can distinguish between two types of theories of intelligence as regard as their acceptance of the existence of a general factor. The origins of single-factor theories can be traced in Spearman's thought, which is the basis, to a lesser or greater extent, of the psychometric tradition, and of tests practice, too. For instance, it is at the base of Wechsler's widespread tests:

"Wechsler referred to Spearman's discovery of *g* as 'one of the great discoveries of psychology' (Wechsler 1944, p. 6). Wechsler's viewpoint as to the importance of *g* did not change over time. Kaufman (2009, p.

[146] Even though the two papers address similar points as regard as factor analysis, Haier et al. (2014) criticize Hampshire et al. (2012) as being victim of the same mistake they pointed out. Hampshire et al. used factor analysis for addressing a non-generalist view of intelligence. According to Haier et al., Hampshire's argument does not hold because, from factor analysis, one cannot say much of anything definitive about *g* or no *g* or its "location" in the brain beyond a demonstration that brain and behavioral data can be modeled in a similar way.

45) writes that when Wechsler visited him in 1975, he told Kaufman's students that 'nothing is more important than *g* for understanding intelligence. Global ability is *the* ability that underlies my IQ tests'." (Benisz et al., 2015, p. 166).

Both WAIS and WISC measure intelligence as a unified phenomenon and produce a single IQ score which is supposed to coincide with general intelligence. This is the reason why Wechsler's tests are called 'single-factor tests' (see Chapter 1).

Within this generalist framework, the psychobiological reality of *g* has been stated by minimizing the role of domain-specific cognitive abilities. Spearman, however, accounted for the role of specific cognitive abilities by developing a two-factors theory which concerns both *g*, which intervenes in every task, and *s*, which intervenes in specific cognitive tasks. However, the emphasis on *g* was already present in Spearman's work:

"Spearman claimed that *g* is a single mental capability measured by all intelligence tests, and that it is some form of generalized mental energy. Specific abilities are capabilities uniquely measured by a particular mental test, for example, mathematical computation. Spearman was interested primarily in what is common among various types of intellectual abilities, rather than in what makes each one unique. He believed that specific abilities do not capture the essence of intelligence and instead proposed that important differences in people's mental test scores are due to just one intellectual capability, mental energy" (see Cianciolo & Sternberg, 2004, p. 3).

Over time the focus has turned more on the general factor than on specific factors, leading to the view that specific abilities play a secondary role insofar as they are strongly influenced by *g*.

Multiple-factor theories are, in many respects, quite different. They originated from the work of Thomson and Thurstone (see Chapter 1), who believed that no such thing as a general intelligence really exists. Thurstone argued that *g* was a statistical artifact resulting from the mathematical procedures used to study it. Thurstone founded with factor analysis seven Primary Mental Abilities (PMA): word fluency, verbal comprehension, spatial visualization, number facility, associative memory, reasoning and perceptual speed. These factors are neither general across all tests nor specific to each test. In other words, those abilities are involved in many tasks (see Schneider & Flanagan, 2015).

The advocates of the multiple-factor view do not accept the reliability of *g* or underestimate its explanatory power. Rather, they focus on the so-called group-factors, shared by some tests only within factor analysis. In multiple-factors tests, the variety of the items may depend on the personal idea one has in mind about intelligence; generally, researchers "extract" several factors of intelligence. Here, *g* is supposed to be an emergent phenomenon whose ontological reality it is not worth assuming.

In sum, what distinguishes the two types of theories is the ontological commitment about general intelligence or, more specifically, about the *g* factor as the underlying mechanism which causes the performance to IQ tests and the cognitive individual

differences. In any generalist theory, *g* represents somehow the real explanatory entity and the depositary of the causal efficacy on intelligent behaviors. At the same time, lower-level causal effects (e.g., genetic causality) supposedly act on *g*, rather than on cognitive processes.[147]

That said, the offer of intelligence theories available on the market is quite heterogeneous. For instance, not every generalist theory overlooks specific cognitive abilities, and not every anti-generalist theory discards the existence of *g* or its hypothetical explanatory power. In §2, I deepen the most important theories of intelligence, their relationship with test practice and with the *g* factor. Indeed, Spearman's and Thurstone's theory have often served as general frameworks to emphasize the generality or the non-generality of intelligence. However, more sophisticated models of human intelligence have been developed over time by both psychometricians and cognitive scientists.

2. Contemporary Theories of Intelligence: An Overview

In Table 7.1 different theories of intelligence are compared according to: a) their inclusion or exclusion of *g*; b) their importance for psychometric practices (for some reviews, see Cianciolo & Sternberg, 2004; Kaufman et al., 2013; Ortiz, 2015; Schneider & Flanagan, 2015). I shall divide these theories, with the help of Kaufman et al. (2013), into three categories according to their relationship with tests:

- Class 1 includes theories that are closely tied to the measurement of intelligence. First, if a theory includes *g*, it is also involved, in many cases, in tests practices. Second, Spearman's *g*, the CHC Theory and the PASS Model (Planning, Attention-Arousal, Simultaneous and Successive) represent the theoretical foundation for nearly all commercial tests of intelligence. The latter two theories, however, are more recent and related to cognitive sciences than Spearman's one. Tests based on the CHC Theory incorporates research on the cognitive mechanisms related to *g*, such as working memory—but it is, in some interpretation, neutral about the real existence of *g* (see §2.2). The PASS model derives from Luria's theory of intelligence. The development of related testing tools is explicitly tied to neuroscience findings.

- Class 2 includes theories that have been elaborated to respond to what is missing in traditional intelligence tests. The theories of Multiple Intelligence (Gardner), Suc-

[147] In Chapter 6, I employed the natural kinds theory for explaining this aspect. According to the HPC model of intelligence, *g* is the basilar causal mechanism that connects lower-level properties (i.e., biological correlates) to higher-level (surface) properties (i.e., cognitive processes and behaviors).

cessful Intelligence (Sternberg) and Emotional Intelligence (Goleman) point to additional abilities (e.g., musical, kinesthetic, artistic, practical abilities and creativity) to be treated with the same importance as the standard analytic abilities measured by most tests.

- Class 3 includes theories grounded in recent neuroscientific research, i.e., the Multiple Mechanisms Approach, the Parieto-Frontal Integration, the Minimal Cognitive Architecture Theory and the Dual-Process Theory. According to Kaufman et al. (2013), these theories, although advancing the scientific understanding of human intellectual differences, are less clearly tied to practical applications in terms of intelligence testing.

	Inclusion of *g*	**Tests**	
Two-factor Theory (Spearman, 1900s-1920s)	✓	✓	
Primary Mental Abilities, PMA (Thurstone, 1930s)	×	✓	
Fluid-Crystallized Intelligence, Gf-Gc (Cattell, 1940s)	✓	✓	
Hierarchical Group Factor Theory (Vernon, 1961)	✓	✓	*Class 1*
Extended Gf-Gc (Horn & Cattell, 1966)	*Not Clear*	✓	
Three-Stratum Theory (Carroll, 1993)	✓	✓	
Cattell-Horn-Carroll Theory (1940s – 2000s)	*Not Clear*	✓	
PASS Model (Luria, Naglieri, Das, Kaufman, 1960s - 2000s)	×	✓	
Multiple Intelligences Theory, Gardner (1983)	×	×	
Successful Intelligence Theory, Sternberg (1997)	×	×	*Class 2*
Emotional Intelligence Theory, Goleman (1995)	×	×	
Minimal Cognitive Architecture (Anderson, 2005)	✓	×	
Dual-Process Theory (~2000)	×	×	
Multiple Cognitive Mechanisms Approach (~2000)	*Not Clear*	×	*Class 3*
Parieto-Frontal Integration Theory (Jung & Haier, 2007)	*Not Clear*	×	

Table 7.1: A comparison between theories of intelligence.

As one might see in Table 7.1, the more we approach the contemporaneity, the less the commitment to *g* is. Nevertheless, one might wonder whether cognitive-oriented theories could accept the existence of a g-like phenomenon and whether they are consistent with the purpose of testing intelligence. Several authors believe that the theories included in Class 2 and in Class 3 will be eventually connected to tests practices. For instance, the Minimal Cognitive Architecture Theory tries to make sense of general intelligence in a cognitive-oriented framework, i.e., the modular theory of mind. As another example, the PASS model represents a case in which cognitive sciences have been

employed for the sake of measuring individual differences (e.g., the CAS test). However, I believe that cognitive-oriented theories are quite distant from tests practice because of a precise reason, that is, they did not identify any general cognitive ability capable of summarizing individual performances as a global test score like IQ.

For clarifying these aspects, a detailed analysis of some theories is required.[148] I first discuss (psychometric) hierarchical theories, which generally include the *g* factor. Then, I analyze the more recent theories coming from cognitive sciences, which generally do not include *g*.

2.1. Hierarchical theories

Many of the theories involved in the measurement of intelligence are generally called 'hierarchical theories'. These theories frequently accept both the two aspects of human intelligence (generality and specificity) and propose hierarchical models aimed at describing the relationship between general and specific abilities (see Jensen, 2002; Kray & Frensch, 2002; Schneider & Flanagan, 2015). Hierarchical theories can accept *g* or not and can imply either generalist or "multiple" views of intelligence. In some cases, it is not clear what is the ontological commitment about the *g* factor.

Figure 7.1 shows some important hierarchical theories and their relationship with the original theories of Spearman and Thurstone.

Let us start with those theories that include the *g* factor: beside Spearman's theory (that is at the base of the PSY-GEN model we have considered so far), *g* is an important aspect of both the Gf-Gc Theory and of the Three-Stratum Theory. Generally, these theories, being intimately related to the psychometric tradition, rely on the existence of the positive manifold and on the difficulties on which Thurstone stumbled.[149]

During 1940, Raymond Cattell elaborated a theory involving two factors: fluid ability (Gf) is generally defined as the flexibility of thought and abstract reasoning capability; crystallized ability (Gc) is defined as the accumulation of knowledge and skills. However, definitions of these factors are slightly different from each other:

"Cattell built upon Spearman's *g* to posit *two* kinds of *g:* fluid intelligence (Gf), the ability to solve novel problems by using reasoning—believed by Cattell to be largely a function of biological and neurological

[148] For the sake of the argument, I do not deepen theories included in Class 2, insofar as their aims and conceptions are quite distant from the present discussion. For some details, see the Conclusions.

[149] When Thurstone tested an intellectually heterogeneous group of children with his PMA test, he did not find that the seven PMA were entirely separate; rather, he found evidence of the existence of the *g* factor as well. Thurstone managed a mathematical solution to make sense of those results, and the final version of his theory accounted for the presence of both a general factor and the seven specific abilities. This paved the way for the hierarchical theories, where Spearman's *g* has been rehabilitated as the top of a hierarchy including group-factors (see Plucker, 2016). See Chapter 1.

factors—and crystallized intelligence (Gc), a knowledge-based ability that is highly dependent on education and acculturation" (Kaufman et al 2013, p. 4).

"Cattell, found evidence for two general factors of intelligence. […] What athletic talent is to the body, fluid intelligence is to the brain. It represents the speed, power, efficiency, and overall integrity of the cerebral cortex. In Cattell's (1987) thinking, *gf* is not an ability itself but an influence on many abilities, particularly those abilities that require controlled attention and on-the-spot problem solving. *gc* is acquired knowledge, particularly information stored in declarative memory" (Schneider & Flanagan, 2015, p. 322).

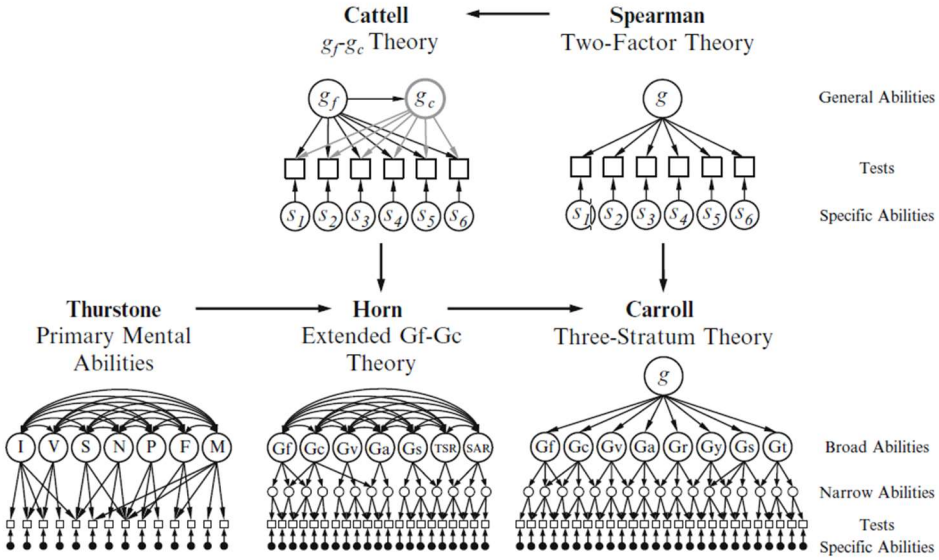


Figure 7.1: The hierarchical structure of some theories of intelligence and their historical and conceptual relationships. From Schneider & Flanagan, 2015, p. 321.

Roughly speaking, fluid intelligence represents a basic biological capacity, a general and heritable potentiality (see Eysenck & Kamin, 1981; Ortiz, 2015). Remarkably, the relationship between fluid and crystallized intelligence is one-way: the former influences the latter, but not the other way around (see Fig. 7.1). The strong genetic influence on fluid intelligence and its generality across different cognitive tasks, make Gf an avatar of Spearman's *g* for all intents and purposes—consider also that Gf and *g* are highly correlated with each other (see Jensen, 1998; Van der Maas et al., 2006).

Carroll's Three-Stratum Theory represents another example that sets *g* at the top of the pyramid.

"Carroll (1993) developed a hierarchical theory based on his in-depth survey of factor-analytic studies composed of three levels or Strata of abilities: (a) Stratum III (General), a Spearman-like *g*, which Carroll considered to be a valid construct based on overwhelming evidence from factor analysis; (b) Stratum II (Broad), composed of eight broad factors, that correspond reasonably closely to Horn's Broad Abilities [see below]; and (c) Stratum I (Narrow), composed of about 70 fairly specific abilities, organized by the broad

factor with which each is most closely associated (many relate to level of mastery, response speed, or rate of learning)" (Kaufman et al., 2013, p. 4).

Conversely, some hierarchical theories do not include *g*, or it is not clear whether they do so. The Extended Gf-Gc Theory (Horn & Cattell, 1966) derives from the work of John Horn, a student of Cattell.[150] It adopts several abilities—quite similar to Thurstone's ones—that are differentially subject to Gf, Gc, and other general factors such as Gv (general visualization), Gs (general speediness or processing speed), Gsm (short-term memory), Glr (long-term retrieval, and Gr (general memory fluency) (see Scheider & Flanagan, 2015, p. 324). According to Kaufman et al. (2013), the diverse broad abilities were treated as equals, not as part of any hierarchy. However, as in the case of the original Gf-Gc Theory, Spearman's *g* sometimes returns to the pitch. As Cianciolo & Sternberg (2004) highlight,

"the best known of these abilities are crystallized ability and fluid ability. […] Relatively more recent depictions of ability hierarchies featuring fluid and crystallized ability show fluid intelligence at the top, equated with Spearman's *g*, and the other abilities below (e.g., Gustafsson, 1984)" (Cianciolo & Sternberg, 2004, p. 7).

The Cattell-Horn-Carroll Theory of Cognitive Ability (CHC) is another case in which the commitment about *g* is debated. It recognizes *g* being atop about 10-16 broad abilities and about 80 narrow abilities. The broad abilities include Gf (fluid intelligence; the ability to solve novel problems), Gq (quantitative knowledge, typically math related), Gc (crystallized intelligence; the breadth and depth of a person's accumulated knowledge of a culture and the ability to use that knowledge to solve problems), Grw (reading and writing), Gsm (short-term memory), Gv (visual processing), Ga (auditory processing), Glr (long-term storage and retrieval), Gs (processing speed), and Gt (decision speed/reaction time). However, the theory maintains the uncertain status of *g*.[151] In many interpretations (e.g., Flanagan, 2007) it is rather neutral about the real existence of *g*:

"In CHC theory, *g* is present but with a deemphasized and uncertain status. This has worked to create an engaged community of scholars, researchers, and practitioners who can talk about cases, discuss research findings, and suggest refinements to the model without having to refight constant battles about the existence of *g*" (Schneider & Flanagan, 2015, p. 327).

[150] Although Cattell's Gf-Gc theory had been proposed in 1941, it was not tested directly until the 1960s with the help Horn. Horn and Cattell, sometimes separately, sometimes together, refined the theory and subjected it to critical tests (see Schneider & Flanagan, 2015, p. 320).

[151] This heterogeneity is likely because CHC has been subject to the refinements of several authors. Moreover, it is in several manners influenced by all the previous psychometric theories: "CHC theory has three parents (*gf-gc* theory, extended Gf-Gc theory, and three-stratum theory) and at least two grandparents (two-factor theory and primary mental abilities). In addition, it has two important first cousins once removed (hierarchical Group Factor theory and triadic theory)" (Schneider & Flanagan, 2015, p. 321).

"The CHC model incorporates both the concept of a general intelligence (all of the different aspects of intelligence are considered to be related to a common '*g*' although this aspect is not often emphasized; see Flanagan et al., 2007) and the concept of many different aspects of intelligence. […] The debate about which is "better," one intelligence versus many aspects of intelligence, still goes on" (Kaufman et al., 2013, p. 4).

What is the relationship between hierarchical theories and intelligence tests? The Stanford-Binet Intelligence Scale (fourth edition, see Thorndike et al. 1986) and the revised versions of Wechsler's tests reflect some aspects of hierarchical theories. Even though those tests seem to agree on the importance of specific cognitive abilities, they are mostly addressed to *different aspects of the same phenomenon*, that is general intelligence.[152]

Several contemporary multiple-factor tests have been based explicitly or implicitly on the ideas of Cattell, Horn, and Carroll. For instance, the Kaufman Adolescent and Adult Intelligence Test (KAIT; Kaufman & Kaufman, 1993) is partially based on the Gf-Gc theory (see Schneider & Flanagan, 2015, p. 320). Instead, Naglieri and Das's (1997) CAS is the primary exceptional test in which the underlying theories, rather than being psychometric, are cognitive-oriented (see §2.2).

Before proceeding in analyzing those recent theories of intelligence, it is important to notice that hierarchical theories are related to factor analysis. In these theories, the structure of human intelligence (its components and the relationships among them) is mainly suggested by correlational data.

"When a large sample of the population, at any age from childhood to old age, is administered a diverse battery of mental tests the covariance structure forms a hierarchy. At the peak of the hierarchy there is a general factor, typically accounting for about 40% to 50% of the test score variance. Below this, there are correlated group factors of ability. These do not attract full agreement between studies, reflecting the different salads of tests' contents in different batteries. At a still lower level in the hierarchy there are specific abilities, which form correlated but separable aspects of the group factors. […] The psychometric studies suggest that there might be different targets for cognitive or broader information processing studies: general variance, and group and specific factor variance" (Deary 2002, p. 152).

In other words, the judgment about the reliability of *g* does not entail, *prima facie*, cognitive and biological research. Rather, statistical analysis represents the reason for assuming a given factor (e.g., Gc and Gf): a latent variable is suggested (and often justified) by factor analysis. Cognitive theories involve quite a different sort of reasoning.

[152] Indeed, Wechsler believed that different aspects of intelligence can be measured by means of different subtests (see Benisz et al., 2015).

2.2. Cognitive theories

As I mentioned in §1, cognitive sciences have sometimes given birth to innovative theories of intelligence. Here, the starting point is not the will of making sense of *g* and IQ, but rather to achieve a better understanding of human intelligence by starting from a given preexistent theoretical framework devised by cognitive sciences. Let us first discuss two cognitive theories that include a general factor of intelligence in their analysis: the Minimal Cognitive Architecture Theory (MCA) and the Parieto-Frontal Integration Theory (P-FIT).

Minimal Cognitive Architecture has been proposed by Anderson (2005) to reconcile general intelligence research, developmental theories of intelligence, multiple intelligence theories and modular theories of mind. According to Anderson, knowledge is acquired through two different processing routes: 1) problem solving, that comprises two uncorrelated processors (i.e., verbal and spatial processors), is constrained by processing speed—it is this constraint that is the basis of general intelligence and the reason why specific abilities are correlated; 2) acquiring knowledge (information processing modules of three dimensional space, syntactic parsing, phonological encoding, and theory of mind)—it is this route that is linked to cognitive development as these modules undergo developmental changes in cognitive competence across the life span.

In Anderson's view, modular processes can be acquired through extensive practice, but the common features of both acquired and innate modules are that they operate automatically and independently of the first route and thus are not constrained by central processing mechanisms.[153] The modular component of Anderson's cognitive theory should allow a reconciliation between Gardner's Theory and the theory general intelligence by acknowledging the importance of domain-specific abilities as well as a central basic processing mechanism (see Kaufman et al., 2013).

Anderson maintains the validity of *g* by referring to the positive manifold. In particular, *g* is reduced to processing speed: processing speed constraints any type of cognitive ability and does not change over development. This is supposedly the origin of the correlations among abilities. However, such a special mention of *g* can be a weakness for the theory, insofar as it relies on the identification of *g* with the speed variable (see §3).[154]

The P-FIT (Jung & Haier, 2007) holds that the neural basis of intelligence is distributed through the brain, but especially in the parietal and frontal regions. Jung and Haier identified brain region activations based on four stages of information processing: 1) temporal and occipital areas acquire sensory information; 2) sensory data are sent to

[153] The author refers to Karmiloff-Smith's (1992) modular conception.

[154] Moreover, this is also a limit for the modular component of the theory: as Kaufman (2011) notices, Anderson does not propose more than just processing speed as a central mechanism and does not propose any domain-general learning mechanisms underlying route 2, focusing instead on Fodor's definition of modules.

regions in the parietal cortex for integration and abstraction; 3) frontal lobes interact with the parietal areas implicated in the second stage for the sake of selecting the best solution (problem-solving and hypothesis testing); 4) the anterior cingulate inhibits alternative responses. This process is dependent upon the fidelity of underlying white matter necessary to facilitate rapid and error-free transmission of data from posterior to frontal brain regions. Jung & Haier (2007) suggest that individual cognitive differences might be accounted for by individual patterns of P-FIT activations: different combinations of brain area activations can lead to the same levels of cognitive performance.[155]

The role of *g* in this theory is not as clear as the authors probably believe; on the one hand, they would like to resolve the secular issue about what and where in the brain is intelligence; but, on the other hand, they propose neurocognitive explanations of information processing by focusing on *individual differences*. The abstract declares:

"We report a striking consensus suggesting that variations in *a distributed network predict individual differences found in intelligence* and reasoning tasks. […] The P-FIT model includes […] the dorsolateral prefrontal cortex (BAs 6, 9, 10, 45, 46, 47), the inferior (BAs 39, 40) and superior (BA 7) parietal lobule, the anterior cingulate (BA 32), and regions within the temporal (BAs 21, 37) and occipital (BAs 18, 19) lobes. White matter regions (i.e., arcuate fasciculus) are also implicated. […] Overall, we conclude that modern neuroimaging techniques are beginning to articulate a biology of intelligence. *We propose that the P-FIT provides a parsimonious account for many of the empirical observations, to date, which relates individual differences in intelligence test scores to variations in brain structure and function*" (Jung & Haier, 2007, p. 135; emphasis added).

So, the authors look for a neurocognitive explanation of a general mental ability (psychobiological *g*), but their analysis is limited to individual differences (related, at most, to the psychometric *g*). Indeed, no single variable is clearly identified with *g*. Rather, intelligence (or better, individual cognitive differences) is localized down a stream of several brain areas and cognitive functions.[156] As I argue in §3, if a general mental ability is splitted up into different neurocognitive phenomena, then it is pointless to call it "general".

Let us now turn to two cognitive theories do not accept the reliability of a general cognitive ability: the Multiple Cognitive Mechanisms Approach and the PASS Model. The Multiple Cognitive Mechanisms Approach is not a coherent theoretical framework, but rather a set of different proposals that explain general intelligence by appealing to cognitive neurosciences. Generally, these approaches hold that the psychometric *g* may

[155] See my discussion in Chapter 6 about cognitive profiles and IQ tests performance.
[156] Even more alarming is that the authors conflate the question about 'what intelligence is' with the question of 'what is the source of individual differences'. One of the opening sentences (p. 135) refer to Neisser et al. (1996) who, according to Jung and Haier, *defined* intelligence in this way: "Individuals differ from one another in their ability to understand complex ideas, to adapt effectively to the environment, to learn from experience, to engage in various forms of reasoning, to overcome obstacles by taking thought". However, this is a definition of intelligence. This misunderstanding has likely led Jung and Haier to think that defining intelligence consists in correlating individual behavioral differences with individual biological differences.

not be comprised of a single cognitive mechanism (i.e., psychobiological *g*) but instead is supported by multiple, interacting mechanisms that become associated with each other throughout the course of development (see e.g., Conway et al., 2011; Hampshire et al., 2012; Kaufman et al., 2009; Kaufman et al., 2013; Van der Maas et al., 2006). Working memory, processing speed, and explicit associative learning represent the three cognitive mechanisms that have received the most attention. As I show in §3, since none of these cognitive processes account for the totality of the g-related variance, they cannot be understood as avatars of the psychobiological *g*. This is the reason why these theories are not committed to the existence of a general cognitive ability such as *g*.

Finally, the most renowned cognitive theory in intelligence research is likely the PASS Model. As I mentioned, it historically derives from Luria's works (1966, 1970, 1973), which focuses on different functional units: attention, simultaneous processing, successive (or sequential) processing and integration. Intelligence refers here to a subset of psychological processes.

> "Luria […] maintained that the brain is complex and that no part of it functions without the cooperation of other parts. Thus, Luria viewed the brain as a functional mosaic, meaning that various parts interact in different combinations to apply varying combinations of cognitive processing abilities (Luria, 1973). Thus, Luria contended that there is no area of the brain that functions without input from other areas. Integration of processing abilities is a key principle of brain function within the Lurian framework" (Otero, 2015, p. 194).

Each process, in Luria's view, is not equally involved in every task. For example, reading comprehension may predominately involve one process, while reading decoding can be strongly dominated by another. As another example, basic math calculation may require more of one process, while math-reasoning tasks may require a different cognitive process (see Otero, 2015).

Das, Naglieri, and Kirby (1994) developed and extended Luria's theory with their PASS Model and operationalized PASS constructs with the Cognitive Assessment System (CAS; Naglieri & Das, 1997). The four processes involved represent an interrelated system of functions (cognitive and neuropsychological constructs) that interact with an individual's base of knowledge and skills (Naglieri & Das, 2002). The processes include:

- Planning: A mental activity that provides cognitive control, use of processes, knowledge and skills, intentionality, and self-regulation (executive function). Planning is central to activities aimed at determining how to solve a problem. This includes self-monitoring, impulse control, and generation of solutions as needed;

- Attention: A mental activity that provides focused, selective cognitive activity over time and resistance to distraction (selective, sustained, and shifting attention). Individuals selectively focus on specific stimuli while inhibiting responses to competing stimuli presented over time;

- Simultaneous: A mental activity by which individuals integrate stimuli into interrelated groups (e.g., visual-spatial tasks). The spatial aspect of simultaneous processing includes the perception of stimuli as a whole as in a recognizable geometric design. Simultaneous processing is similarly involved in grammatical statements that demand the integration of words into a whole idea. This integration involves comprehension of word relationships and prepositions;

- Successive: A mental activity by which individuals integrate stimuli in a specific serial order to form a chain-like progression.

Luria's model also represents the theoretical basis of the Kaufman Assessment Battery for Children (K-ABC; Kaufman & Kaufman, 1983). As Kaufman et al. (2013) notice, the key contributions of the K-ABC were, first, to finally produce an IQ test built on theory, and, second, to switch the emphasis from the *content* of the items (verbal vs. nonverbal) to the *process* that children use to solve problems (e.g., sequential vs. simultaneous).

Apparently, there is no room for *g* in the PASS Model. Indeed, it is explicitly proposed as an alternative to the theory of general intelligence: according to its architects, the measurement of *g* is insufficient for examining the special individuals' cognitive problems and that, although *g* has been shown to be a good predictor of achievement for groups of children, a different conceptualization of intelligence can predict achievement more effectively (Naglieri & Das, 2002).[157] It is worth considering that, for the authors, the relationship between tests practice and tests theory has been stagnant until recent years:

"The Wechsler and Binet tests represent a traditional IQ testing technology that rests on the concept of general ability and has not changed since Binet and Simon introduced their first scale in 1905 and Wechsler published his first test in 1939. Despite cosmetic modifications and improved standardization samples the Fourth Edition of the Stanford-Binet and the latest revisions of the Wechsler Scales […] are essentially the same as their respective early versions" (Naglieri & Das, 2002, p. 58).

This brief review about cognitive theories of intelligence leaves several questions open: Why do some cognitive theories adopt the *g* factor, while others not? What does it mean that *g* can be reduced to, for instance, processing speed? What does it mean that the positive manifold arises from the interactions among different processes? To better understand the role played by *g* in cognitive theories, one must turn to the contribution that cognitive sciences have made in the quest for the psychobiological nature of *g*.

[157] I return to this point in the Conclusions.

3. The Psychobiological Nature of *g*

In the last few decades, cognitive scientists and psychometricians have often worked together for making sense of psychometric practices by relating *g* and IQ to neurocognitive constructs. Achieving a comprehensive theoretical framework, thought, is not here the priority, since this research is mainly driven by statistical analyses. Indeed, the PSYGEN model and the hierarchical theories committed to *g* are here taken as reliable.

Several authors have addressed the neurobiological correlates of intelligence from as many different perspectives. The general idea is: by finding a cognitive or a biological phenomenon which accounts for *g*—read: for (most of) the g-related variance—one can say she has found the real foundation of *g*, that is, a single phenomenon capable of explaining the positive manifold *and* in identifying the very nature of general intelligence. A less ambitious purpose, however, consists in looking for neurobiological correlates of *g* to better understand it or to find some sort of process that could, eventually, explain the positive manifold.[158]

Classical and recent attempts to strengthen the reliability of *g* consist of studying, then, the correlations between the *g* factor (or performance to tests) and some well-established neurocognitive phenomena or neurobiological variables (for some review, see Cianciolo & Sternberg, 2004; Pretz & Sternberg, 2005; Williams et al., 2008). Associations have been found, for instance, with:

- The speed and efficiency of brain functioning inferred from reaction time or assessed by electrical propagation of nerve impulses through the brain (Deary & Caryl, 1997; Engle et al., 1999; Grudnik & Kranzler, 2001; Jensen, 1998, 2002; Salthouse, 1996);
- Working memory (Conway et al., 2002; Conway et al., 2011; Engle et al., 1999; Gray & Thompson, 2004; Kyllonen & Christal, 1990);
- Cognitive constructs such as problem-solving, meta-cognition, attention and associative learning (Gray & Thompson, 2004; Kaufman et al., 2009; Kray & Frensch, 2002; Stadler et al., 2015; Sternberg & Frensch, 1990; Tamez et al., 2008; Williams, et al., 2008; Williams & Pearlberg, 2006). Explicit associative learning has recently been associated with *g,* sometimes statistically independently of working memory and processing speed. It involves the ability to remember and voluntarily recall specific associations between stimuli;
- The commonality of frontal lobe recruitment and fronto-parietal integration across a wide range of cognitive demands, including intelligence (Duncan et al., 2000; Duncan et al., 1995; Jung & Haier, 2007; Kane & Engle, 2002; Naghavi & Nyberg, 2005);

[158] On closer inspection, the two things are slightly different from each other (see §3.1).

- Neurobiological variables, like glucose metabolic rate and electrocortical activity (Deary & Caryl, 1997; Haier, 2003; Haier et al., 2003), and anatomic variables, such as brain size (McDaniel, 2005; MacLullich et al., 2002; Vernon et al., 2000; Wickett et al., 2000).

The aim of this chapter is not to provide a complete review of this literature. As instances, I shall mainly focus on the most renowned findings, i.e., about processing speed and working memory.

Several scholars refer to processing speed as a cognitive mechanism which may account for *g*. Participants with higher *g* scores tend to respond faster in simple, choice reaction time and inspection time paradigms. Salthouse offers two reasons for why speed should be so important:

"First, if you are slow to process information, and you cannot control the rate at which it is presented, then you are likely to miss information, some of which may be needed for the behavior in which you are engaged. Second, coordination between two different tasks is likely to be impaired if you are slow, because you may take so long on one task that you forget information that is needed to perform the other task" (Williams et al., 2008, p. 224).

Others advocate that intelligence is synonymous with working memory since it is strongly correlated with *g* both from a performance and neurological perspectives. Even if there is no full agreement about the working memory construct (Williams et al., 2008), it is generally assumed that information is maintained in a memory storage during the information processing.

"Many researchers assume that […] more capable individuals […] have greater working memory 'capacity' and/or a more effective attentional or executive system that allows the memory system to be less disrupted when simultaneous processing is required" (Williams et al., 2008, p. 225).

As Jensen (2002, p. 52) notices, the functional basis of why and how all these physical variables are correlated with *g* is not yet known. According to him,

"The explanation for it in causal rather than merely correlational terms is now the major research task for the further development of *g* theory. Some of the as yet inadequately investigated and unproved hypotheses that have been put forth to explain the relationship of *g* to brain variables involve the total number of neurons, the number of connections between neurons (dendritic arborization), nerve conduction velocity, the degree of myelination of axons, the number of glial cells, and brain chemistry (neurotransmitters, ionic balance, hormonal effects, and so on). The *g* factor at the level of psychometrics is now well established. Discovering its causal explanation, however, obviously requires that investigation move from psychology and psychometrics to anatomy, physiology, and biochemistry" (Jensen, 2002, p. 52).

Therefore, Jensen recognizes the necessity of moving from mere correlations to causal explanations of *g* and believes that this shift will eventually take place. Some authors are less optimistic than Jensen. For instance, Frank Ramus (2017) leaves little room for such neurocognitive explanations of *g*:

"Every attempt to reduce general intelligence to a single cognitive (processing speed, working memory, etc.) or biological (brain volume, nerve conduction velocity, etc.) construct has failed, each construct showing moderate correlation with *g* and being best described as simply one contributor to the *g* factor (e.g., Mackintosh, 2011)" (Ramus, 2017).

As the reader might notice, Ramus distinguishes between: a) reducing general intelligence to something else, and b) finding moderate correlations between *g* and contributing constructs. This paves the way for a distinction between *reducing* the psychobiological *g* to another psychobiological aspect of human cognition and just *explaining* the psychometric *g*. How is it possible to reduce *g* to another phenomenon?

3.1. Theoretical criteria for a g-reduction

As I mentioned above, two different things can be identified in the literature about the biological correlates of *g*: on the one hand, the identification of a phenomenon that accounts for most of the g-related variance; on the other hand, a phenomenon that explains some aspects of *g* and, hence, accounts for some part of the g-related variance. In other words, the former pertains to a complete reduction of *g* to another phenomenon, while the latter regards an explanation of the positive manifold by means of several different aspects:

- g-Reduction: A single aspect *x* of human biology, or neurocognitive architecture, is *identical* to *g*. Consequently, whenever we refer to *g*, to general intelligence or to IQ, we are actually referring to *x*. If one finds a strict correlation between *x* and *g* (i.e., *x* accounts for almost the whole g-related variance), then *x* represents general intelligence.

- g-Explanation: A set of distinct aspects [*x*, *y*, *z*…*n*] of human biology, if taken together, explain the positive manifold. It is worth reminding that the positive manifold is an empirical phenomenon that can arise for several reasons. In the case of the g-explanation, the reason why the performance to tests are intercorrelated consists of a specific relationship between *x*, *y*, *z*, etc. However, this set of biological aspects does not coincide to general intelligence.

As I will show, nothing like a g-reduction can be accounted for by empirical data; conversely, some g-explanations have been proposed by interpreting *g* as reflecting a causal network of distinct interacting aspects (see §3.2). I believe, however, that a g-reduction is required to make sense of general intelligence within an ontological framework—otherwise, a realist theory of general intelligence cannot be provided.

A first reason why a g-reduction seems to be important concerns the fact that the PSYGEN approach conceives *g* as a single entity: as I discussed in the previous chapters, for instance, *g* causes individual IQ differences; *g* is the target of the genetic influences; *g* does not change over lifetime, to some extent. In my view, any realist explanation of *g* should maintain this sort of "solidity" for *g*.

By contrast, without a g-reduction the unitary nature of *g* would be lost. Indeed, if no *x* identical to *g* exists, then one needs another entity, let us say *y*, to account for *g.* In other words, the biological aspect *x* representing *g* should account for the entire g-related variance: if there is no (almost) perfect overlap between *x* and *g*, then another phenomenon *y* should be invoked for the sake of accounting for *g*. It would follow that *g* is not a unitary general cognitive ability with clear biological bases.[159]

A second reason to seek a g-reduction is more "pragmatic" and concerns the endless debate about the magnitude of the obtained empirical correlations. Many, if not most, of the explanatory concepts mentioned above (e.g., processing speed and working memory) show moderate correlations with intelligence-test performance. I agree with Kray & Frensch (2002) and Hunt (1980):

"The argument between the generalist and the specialist view does, at times, take some of the aspects of an argument over whether a glass is half full or half empty" (Hunt, 1980, p. 466).

"Researchers preferring a strict interpretation of *g* […] often interpret correlations between their proposed construct and psychometric test performance that are above .4 as supporting their view […]. On the other hand, researchers opposing a strict interpretation of *g* often interpret correlations of less than .2 in favor of their own view" (Kray & Frensch, 2002, p. 211).

In the light of these controversies, Kray & Frensch argue that the magnitude of the obtained correlations cannot be taken as a reasonable basis for evaluating the appropriateness of differing accounts of what the cognitive manifestation of *g* might be. Rather,

[159] Kray & Frensch (2002, pp. 184-186) raise similar concerns. According to them, two different interpretations of the statement that "there exists a general ability of intelligence" can be made: 1) there exists one and only one source of all intelligent behavior (strict interpretation); 2) there exists one source that *affects* all intelligent behavior (loose interpretation). The loose interpretation implies that it is possible that behavior is influenced by other sources as well. According to the authors, the assumption of a general ability makes sense only if the impact of *g* on all forms of intelligent behavior is large relative to the impact of other abilities. Because of this, Kray and Frensch do not differentiate between the two possible interpretations of g. Instead, they interpret the concept of a "general ability *g* of intelligence" as meaning that there exists a source or ability such that the influence of this source or ability on all forms of intelligent behavior is large relative to the impact of other sources or abilities.

some sort of theoretical principle is required, about the exact nature of such a general ability.

What kind of evidence might constitute a support for a g-reduction? For providing a possible answer, I shall take Kray and Frensch (2002) criteria for evaluating neurocognitive accounts of *g*. Indeed, their analysis sounds quite reminiscent of what I called a "g-reduction".

"By asking what the nature of *g* might be, we are searching for a cognitive manifestation of a general ability. More specifically, we are asking which mental processes and representations or which properties of mental processes and representations might be primarily responsible for intelligent behavior" (Kray & Frensch, 2002, p. 187).

The five criteria proposed by Kray and Frensch must be met before any cognitive construct can truly be considered a cognitive manifestation of *g*:

- Criterion 1 – Theoretical foundation: the account must be theoretical rather than empirical. For instance, if the assumption that speed of mental processing is a potential cognitive manifestation of *g* is to be acceptable, it needs to be spelled out how exactly the speed of processing is realized in the cognitive system.
- Criterion 2 – Multiple measures of *g*: in empirical research relating *g* to the proposed account, *g* must be measured in multiple ways. This is important to ensure the external validity of *g*.
- Criterion 3 – Control of Third Variables: any empirically observed relation between the proposed account and *g* must not be due to the influence of third variables. Indeed, potential third variables might modulate an empirically observed relationship between an account and *g* (i.e., spurious correlations).
- Criterion 4 – Direction of causality: the direction of causality must be demonstrated empirically. All the proposed theoretical accounts establish a causal direction between the construct of interest and psychometric *g*. At the same time, empirical studies are usually based on cross-sectional data and unlikely to prove the assumed direction of causality. Hence, longitudinal research designs are required to clarify whether the proposed account is a cause rather than a consequence of intelligent behavior.
- Criterion 5 – Theoretical plausibility: the proposed theoretical relation between the account and *g* must be plausible. This entails a rationale for why a relationship between the construct of interest and *g* should hold.

In their paper, the authors analyze several cognitive constructs under the five criteria just described. Table 7.2 summarizes the resulting analysis.

Construct	*Criterion 1 Theoretical Rationale*	*Criterion 2 Dependent Variables*	*Criterion 3 Third Variables*	*Criterion 4 Causality*	*Criterion 5 Rationale of Relation*
Attentional resources	medium	single	Yes, strategies	—	low
Speed of information processing	medium	two/three	No	—	low
Speed of visual processing	low	two/three	No	Yes	low
Cognitive components	medium	multiple	No	—	medium
Metacognition	medium	multiple	Yes, contents	—	medium
Goal selection	low	two/three	—	—	medium
Inhibitory processing	medium	multiple	—	—	medium

Table 7.2: Some candidates for the role of cognitive manifestations of *g*.
From Kray & Frensch, 2002, p. 161.

Let us see some details by taking processing speed as an explicative case. The theoretical concept underlying a speed-of-processing account (Criterion 1) is straightforward. Jensen (1984), argued that the cognitive manifestation of *g* is a global property of the brain associated with the neural efficiency of the cerebral cortex (i.e., the number of neurons activated by the environment, and the rate of oscillation between refractory and excitatory phases of neural processes). Individual differences in reaction time are viewed as reflecting a "hardware" component of the cognitive system that is independent of knowledge, skills, or cultural background.

However, limitations are often related to (a) the lack of a clear definition of the construct, and (b) low validity. Psychometric and experimental tasks are often used as indicators of individual differences in hypothetical constructs that generally involve a relatively complex sequence of processes. For instance, traditional perceptual speed tests measure not only speed of processing; they also measure the ability to coordinate visual and working memory processes. Moreover, there is also empirical evidence indicating that taking *more* time sometimes leads to a greater likelihood of solving a problem (see Sternberg & Davidson, 1983).

Criterion 2 is generally met by proponents of a speed-of-processing account. It demands that researchers use multiple measures of *g* (e.g., experimental tasks) to enhance the validity of measurement. This requirement is consistent with many modern intelligence theories that allow for multiple types of intelligence (e.g., Gardner and Sternberg theories). Researchers typically use more than one IQ test as indicators of mental speed (e.g., Raven's Matrices Tests). However, like most IQ tests, Raven's test captures only a part of intelligent behavior. Thus, one should keep in mind that the correlations obtained cannot be interpreted unless one accepts that the intelligence tests indeed measure intelligence.

About the influence of third variables (Criterion 3), in the case of processing speed, there are few alternative accounts based on the possible influence of third variables. In general, for some of the discussed theoretical constructs, there are possible alternative sources of individual differences that might affect the amount of variance in *g* explained by the proposed construct (e.g., individual differences in strategy choice, training or pre-experimental knowledge).

Criterion 4, asking for empirical evidence confirming the assumed direction of causality, is not generally met by speed-of-processing accounts (e.g., lack of longitudinal studies). In all theoretical accounts, it is assumed that the proposed theoretical construct is the source rather than the consequence of human intelligence. However, most of the reported empirical findings are based on cross-sectional studies. Longitudinal research is needed to (a) clarify the direction of causality, and (b) determine the relative impact of the proposed construct to individual differences in *g*.

Equally problematic seems to be the lack of a plausible theoretical reason for the relation between mental speed and *g* (Criterion 5). One of the main difficulties in finding determinants of human intelligence consists of coming up with theoretical accounts that capture the full range of intelligent behavior. According to Kray and Frensch, the proposed construct should be not too specific (e.g., visual inspection time) but also not too broad (attentional resources) to be considered a serious candidate for capturing the essence of *g*. Many of the accounts discussed cannot fulfill this criterion.

In sum, at best, three of the five evaluation criteria are met by speed-of-processing accounts of human intelligence. Thus, mental speed seems not to be the only source of variance accounting for individual differences in intelligence test performance. Therefore, it is not a good choice for such a "basic" process.[160]

Similar concerns can be raised for working memory. It has been suggested that reasoning ability is little more than working memory capacity (Kyllonen & Christal, 1990), but this strong claim has been strenuously criticized (Ackerman et al., 2005; Conway et al., 2003; Gignac & Watkins, 2015).

> "Recent theoretical accounts, have retreated from the strong view that fluid intelligence and working memory capacity are essentially isomorphic, in part because studies have shown that working memory tasks are far from perfectly correlated with intelligence. This means that much of the variance in *g* must be explained by factors other than working memory" (Williams et al., 2008, p. 226).

For instance, Engle and his colleagues recently proposed that the reason why working memory performance predicts fluid intelligence is not because memory *per se* is used to solve reasoning problems; rather, it is because working memory tasks measure how well an individual's attentional processes function under conditions of interference, distraction, or conflict (see Hasher et al., 2007; Kane et al., 2007; Unsworth & Engle, 2005).

Having provided similar analyses for many cognitive constructs, Kray and Frensch (2002, p. 211) conclude that none of the information-processing concepts they considered meets all five criteria; therefore, none of them can be considered a cognitive manifestation of *g*.

[160] The view that mental speed is the single source of individual differences in *g* seems to be implausible to many scholars. While there is little doubt that a substantial fraction of the variance in intelligence scores is related to measures of processing speed, numerous investigators have questioned its adequacy as a complete account of *g* (e.g., Stankoff & Roberts, 1997; Stankov, 2002; Williams et al., 2008).

Even if the proposed criteria seem to be satisfying, further criteria can be included to ensure the g-reduction. For instance, I previously pointed at the magnitude of the empirical relationship.[161] As I previously argued, the correlations between *g* and *x* should be substantial—otherwise, another construct *y* should be invoked for accounting for the remaining variance. Since the correlations between the most studied cognitive constructs and *g* are far from being perfect—but rather they are generally attested between .2 and .4—this criterion is generally not satisfied. That is, there is no cognitive process, or neurobiological variable, which is identical to *g*. In a nutshell, right now no such thing as a g-reduction can be expected to happen.

One last remark can be raised about a clear definition of the constructs under examination. As I mentioned in Chapter 5, if behavioral phenotypes and psychological constructs are not well defined, then something will likely go wrong in empirical research. This applies both to IQ, general intelligence, mental disorders, and cognitive processes. If one analyzes correlations between, say, the presence of a neurotransmitter and a mental disorder, and if the disorder is not well-defined, one will likely obtain confounding data (e.g., spurious correlations and low replicability-rates). If general intelligence is not a general and uniform phenotypic trait—but rather a cluster of different traits (e.g., cognitive components)—then analyzing statistical associations between IQ and SNPs could be problematic in several respects. In sum, if one tries to reduce *g* to another phenomenon, both the two must be clearly defined.

What does all this imply for contemporary theories of general intelligence? I submit that general intelligence should be subtyped to different cognitive processes, not necessarily independent of each other, but largely autonomous from a causal and a developmental viewpoint (see Chapter 6). This gives strength to theories of intelligence like the PASS Model and the Multiple Cognitive Mechanisms Approach, which do not appeal to any general mental ability.

Before drawing any conclusion, however, another step deserves to be taken: what is *g*, after all? Why are cognitive tests correlated with each other? Can cognitive sciences provide a g-explanation? In the next paragraphs, I delineate two possible hypotheses about the nature of *g* that remain to be addressed: a) *g* reflects interactions between several internal and external aspects characterizing humans' beings (*causal-network-g*); and b) *g* is a "distillate" of tests variance (*distillate-g*). As I shall show, the first one does not require an ontological commitment about *g* but, nevertheless, allow to include it in some sort of explanation. The last hypothesis, instead, completely dismiss the possibility of an ontological commitment about *g*.

[161] According to Kray and Frensch, this criterion is not usable because a comparison of correlation coefficients between studies that differ in ability ranges and sampling strategies might be severely flawed. However, this is a contingent fact: ideally, a test for intelligence should include items capable of measuring every aspect related to intelligence and only them (see Chapter 1).

3.2. Developmental explanations of g

The Multiple Cognitive Mechanism Approach holds that the psychometric *g* may not be comprised of a single cognitive mechanism (i.e., psychobiological *g*) but instead is supported by multiple, interacting mechanisms that become associated with each other throughout the course of development (see §2.2). I call this interpretation of *g* the *causal-network-g*.

This account of *g* characterizes the mutualist model proposed by van der Maas et al. (2006, 2014). The authors agree that the positive manifold is a robust empirical phenomenon. However, they propose an explanation based on a developmental model involving beneficial relationships between cognitive processes. In other words, they say, the mutual influences between cognitive processes represent a plausible mechanism that gives rise to the positive manifold but that does not include *g* as a latent quantitative variable representing variable such as speed of processing.

> "In the initial phase of development, cognitive processes are uncorrelated. During development, the positive manifold emerges as a consequence of mutually beneficial interactions between these processes. Factor analysis of data generated by this dynamical process suggests the presence of a dominant factor […]. Interestingly, under certain circumstances, the mutualism model and the factor models are statistically equivalent, in the sense that they produce the same covariance structure […]. However, the mutualism or cooperation between processes is conceptually very different from the *g* explanation, in terms of a single quantitative dimension" (van der Maas et al., 2006, p. 855).

According to the authors, there is nothing wrong with using the *g* factor as a summary index variable (e.g., if it allows successful predictions), as long as we do not assume that this variable relates to a single underlying quantitative process or capacity.

A similar proposal has been made by Hampshire et al. (2012). According to them, human intelligence is not unitary. Rather, the basis of the higher-order component called *g* may be accounted for by cognitive tasks co-recruiting multiple functionally dissociable brain networks. Cognitive components reflect the way in which the brain regions implicated in intelligent behaviors are organized into functionally specialized networks. However, cognitive tasks recruit a combination of these functional networks. The authors conclude that human intelligence is most parsimoniously conceived of as an emergent property of multiple specialized brain systems, each of which has its own capacity.

Lautrey (2002) suggests a similar explanation for *g* by considering it a developmental property (i.e., a general factor of development).[162] In this explanation, processing capacity and processing speed play a role in the apparent generality of intelligence. However, they are not interpreted as a single mental process, but rather as global features of the neural system realized by many distinct aspects of the biological architecture of the brain. For

[162] Lautrey's general purpose is to combine psychometric data with the developmental theory of intelligence provided by Piaget.

Lautrey, the notion of an upper limit in the processing capacity fits better with the observations than a general structure. It should not, however, be inferred that this processing capacity corresponds to a unitary cognitive mechanism. Rather, the development of processing capacity can be explained by the maturation of the central nervous system (e.g., myelinization), but also by environmental factors (e.g., automatization of information processing with exercise, the discovery of metacognitive strategies, and the influence of instruction which increases knowledge simultaneously in various fields).

"All these factors, maturational and environmental, covary with age and there are interactions between some of them, for example, via pruning, between the waves of dendritic connections and exercise. It is thus illusory to search for a single, general-purpose, elementary process, that would account for the upper limit of processing capacity and thus for the existence of a general factor of development. The increase in processing speed, sometimes advanced as an elementary mechanism susceptible to play this role […] results from changes in the complete set of these factors and is thus only one global indicator of development, as global as mental age. Explaining the general factor of development by an increase in processing speed adds little more than explaining it by an increase in mental age" (Lautrey, 2002, pp. 144-145).

Similar explanations of *g* have been proposed by Cattell (1971) and Ackerman (1996), too. Cattell provided an explanation of the nature of *g* in terms of broad sets of influences that go beyond the cognitive domain. An early proposal of Cattell was that Gf and sensory processes interact with the environment from the early stages of human development.[163] The outcome of these interactions is the development of crystallized intelligence (Gc). The full structure of human abilities is, therefore, the result of history. Ackerman, in turn, proposed in 1996 another model of intelligence in historical terms. Accordingly, *g* is the outcome of ontogenesis consisting of many loosely related elements that have been selected by processes that have relatively little to do with cognitive abilities.

As Stankov (2002) summarizes, it does follow that at any point in development *g* captures a mishmash of different things that are continuously changing. The author concludes:

"there is no single cognitive process that can explain the presence of *g*. Even a small number of core processes is unlikely to suffice for this purpose. It is a mixture of many different processes (including noncognitive influences) that are known to change in the course of development. The search for a single biological basis of *g* might be a futile exercise" (Stankov, 2002, p. 35).

A similar conclusion has been recently drawn by Ramus (2017). He explains the positive manifold (i.e., the psychometric *g*) by noticing that there is no test which is a pure measure of a cognitive function or construct. The relationship between cognitive

[163] Interests play an important part in the choice of these interactions, by reinforcing some types of activities and eliminating others. In his many writings on the development of cognitive abilities, Horn (1985) extended the list of influences beyond interests to include a host of proximal and distal causes (see Stankov, 2002).

functions and test scores is many-to-many: each test or subtest (e.g., mathematical, verbal, or spatial) recruits several cognitive functions, and each cognitive function is involved in several tests. This is the reason why test scores are positively correlated. According to Ramus, interactions within the brain are so strong that positive correlations among cognitive functions should be expected:

"Each brain function or property (e.g., frontal grey matter volume, nerve conductance velocity, dopamine synthesis, etc.) influences several cognitive functions, thereby inducing intrinsic positive correlations between cognitive functions. One step further back, each gene expressed in the brain (e.g., genes that code for neurotrophic factors, transcription factors, and any molecule involved in neurotransmission) typically influences several brain functions and properties, thereby inducing positive correlations between them. In parallel, many environmental factors (e.g., nutrition, socioeconomic status, education, diseases…) influence more than one brain or cognitive function, thereby inducing further correlations" (Ramus, 2017).

Are developmental theories of *g* convincing? From an empirical perspective, they seem to be quite solid: as Van der Maas and colleagues (2006) notice, the best way to falsify the mutualist model is to find a variable that correlates perfectly with *g* (likely, this is also true for the other developmental models I analyzed in this paragraph). As I previously argued, nothing like a g-reduction is justified by empirical data. From a theoretical point of view, more remarks can be made.

First, developmental theories accept and refuse the *g* factor at the same time: on the one hand, they explain the psychometric *g*; on the other hand, they dismiss the psychobiological *g*. Thus, they provide a g-explanation: a set of distinct biological or cognitive aspects [*x, y, z…n*], if taken together, explain the positive manifold. In other words, the reason why the performance to tests are intercorrelated consists of a relationship between *x, y, z*, etc. Second, they seem more consistent with biological sciences than the PSY-GEN model. It is nothing new that organisms are characterized by complex networks of causal interactions, which include their environment. Developmental explanations of *g* account for this aspect. Figure 7.2 represents in graphical terms a possible interpretation of such developmental models, including some aspects I have discussed in the previous chapters.

This summary of different theoretical accounts includes cognitive theories of intelligence like the PASS Model and the Multiple Cognitive Mechanisms Approaches and accounts for empirical correlations between many variables and performance to tests. Let us see some details:

- Genes and Environment: Individual allelic variation and environment interact in such a way to produce the experience to which an individual is exposed over the lifetime. Experiences influence genetic expression, and genetic expression has some effects on the peculiar experience to which individuals are exposed (see genes-environment interactions, Chapter 4).

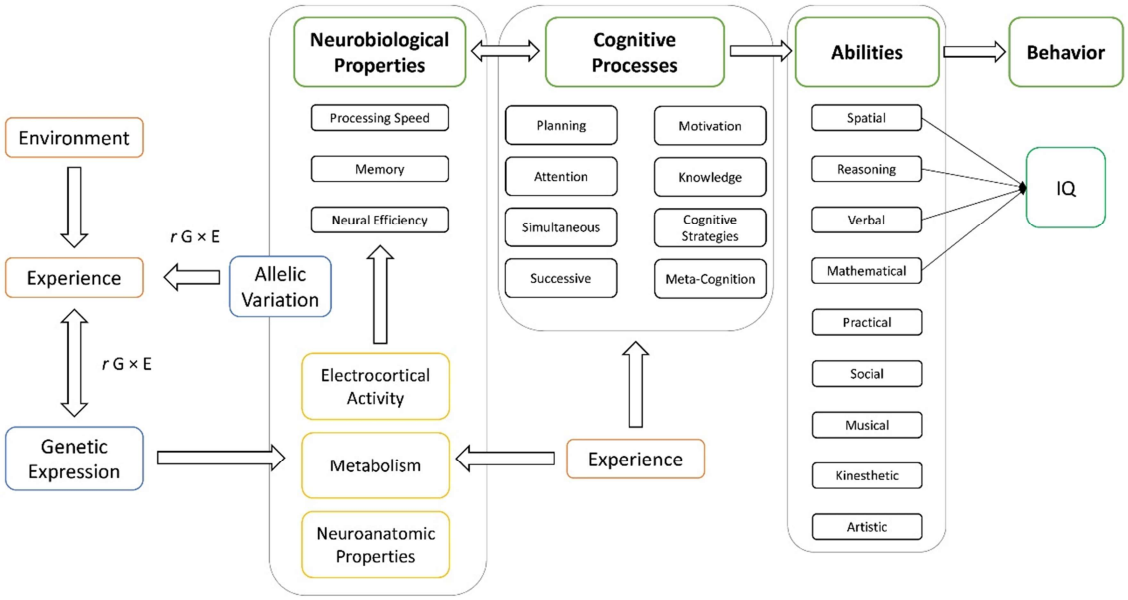


Figure 7.2: A developmental explanation of the psychometric *g*.

- Neurobiological Properties: Genetic expression implements properties of the neural system, like metabolic reactions, electrocortical activities and neuroanatomic properties (e.g., brain size).[164] These neurobiological features are influenced by experience as well: mammals' neural system is very plastic, especially within prefrontal cortex areas—generally associated with complex cognitive activities—, and open to modification over the entire lifetime. Systemic properties of the brain are likely the material bases of cognitive systemic properties such as processing speed, memory, and efficiency.

- Cognitive Processes: They form an open set of interacting high-level phenomena and coincide, to a lesser or greater extent, with the processes included in the cognitive theories of intelligence (the PASS Model, especially).[165] Higher-level processes are subject to environmental influences (i.e., educative experiences). The important aspect is that there is a mutual relationship between neurobiological properties and cognitive processes: that is, higher-level processes are affected and can affect lower-level systemic properties. For instance, individual cognitive strategies can modify over development the neural efficiency or some other aspect of the in-

[164] The inclusion of anatomical variables makes sense of their correlations with psychometric *g* (see §3).
[165] However, this set is open to revision in the light of new theories or empirical findings. For instance, the biological bases of several cognitive constructs, how genes influence them, and whether they are phenotypic traits narrow-sense, is far from being clear.

formation retrieval. On the other hand, being every cognitive process and integrative processes, they influence any Ability. For instance, planning, motivation, and acquired knowledge affect the set of abilities involved in IQ tests (see below).

- Abilities: This is an open set of skills aimed at resolving specific tasks. Some of them are related to psychometric tests and bring about individuals' IQ scores. Other skills are not related to tests, but rather are the object of non-psychometric theories of intelligence (e.g., Gardner's Multiple Intelligence and Sternberg's Successful Intelligence theories).

- Psychometric *g*: It arises from the mutual interactions between any element of the sketch throughout the development. Empirical correlations between variables are due to those mutual relationships.

The proposed sketch is consistent with the splitting-strategy proposed in Chapter 6: subtyping to distinct aspects something that seems *prima facie* a unified phenomenon (i.e., general intelligence). Test solving seems to be due to many cognitive abilities and it is pointless to categorize them as a single thing. As I discussed, it seems empirically more productive to search for genes that act on neurobiological properties rather than genes that act on general intelligence; indeed, the targets of genetic and environmental influences are likely those properties (see Chapter 5). General intelligence, at the very end, seems to be the name we give to the set of all these entities and their interactions.

3.3. The distillate-g

Before concluding, one last hypothesis about *g* should be considered. As I explained in §1, cognitive sciences are mainly interested in the relationship between *g*, cognitive processes and behaviors (e.g., mechanistic or causal explanations). Conversely, psychometrics is mainly interested in the role of *g* in individual differences. So far, I did not emphasize the fact that *g* represents variance among test performances. Factor analysis is often conceived as a technique capable of highlighting latent variables in a correlation matrix.[166] The tendency to reify a factor comes often naturally: in a sense, it is natural to think that the identified factor is something that *exists* and that is detected by factor analysis. However, a general factor is not necessarily something that arises *in addition* to other variables.

[166] This is, in fact, how Spearman "discovered" *g,* since he found that scores on all mental tests (regardless of the testes domain) tend to load on one major factor. See Chapter 1.

Jensen (2002) explains the topic quite clearly, although, as I have stated on several occasions, the concept of variance can be very puzzling. According to Jensen, factor analysis is a procedure of "distillation": it identifies a factor reflecting the variance that all the different intellectual measures have in common.

"It could never have been discovered with $N = 1$, because it reflects *individual differences* in performance on tests or tasks that involve anyone or more of the kinds of processes just referred to as *intelligence.* […] The *g* factor should be thought of not as a *summation* or *average* of an individual's scores on a number of diverse tests, but rather as a *distillate* from such scores. Ideally, it reflects only the variance that all the different tests measure in common" (Jensen, 2002, p. 40).

Deary (2002) connects this definition of *g* to hierarchical theories of intelligence:

"The psychometric studies suggest that there might be different targets for cognitive or broader information processing studies: general variance, and group and specific factor variance. [However,] people conducting reductionistic validity studies need to be aware of the limitations of psychometric studies. […] It must be recalled that the hierarchical structure of the covariance of ability test scores exists as a finding that is not necessarily *isomorphic* with anything *in people's heads*; the three-level hierarchy is a taxonomy of tests, *not of human's mental structures* (not necessarily, anyway)" (Deary 2002, pp. 152-153; emphasis added).

In the light of this definition of *g*, Stankov (2002, p. 34) states that *g* gains its strength from the fact that it captures a little bit of variance from many tests. What remains within each test to be picked up by *g* is a distillate that may be psychologically uninterpretable.

Variance is a feature of a population and regards individual differences. If *g* reflects individual differences in a population, how could it be a biological characteristic? If taken seriously, the hypothesis that *g* is nothing more than a "distillate" of tests variance strongly suggests that any reference to the psychobiological *g* should be dismissed as soon as possible. Indeed, there is no such concrete thing as *g* in individual organisms: *g* is, rather, an abstract entity, perhaps what the critics have often called "a statistical artifact". Further research should be devoted to clarifying the ontological status of an entity like a *distillate* from tests scores, something that solely reflects the variance that all the different tests measure in common.

Conclusion: One Last Question

My inquiry about human intelligence has mostly assumed that the general factor of intelligence is not just a statistical artifact. I asked, rather, whether a realist standpoint about general intelligence is tenable. This requires for general intelligence to be grounded on a general cognitive mechanism, such as *g*, which represent the nexus of many causal influences: from genes to *g* and from *g* to cognitive processes and behaviors—in philo-

sophical terms, the homeostatic mechanisms of an HPC called 'general intelligence'. Unfortunately, there is no a neurobiological entity that might account for *g* in this sense, neither from an empirical nor from a theoretical perspective.

Both the accounts of the distillate-g and the causal-network-g can explain why the *g* factor is not reflected in empirical research from cognitive neurosciences. Hence, they could be regarded as instrumentalist with respect to *g*. Indeed, none of them accept *g* as a genuine neurocognitive or biological entity—that is, there is no something like *g* in humans' brain. However, the developmental account allows a charitable interpretation of *g* according to which *g* reflects something more than a mere statistical entity, that is, the interaction among biological and environmental aspects throughout development. What does this imply for intelligence theories? This implies that, at least, their ontological commitment about *g* should be revised to some extent. Therefore, I suggest, we must avoid any reference to general intelligence since it does not represent a rewarding posit for science and its aim of carving nature at its joints.

One question remains open: Can general intelligence be regarded as a valuable instrumental concept? Can general intelligence have some sort of intrinsic value for clinical, educational, or social practices? For instance, one might think that, if there is *a precise way* to estimate the causal-network-g, *and* if this interaction has *predictive power* with respect to some useful variable (e.g., social, educational or clinical), we could maintain the concept of general intelligence as it is instrumentally defined. I will briefly address this question in the Conclusions.

Conclusions

Psychometrics and behavioral genetics have together developed one of the most remarkable and lasting framework for analyzing the complex trait we call 'intelligence'. Most of the psychometricians refer to intelligence as a general mental ability, namely the *g* factor. Behavioral geneticists describe intelligence as a quantitative phenotypic trait which is related to the small effect of several genes, namely, the IQ. These two major tenets have enabled scientists from many research areas to analyze intelligence by means of statistical and empirical methods.

In my thesis, I tackled this framework by arguing that it is not capable of describing human intelligence in ontological terms. Indeed, both psychometricians and geneticists have mainly relied, in their work, on statistical approaches such as factor analysis and the analysis of variance. By contrast, they have rarely uncovered their ontological commitment and preferred to adopt an instrumental conception of that psychological construct. At the same time, researchers have frequently sought the biological correlates of the *g* factor and the genetic bases of individual differences in IQ, clearly importing thus a realist viewpoint involving an ontological commitment to *g*.

For advancing my criticisms, I have shed light on the ontological assumptions which underlie the PSY-GEN approach to the study of intelligence. What does it mean that intelligence is general and quantitative? With respect to behavioral genetics, I tried to make sense of the conception of intelligence as a quantitative phenotypic trait by asking whether it is related to developmental quantitative phenomena. In this respect, the additivity assumption, widespread among geneticists, served as my main analytic, critical target. With respect to psychometrics, I tried to make sense of the conception of intelligence as a general cognitive ability by asking whether there is any aspect of the human neurobiology capable of accounting for the so-called *psychobiological g*. With the help of the natural kinds theory, I argued that such a phenomenon should represent the causal mechanism connecting biological, cognitive, and behavioral aspects of intelligence. Thus, it represents the basis for prediction, generalization, and for the scientific reference to 'human intelligence' as well.

I believe that a theory of general intelligence that overlooks biological and causal aspects cannot be taken seriously. Therefore, the advocates of *g*, for arguing in favor of its real existence, should make several empirical and theoretical steps aimed at avoiding the reference to general intelligence as merely instrumental. First, they should rule out misleading conceptions about *g*. For instance, *g* cannot be considered *quantitative*, since what is properly quantitative is the variation of the IQ in a population. Then, *g* should be

conceived as a unitary entity since it cannot be reduced to many different aspects of human cognition without losing its "generality". Second, the advocates of *g* should explain how genetic, epigenetic, and environmental causes act upon a single biological phenomenon. Third, they should identify this phenomenon, and show that, in any reasonable sense, it is *g*. This identification requires, for instance, the candidate phenomenon being suitable to account for the *g*-related variance in factor analysis. Moreover, this identification requires that what IQ tests measure is, at the very end, such a phenomenon and that the phenomenon accounts for individual differences in IQ tests. A *g* satisfying all these demands would represent: a) the causal source of the clustering of biological, cognitive, and behavioral aspects; b) the causal source of the positive manifold; c) the basis of prediction and generalization; d) the ontological reason why intelligence is a general cognitive ability; e) the reason why we can use the term "intelligence" in an ontologically-committed manner.

I argued that both biological and cognitive sciences deny the existence of general intelligence within the human organism. Concerning biology, rather than being a unitary phenotypic trait in the narrow sense, IQ is better regarded as a quantification over different aspects of human cognition. I argued for this conclusion by connecting the definition of a proper phenotypic trait with the analysis of genetic causation on behavioral and cognitive aspects. Concerning cognitive sciences, rather than being a single causal mechanism underlying intelligent behavior, the *g* factor is better understood as a correlational phenomenon arising from interactions between several aspects of human cognitive development. According to this developmental explanation of the psychometric *g*, the term 'intelligence' refers to a causal network.

The question, now, is: Even though there is no such thing as general intelligence, there is any sense in which it is fruitful to use this term? Can we adopt a plurality of definitions of intelligence in relation to different practical purposes? As Ludwig (2015) notices, different explanatory interests within a scientific discipline lead to different conceptual needs. It is then not surprising that some psychologists prefer to work with an ontology that includes a general intelligence while others insist on an ontology of multiple intelligences. From an ontological perspective, I argued, the notion of general intelligence makes no sense. Hence, the present question pertains, rather, to epistemic practices, e.g., prediction and generalization, and to pragmatic concerns, e.g., clinical or social.

Two important aspects of general intelligence, in these respects, can be highlighted: first, IQ scores are good predictors of school achievement; second, IQ tests can discriminate, even broadly, between normal and pathological intellectual conditions. As Deary (2002) states, mental ability differences are significant predictors of educational, occupational, and social outcomes, with effects sizes that are typically moderate to large. This suggests that mental tests and the IQ construct have at least some practical importance. The same might apply to the causal-network-g: if there is *a precise way* to estimate the causal-network-g, *and* if it has some sort of *predictive power* with respect to some useful

variable (e.g., social, educational, or clinical), we could maintain the concept of general intelligence as it is instrumentally defined. Several doubts, though, have been raised about this possibility.

A methodology for evaluating the causal-network-g could lie in tests inspired by cognitive theories (e.g., the PASS Model). However, remarkably, these tests do not produce a single score; then, they are not good assessments of *g*, and hence they cannot be employed for predictions with the causal-network-g. Consequently, classical IQ tests must represent the putative way for estimating *g*, and then, the causal-network-g. In this case, however, the value of IQ as a good avatar of human intelligence has been often questioned. In my thesis I did not analyze, if not tangentially, theories of intelligence which include non-cognitive aspects in the domain of human intelligence (e.g., Gardner and Sternberg's theories). Nevertheless, these theoretical alternatives exist, and they can be taken as testimonies of the inadequacy of IQ tests in accounting for intelligence.

This is also related to how IQ tests are "composed". According to Naglieri and Das (2015), the predictive power of IQ tests derives from a circular reasoning, since the tests used to measure intelligence are remarkably similar to achievement tests. If there is not a clear distinction between mental ability and achievement, then any child who does not have an adequately enriched educational experience will be disadvantaged when assessed with a so-called "ability" test.

Schneider and Flanagan (2015), in turn, stress that, even if organizational psychology has a long tradition of using IQ tests for personnel selection, this can be due to how difficult is to demonstrate the differential validity of specific abilities in job performance. However, if all test scores are determined solely by specific abilities, it is hard to justify the practice of calculating IQ. Why? If only specific abilities are measured, then a person's IQ would depend on the arbitrary choices made by test developers as to which abilities are measured.

According to Van der Maas and colleagues (2014), there is no such thing as a separate latent variable that we could honor with the term "intelligence", and it is questionable whether one should, in fact, use the word "intelligence measurement" at all in such a situation. However, if one insists on keeping the terminology of measurement around, there is little choice except to bite the bullet: interpreted as an index, intelligence is just whatever IQ-tests measure.

However, Schneider and Flanagan (2015) account for a possible usage of IQ tests for prediction. According to them, this practice would be justifiable if IQ is seen not as a measure of Spearman's notion of *g,* but as Binet's notion of *intelligence in general*. If there are a few broad abilities that influence performance in many tasks (e.g., verbal ability, spatial ability, logical reasoning), a simple average of these important abilities is a useful summary of a person's capacity to act intelligently. This is true regardless of the existence of *g*.

It is not my aim to dismiss any usage of the term 'general intelligence', once provided IQ tests are suitable to estimate the causal-network-g. In this case, the IQ and the causal-network-g would represent a summary of many variables, and some prediction could be made by them. Hence, an instrumental notion of general intelligence might be kept. However, I contend that this approach is promising. The question is: Is IQ's predictive power enough? A major tenet of my thesis is that, if IQ represents a collection of heterogeneous cognitive aspects, then correlational analyses about IQ can be misleading. For instance, conducting genetics studies on a phenotypic trait requires a good definition of the trait under examination. A fine-grained analysis, which looks at specific cognitive phenotypes, might find much stronger correlations and might allow for much better predictions with respect to social variables.

Furthermore, single-score tests can lead to harmful consequences. Some scholars have highlighted limits in the clinical utility of IQ tests since they are unsuited for evaluating specific developmental problems, which are likely the real cause of low scoring in tests**.** According to Naglieri and Das (2002), it is well documented in the literature that individuals with frontal lobe damage can earn average scores on a traditional IQ test. Similarly, many dyslexic children earn average or higher IQ scores but experience significant difficulty in reading. These examples clearly illustrate that, whereas these persons may have specific cognitive problems that relate to performance deficiencies, little value is obtained from a measure of general intelligence.

In education, classifying people with a general intelligence scale is often perceived as a limiting strategy. This is the reason why, Gardner's theory of multiple intelligences is so popular among pedagogues: indeed, it allows educators to concentrate on individual strengths and weaknesses.

Finally, as Ludwig (2015) notices, different accounts of intelligence do not only reflect different explanatory interests, but also often non-epistemic (e.g. moral, educational, political) values. The critics of general intelligence consider it a dangerous platform for pseudo-scientific justifications of racism and sexism. The history of the twentieth century is plenty of ideological usages of intelligence tests. The risk of similar fallout is always open. If general intelligence makes no ontological sense, I believe we should consider the hypothesis of ruling out it from the scientific vocabulary.

References

Ackerman, P. L. (1996). A theory of adult intellectual development: Process, personality, interests, and knowledge. *Intelligence*, *22*(2), 227-257.

Ackerman, P. L., Beier, M. E., & Boyle, M. O. (2005). Working memory and intelligence: The same or different constructs?. *Psychological Bulletin, 131*, 30-60.

Alford, J. R., Funk, C. L., & Hibbing, J. R. (1978). Are political orientations genetically transmitted? *American Political Science Review*, *99*, 153-157.

American Psychiatric Association (2000). *Diagnostic and Statistical Manual of Mental Disorders* (4th ed., text rev.). Washington DC: Author.

American Psychiatric Association (2013). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed.). Arlington, VA: American Psychiatric Publishing.

Anderson, J. (1938). The problem of causality. *The Australasian Journal of Psychology and Philosophy*, *16*(2), 127-142.

Anderson, M. (2005). Marrying intelligence and cognition. A Developmental View. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and Intelligence* (pp. 268-287). Cambridge, UK: CUP.

Baetu, T. M. (2015a). When is a mechanistic explanation satisfactory? Reductionism and antireductionism in the context of mechanistic explanations. In *Romanian Studies in Philosophy of Science* (pp. 255-268). Springer International Publishing.

Baetu, T. M. (2015b). The completeness of mechanistic explanations. *Philosophy of Science*, *82*(5), 775-786.

Baetu, T. M. (2016). From interventions to mechanistic explanations. *Synthese*, *193*(10), 3311-3327.

Barnes, B., & Dupré, J. (2008). *Genomes and What to Make of Them*. Chicago-London: University of Chicago Press.

Barrett, L. F. (2006). Are Emotions Natural Kinds? *Perspective on Psychological Science*, *1*(1), 28-58.

Bateson, P. (1987). Biological approaches to the study of behavioural development. *International Journal of Behavioral Development*, *10*(1), 1-22.

Bateson, W. (1909). *Mendel's Principles of Heredity*. Cambridge, UK: CUP.

Bearden, C. E., & Freimer, N. B. (2006). Endophenotypes for psychiatric disorders: ready for primetime?. *Trends in genetics*, *22*(6), 306-313.

Bechtel, W. (1994). Levels of description and explanation in cognitive science. *Minds and Machines*, *4*(1), 1-25.

Bechtel, W., & Abrahamsen, A. (2005). Explanation: A mechanist alternative. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *36*(2), 421-441.

Benisz, M., Dumont, R., & Willis, J. O. (2015). From psychometric testing to clinical assessment: Personalities, ideas, and events that shaped David Wechsler's views of intelligence and its assessment. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 163-179). New York: Springer.

Benjamin, J., Ebstein, R. P., & Belmaker, R. H. (2002). Genes for human personality traits: "Endophenotypes" of psychiatric disorders?. In J. Benjamin, R. P. Ebstein & R. H. Belmaker (Eds.), *Molecular Genetics and Human Personality* (pp. 333-344). Washington D.C.: American Psychiatric Publishing.

Benyamin, B, & Visscher, P. (2014). *Intelligence inheritance – three genes that add to your IQ score*. <http://theconversation.com/intelligence-inheritance-three-genes-that-add-to-your-iq-score-31397>

Binet, A. (1909). *Les idées modernes sur les enfants*. Paris: Flammarion.

Binet, A., & Simon, T. (1905). New Methods for the Diagnosis of the Intellectual Level of Subnormals. In E. S. Kite (Ed.), *The Development of Intelligence in Children (The Binet-Simon scale)*. Baltimore: Williams & Wilkins, 1916.

Bird, A., & Tobin, E. "Natural Kinds", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/natural-kinds/>.

Block, N. (1995). How heritability misleads about race. *Cognition, 56,* 99-128.

Boakes, R. (1984). *From Darwin to behaviourism: Psychology and the minds of animals*. Cambridge, UK: CUP.

Boomsma, D., Busjahn, A., & Peltonen, L. (2002). Classical twin studies and beyond. *Nature reviews. Genetics*, *3*(11), 872-882.

Bouchard, T. J. (2004). Genetic Influence on Human Psychological Traits. A Survey. *Current Directions in Psychological Science, 13*(4), 148-151.

Bouchard, T. J., & McGue, M. (2003). Genetic and environmental influences on human psychological differences. *Journal of neurobiology*, *54*(1), 4-45.

Bouchard, T. J., Lykken D. T., McGue, M., et al. (1990). Sources of Human Psychological Differences: The Minnesota Study of Twins Reared Apart. *Science, 250*(4978), 223-228.

Bourrat, P., Lu, Q., & Jablonka, E. (2017). Why the missing heritability might not be in the DNA. *BioEssays 39*, 1700067.

Boyd, R. (1991). Realism, Anti-Foundationalism and the Enthusiasm for Natural Kinds. *Philosophical Studies, 61*(1/2), 127-148.

Boyd, R. (1999). Homeostasis, Species, and Higher Taxa. In R.A. Wilson (Ed.), *Species: New Interdisciplinary Essays* (pp. 141-185). Cambridge, US: MIT Press.

Brenner, S. (1974). The genetics of Caenorhabditis elegans. *Genetics*, *77*(1), 71-94.

Brody, N. (2000). g and the one-many problem: is one enough? *Novartis Foundation Symposium*, 233, 122-129.

Bronfenbrenner, U., & Ceci, S. J. (1994). Nature-nurture reconceptualized in developmental perspective: A bioecological model. *Psychological review*, *101*(4), 568-586.

Burian, R. M., & Kampourakis, K. (2013). Against "Genes For": Could an Inclusive Concept of Genetic Material Effectively Replace Gene Concepts?. In K. Kampourakis (Ed.), *The philosophy of biology*. A Companion for Educators (pp. 597-628). Springer.

Burkart, J. M., Schubiger, M. N., & van Schaik, C. P. (2017). The evolution of general intelligence. *Behavioral and Brain Sciences*, *40, 1-24*.

Butcher, L. M., Davis, O. S. P., Craig, I. W., & Plomin, R. (2008). Genome-wide quantitative trait locus association scan of general cognitive ability using pooled DNA and 500k single nucleotide polymorphism microarrays. *Genes, Brain and Behavior*, *7*, 435-446.

Cairns, R. B., Gariepy, J. L., & Hood, K. E. (1990). Development, microevolution, and social behavior. *Psychological Review*, *97*(1), 49-65.

Callier, S. L., & Bonham, V. L. (2015). Taking a Stand: The Genetics Community's Responsibility for Intelligence Research. *Hastings Center Report*, *45*(S1), S54-S58.

Campbell, J. (2007). An interventionist approach to causation in psychology. *Causal learning: Psychology, philosophy, and computation*, 58-66.

Campbell, J. K., O'Rourke, M., & Slater, M. H. (Eds.). (2011). *Carving nature at its joints. Natural kinds in metaphysics and science*. Cambridge, MA, & London: MIT Press.

Canivez, G. L. (2013). Psychometric versus actuarial interpretation of intelligence and related aptitude batteries. In D. H. Saklofske, C. R. Reynolds, & V. L. Schwean (Eds.), *The oxford handbook of child psychological assessments* (pp. 84-112). New York: Oxford University Press.

Carlborg, O., & Haley, C. S. (2004). Epistasis: too often neglected in complex trait studies?. *Nature reviews. Genetics*, *5*(8), 618-625.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, UK: CUP.

Cattell, J. M. (1890). Mental tests and measurements. *Mind, 15*, 373-380.

Cattell, R. B. (1948). The Fate of National Intelligence; tests of a thirteen years prediction. *Eugenics Review, 42*(3), 136-148.

Cattell, R. B. (1971). *Abilities: Their structure, growth and measurement.* Boston: Houghton-Mifflin.

Cattell, R. B. (1982). *The Inheritance of Personality and Ability*. New York: Academic Press.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., et al. (2012). Most reported genetic association with general intelligence are probably false positive. *Psychological Science, 20*(10), 1-10.

Chabris, C. F., Lee, J. J., Benjamin, D. J., et al. (2013). Why it is hard to find genes associated with social science traits: theoretical and empirical considerations. *American journal of public health*, *103*(S1), S152-S166.

Chabris, C. F., Lee, J. J., Cesarini, D., et al. (2015). The fourth law of behavior genetics. *Current Directions in Psychological Science*, *24*(4), 304-312.

Champagne, F. A., & Mashoodh, R. (2009). Genes in context: Gene–environment interplay and the origins of individual differences in behavior. *Current Directions in Psychological Science*, *18*(3), 127-131.

Charney, E. (2008). Genes and ideologies. *Perspectives on Politics*, *6*(2), 299-319.

Cheung, V. G., Conlin, L. K., Weber, T. M., et al. (2003). Natural variation in human gene expression assessed in lymphoblastoid cells. *Nature Genetics, 35,*422-425.

Cianciolo, A. T., & Sternberg, R. J. (2004). *Intelligence. A Brief History*. Oxford: Blackwell.

Cock, A. G. (1973). William Bateson, Mendelism and Biometry. *Journal of the History of Biology*, *6*(1), 1-36.

Cole, M. (2006). *The Illusion of Culture-free intelligence Testing*. San Diego: University of California.

Cole, M., Gay, J., Glick, J. A., & Sharp, D. W. (1971). *The cultural context of learning and thinking: an exploration in experimental anthropology*. New York: Basic Books.

Conway, A. R. A., Cowan, N., Bunting, M. F., et al. (2002). A latent variable analysis of working memory capacity, short-term memory capacity, processing speed, and general fluid intelligence. *Intelligence*, *30*, 163-183.

Conway, A. R. A., Getz, S. J., Macnamara, B., & Engel de Abreu, P. M. J. (2011). Working memory and intelligence. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence (pp. 394-418)*. New York: CUP.

Conway, A. R., Kane, M. J., & Engle, R. W. (2003). Working memory capacity and its relation to general intelligence. *Trends in cognitive sciences*, *7*(12), 547-552.

Cook Jr, E. H., & Scherer, S. W. (2008). Copy-number variations associated with neuropsychiatric conditions. *Nature*, *455*(7215), 919-923

Cordell, H. J. (2002). Epistasis: what it means, what it doesn't mean, and statistical methods to detect it in humans. *Human molecular genetics*, *11*(20), 2463-2468.

Cowan, R. S. (1972a). Francis Galton's contribution to genetics. *Journal of the History of Biology*, *5*(2), 389-412.

Cowan, R. S. (1972b). Francis Galton's statistical ideas: the influence of eugenics. *Isis*, *63*(4), 509-528.

Craver, C. (2002). Interlevel Experiments and Multilevel Mechanisms in the Neuroscience of Memory. *Philosophy of Science, 69*(3), S83-S97.

Craver, C. (2009). Mechanisms and natural kinds. *Philosophical Psychology, 22,* 575-594.

Craver, C. (2015). Levels. In *Open MIND*. Frankfurt am Main: MIND Group.

Craver, Carl and Tabery, James, "Mechanisms in Science", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/science-mechanisms/>.

Crow, J. F. (1990). How Important is Detecting Interaction?. *Behavioral and Brain Sciences, 13*, 126-127.

Crusio, W. E. (1990). Estimating heritabilities in quantitative behavior genetics: A station passed. *Behavioral and Brain Sciences, 13,* 127-128.

Dar-Nimrod, I., & Heine, S. J. (2011). Genetic essentialism: on the deceptive determinism of DNA. *Psychological bulletin*, *137*(5), 800-818.

Darwin, C. (1859). The origin of species by means of natural selection: or, the preservation of favored races in the struggle for life, 2009.

Das, J. P., Naglieri, J. A, & Kirby, J. R. (1994). *Assessment of cognitive process: The PASS theory of intelligence*. Boston: Allyn & Bacon.

Davis, O. S. P., Butcher, L. M., Docherty, S. J., et al. (2010). A three-stage genome-wide association study of general cognitive ability: Hunting the small effects. *Behavior Genetics, 40*, 759-767.

Dawkins, R. (1982). *The Extended Phenotype*. Oxford: OUP.

Deary, I. J. (2002). g and cognitive elements of information processing: An agnostic view. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 151-181). Mahwah: Lawrence Erlbaum.

Deary, I. J., & Caryl, P. G. (1997). Neuroscience and human intelligence differences. *Trends in Neurosciences*, *20*(8), 365-371.

Detterman, D. K. (2002). General intelligence: Cognitive and biological explanations. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 223-243). Mahwah: Lawrence Erlbaum.

Di Franco, G., & Marradi, A. (2003). *Analisi fattoriale e analisi in componenti principali*. Roma: Bonanno.

Dickson, S. P., Wang, K., Krantz, I., et al. (2010). Rare variants create synthetic genome-wide associations. *PLoS biology*, *8*(1), e1000294.

Dobzhansky, T. (1970). *Genetics of the evolutionary process*. New York & London: Columbia University Press.

Downes, S. M., "Heritability", *The Stanford Encyclopedia of Philosophy* (Spring 2017 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2017/entries/heredity/>.

Duncan, J., Burgess, P. & Emslie, H. (1995). Fluid intelligence after frontal lobe lesions. *Neuropsychologia, 33*(3), 261-268.

Duncan, J., Seitz, R. J., Kolodny, J., et al. (2000). A neural basis for general intelligence. *Science, 289*(5478), 457-460.

Dupré, J. (2013). I—Living Causes. *Aristotelian Society Supplementary Volume, 87*, 19-37.

Dupré, J., & Nicholson, D. J. (forthcoming). *Everything Flows. Toward a Process Philosophy of Biology*. OUP.

East, E. M. (1910). A Mendelian interpretation of variation that is apparently continuous. *The American Naturalist*, *44*(518), 65-82.

Edgeworth, F. (1893). Statistical Correlation Between Social Phenomena. *Journal of the Royal Statistical Society, 56*(4), 670-675.

Eichler, E. E., Flint, J., Gibson, G., et al. (2010). Missing heritability and strategies for finding the underlying causes of complex disease. *Nature reviews. Genetics*, *11*(6), 446-450.

Eley, T. C., & Rijsdijk, F. (2005). Introductory guide to the statistics of molecular genetics. *Journal of Child Psychology and Psychiatry*, *46*(10), 1042-1044.

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. (1999). Working memory, short-term memory, and general fluid intelligence: a latent-variable approach. *Journal of experimental psychology: General*, *128*(3), 309.

Esping, A., & Plucker, J. A. (2015). Alfred Binet and the Children of Paris. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 153-161). New York: Springer.

Evans, J. S. B. T., & Frankish, K. (2008). *In Two Minds: Dual Processes and Beyond*. New York: Oxford University Press.

Eyre-Walker, A. (2010). Genetic architecture of a complex trait and its implications for fitness and genome-wide association studies. *Proceedings of the National Academy of Sciences*, *107*(S1), 1752-1756.

Eysenck, H. J., & Kamin, L. (1981). *Intelligence: The battle for the mind*. London: Macmillan.

Feldman, R. S. (2015). *Essentials of Understanding Psychology* (11th ed.). New York: McGraw-Hill.

Fentress, J. C. (1981). Order in ontogeny: Relational dynamics. In K. Immelmann, G. W. Barlow, L. Petrinovich & M. Main (Eds.), *Behavioral Development. The Bielefeld Interdisciplinary Project* (pp. 338-371). Cambridge: CUP.

Ferraguti, M., & Castellacci, C. (Eds.). (2011). *Evoluzione. Modelli e processi*, Milano-Torino: Pearson.

Fisher, R. A. (1918). The Correlation Between Relatives on the Supposition of Mendelian Inheritance. *Transactions of the Royal Society of Edinburgh, 52,* 399-433.

Fisher, R. A. (1919). The causes of human variability. *Eugenics Review*.

Fisher, R. A., Immer, F. R., & Tedin, O. (1932). The genetical interpretation of statistics of the third degree in the study of quantitative inheritance. *Genetics*, *17*(2), 107.

Fisher, S. E. (2006). Tangled webs: Tracing the connections between genes and cognition. *Cognition*, *101*(2), 270-297.

Floyd, R. G., & Kranzler, J. H. (2012). Processing approaches to interpretation of information from cognitive ability tests: A critical review. In D. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (3rd ed.) (pp. 497-523). New York: Guilford Press.

Fodor, J., & Piattelli Palmarini, M. (2010). *What Darwin got wrong*. New York: Farrar, Strauss & Giroux.

Fonagy, P., Gergely, G., & Jurist, E. L. (Eds.). (2004). *Affect regulation, mentalization and the development of the self*. London: Karnac books.

Fowler, J. H., Baker, L. A., & Dawes, C. T. (2008). Genetic variation in political participation. *American Political Science Review*, *102*, 233-248.

Fraga, M. F. et al. (2005). Epigenetic differences arise during the lifetime of monozygotic twins. *Proceedings of the National Academy of Sciences USA, 102*, 10604-10609.

Franklin-Hall, L. R. (2015). Natural kinds as categorical bottlenecks. *Philosophical Studies*, *172*, 925-948.

Fuller, J. L., & Thompson, W. R. (1960). *Behavior genetics.* New York: John Wiley & Sons.

Fuller, J. L., & Thompson, W. R. (1978). *Foundations of behavior genetics.* The C.V. Mosby Company

Fuller, T., Sarkar, S., & Crews, D. (2005). The use of norms of reaction to analyze genotypic and environmental influences on behavior in mice and rats. *Neuroscience & Biobehavioral Reviews*, *29*(3), 445-456.

Furnham, A. (2015). Intelligence: One Thing or Many? <https://www.psychologytoday.com/blog/sideways-view/201508/intelligence-one-thing-or-many>

Futuyma, D. J. (1986). *Evolutionary Biology*. Sunderland, MA: Sinauer Associates.

Galton, F. (1874). *English Men of Science. Their Nature and Nurture.* London: Macmillan & Co.

Galton, F. (1883). *Inquiry into human faculty and its development*. London: Macmillan Press.

Gardner, H. (1983). *Frames of mind: The theory of multiple intelligences*. New York: Basic books.

Garlick, D. (2002). Understanding the nature of the general factor of intelligence: The role of individual differences in neural plasticity as an explanatory mechanism. *Psychological Review, 109*, 116-136.

Gelman, S. A. (2004). Psychological essentialism in children. *Trends in Cognitive Sciences, 8*, 404-409.

Gibbard, A. (2001). Genetic plans, genetic differences, and violence: some chief possibilities. In D. Waddermann & R. Wachbroit *(Eds.), Genetics and Criminal Behaviour* (pp. 169-197). Cambridge: CUP.

Gifford, F. (1990). Genetic traits. *Biology and Philosophy*, *5*(3), 327-347.

Gifford, F. (2002). Understanding genetic causation and its implications for ethical issues concerning medical genetics. In *Mutating Concepts, Evolving Disciplines: Genetics, Medicine, and Society* (pp. 109-125). Springer Netherlands.

Gignac, G. E., & Watkins, M. W. (2015). There may be nothing special about the association between working memory capacity and fluid intelligence. *Intelligence*, *52*, 18-23.

Glennan, S. (2002). Rethinking mechanistic explanation. *Philosophy of science*, *69*(S3), S342-S353.

Glutting, J. J., Watkins, M. W., Konold, T. R., & McDermott, P. A. (2006). Distinctions without a difference: The utility of observed versus latent factors from the WISC-IV in estimating reading and math achievement on the WIAT-II. *Journal of Special Education, 40*, 103-114.

Goddard, H. H. (1917). Mental Tests and the Immigrant. *Journal of Delinquecy, 2*, 243-277.

Godfrey-Smith, P. (1999). Genes and codes: lessons from the philosophy of mind?. In V. Hardcastle (Ed.), *Where Biology Meets Psychology: Philosophical Essays* (pp. 305-331). Cambridge, MA: MIT Press.

Godfrey-Smith, Peter and Sterelny, Kim, "Biological Information", *The Stanford Encyclopedia of Philosophy* (Summer 2016 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/sum2016/entries/information-biological/>.

Goldstein, D. B. (2009). Common genetic variation and human traits. *New England Journal of Medicine*, *360*(17), 1696-1698.

Goldstein, S., Princiotta, D., & Naglieri, J. A. (Eds.). (2015). *Handbook of intelligence. Evolutionary theory, historical perspective, and current concepts*. New York: Springer.

Goleman, D. P. (1995). *Emotional intelligence: Why it can matter more than IQ for character, health and lifelong achievement*. New York: Bantam Books.

Gollin, E.S. (1985). Ontogeny, phylogeny, and causality. In E.S. Gollin (Ed.), *The comparative development of adaptive skills: Evolutionary implications* (pp. 1-17). Hillsdale: Erlbaum.

Gottlieb, G. (1991). Experiential canalization of behavioral development: Theory. *Developmental psychology*, *27*(1), 4-13.

Gottlieb, G. (1992). *Individual development and evolution: The genesis of novel behavior*. New York: Oxford University Press.

Gottlieb, G. (1995). Some conceptual deficiencies in 'developmental' behavior genetics. *Human Development*, *38*(3), 131-141.

Gottlieb, G. (2003). On making behavioral genetics truly developmental. *Human Development, 46*(6), 337-355.

Gould, S. J. (1981). *The Mismeasure of Men*. New York: Norton & Co.

Grant, S. G. (2003). Systems biology in neuroscience: bridging genes to cognition. *Current Opinion in Neurobiology*, *13*, 577-582.

Gray, J. R., & Thompson, P. M. (2004). Neurobiology of intelligence: science and ethics. *Nature reviews. Neuroscience*, *5*(6), 471.

Gray, R. (1992). Death of the gene: Developmental systems strike back. In *Trees of life* (pp. 165-209). Springer Netherlands.

Greenwood, J. D. (2015). Intelligence defined: Wundt, James, Cattell, Thorndike, Goddard, and Yerkes. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 123-135). New York: Springer.

Griesemer, J. R. (2000). Reproduction and the reduction of genetics. In P. Beurton, R. Falk & H. J. Rheinberger (Eds.), *The concept of the gene in development and evolution: Historical and epistemological perspectives* (pp. 240-285). Cambridge, UK: CUP.

Griffing, B. (1950). Analysis of quantitative gene action by constant parent regression and related techniques. *Genetics*, *35*(3), 303.

Griffiths, P. E. (2002). What is Innateness? *The Monist, 85*(1), 70-85.

Griffiths, P. E. (2004). Emotions as natural and normative kinds. *Philosophy of science*, *71*(5), 901-911.

Griffiths, P. E., & Gray, R. D. (1994). Developmental systems and evolutionary explanation. *The Journal of Philosophy*, *91*(6), 277-304.

Griffiths, P., & Stotz, K. (2013). *Genetics and philosophy: An introduction*. Cambridge, UK: CUP.

Grudnik, J. L., & Kranzler, J. H. (2001). Meta-analysis of the relationship between intelligence and inspection time. *Intelligence, 29*, 523-535.

Gudbjartsson, D. F., Walters, D. F., Thorleifsson, H. S., et al. (2008). Many sequence variants affecting diversity of adult human height. *Nature Genetics*, *40*, 609-615.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill.

Guilford, J. P. (1982). Cognitive psychology's ambiguities: Some suggested remedies. *Psychological Review, 89*, 48-59.

Hacking, I. (1991). A Tradition of Natural Kinds. *Philosophical Studies, 61*(1/2), 109-126.

Haier, R. J. (2003). Brain imaging studies of intelligence: Individual differences and neurobiology. In R. J. Sternberg, J. Lautrey & T. I. Lubart (Eds.), *Models of intelligence: International perspectives*. Washington, DC: APA.

Haier, R. J., Karama, S., Colom, R., et al. (2014). A comment on "fractionating intelligence" and the peer review process. *Intelligence*, *46*, 323-332.

Haier, R. J., White, N. S., & Alkire, M. T. (2003). Individual differences in general intelligence correlate with brain function during nonreasoning tasks. *Intelligence*, *31*(5), 429-441.

Hampshire, A., Highfield, R. R., Parkin, B. L., & Owen, A. M. (2012). Fractionating human intelligence. *Neuron*, *76*(6), 1225-1237.

Hasher, L., Lustig, C., & Zacks, R. (2007). Inhibitory mechanisms and the control of attention. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. Towse (Eds.), *Variation in working memory* (pp. 227-249). New York: OUP.

Haslam, N. (2014). Natural Kinds in Psychiatry: Conceptually Implausible, Empirically Questionable, and Stigmatizing. In H. Kincaid & J. A. Sullivan (Eds.), *Classifying Psychopathology. Mental kinds and natural kinds* (pp. 11-28). Cambridge, US: MIT Press.

Hemani, G., Knott, S., & Haley, C. (2013). An evolutionary perspective on epistasis and the missing heritability. *PLoS genetics*, *9*(2), e1003295.

Herrnstein, R. J. (1973). *IQ in the meritocracy*. Boston: Atlantic Monthly Press.

Herrnstein, R. J., & Murray, C. A. (1994). *The bell curve: Intelligence and class structure in American life*. New York: Free Press.

Hibberd, F. J. (2014). The metaphysical basis of a process psychology. *Journal of Theoretical and Philosophical Psychology*, *34*(3), 161-186.

Hirsch, J. (1997). Some history of heredity-vs-environment, genetic inferiority at Harvard (?), and The (incredible) Bell Curve. *Genetica, 99,* 207-224.

Hirschhorn, J. N., Lohmueller, K., Byrne, E., & Hirschhorn, K. (2002). A comprehensive review of genetic association studies. *Genetics in Medicine*, *4*(2), 45-61.

Hood, K. E., Halpern, C. T., Greenberg, G., & Lerner, R. M. (Eds.). (2010). *Handbook of Developmental Science, Behavior, and Genetics.* Wiley-Blackwell.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized general intelligences. *Journal of Educational Psychology, 57,* 253–270.

Humphreys, L. G., & Stark, S. (2002). General Intelligence: Measurement, Correlates, and Interpretations of the Cultural-Genetic Construct. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 87-116). Mahwah: Lawrence Erlbaum.

Hunt, E. (1980). Intelligence as an information-processing concept. *British Journal of Psychology, 71,* 449-474.

Hunt, E. (2011). *Human intelligence*. New York: CUP.

Jablensky, A. (2006). Subtyping schizophrenia: implications for genetic research. *Molecular Psychiatry, 11,* 815-836.

Jacquard, A. (1983). Heritability: One Word, Three Concepts. *Biometrics, 39*(2), 465-477.

Jamieson, A., & Radick, G. (2013). Putting Mendel in his place: how curriculum reform in genetics and counterfactual history of science can work together. In K. Kampourakis (Ed.), *The philosophy of biology*. A Companion for Educators (pp. 577-595). Springer.

Jencks, C. (1980). Heredity, Environment, and Human Policy Reconsidered. *American Sociological Review, 45,* 723-736.

Jensen, A. R. (1969). How Much Can We Boost IQ and Scholastic Achievement? *Harvard Educational Review, 39*, 1-123.

Jensen, A. R. (1980). *Bias in Mental Testing*. New York: The Free Press.

Jensen, A. R. (1983). *Straight talk about mental tests*. London: Methuen.

Jensen, A. R. (1984). Test validity: g versus the specificity doctrine. *Journal of Social Biological Structures, 7*, 93-118.

Jensen, A. R. (1998). *The* g *factor: The science of mental ability*. Westport, CT: Praeger.

Jensen, A. R. (2002). Psychometric g: Definition and Substantiation. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 39-54). Mahwah: Lawrence Erlbaum.

Johannsen, W. (1903). Heredity in populations and pure lines. *Classic Papers in Genetics*, 20-26.

Johannsen, W. (1911). The Genotype Conception of Heredity. *The American Naturalist 45*(531), 129–159.

Johnson, K. A., & Goody, R. S. (2011). The original Michaelis constant: translation of the 1913 Michaelis–Menten paper. *Biochemistry*, *50*(39), 8264-8269.

Johnston, T. D. (1988). Developmental explanation and the ontogeny of birdsong: Nature/nurture redux. *Behavioral and Brain Sciences*, *11*(4), 617-630.

Jordan, B. (2000). *Les imposteurs de la génétique*. Paris: Editions du Seuil.

Joseph, J. (2001). Separated twins and the genetics of personality differences: A critique. *The American journal of psychology*, *114*(1), 1.

Joseph, J. (2004). *The gene illusion: genetic research in psychiatry and psychology under the microscope*. New York: Algora.

Joseph, J. (2006). *The missing gene: Psychiatry, heredity, and the fruitless search for genes*. New York: Algora Publishing.

Juel-Nielsen, N. (1965). Individual and environment: A psychiatric-psychological investigation of monozygotic twins reared apart. *Acta Psychiatrica and Neurologica Scandinavica*, 183, 1980.

Jung, R. E., & Haier, R. J. (2007). The Parieto-Frontal Integration Theory (P-FIT) of intelligence: converging neuroimaging evidence. *Behavioral and Brain Sciences*, *30*(2), 135-154.

Kane, M. J., & Engle, R.W. (2002). The role of prefrontal cortex in working-memory capacity, executive attention, and general fluid intelligence: An individual differences perspective. *Psychonomic Bulletin and Review, 9*(4), 637-671.

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, C. Jarrold, M. J. Kane, A. Miyake & J. Towse (Eds.), *Variation in working memory* (pp. 21-48). New York: OUP.

Karmiloff-Smith, A. (1992). *Beyond Modularity*. Cambridge, MA: MIT Press/Bradford Books.

Kaufman, A. S. (2009). *IQ testing 101*. New York: Springer.

Kaufman, A. S., & Kaufman, N. L. (1983). *KABC: Kaufman Assessment Battery for Children*. Circle Pines, MN: American Guidance Service.

Kaufman, A. S., & Kaufman, N. L. (1993). *Kaufman Adolescent and Adult Intelligence Test (KAIT)*. Circle Pines: American Guidance Service.

Kaufman, J. C., Kaufman, S. B., & Plucker, J. A. (2013). Contemporary theories of intelligence. In D. Reisberg (Ed.), *The Oxford Handbook of Cognitive Psychology* (pp. 811-822). OUP.

Kaufman, S. B. (2011). Intelligence and the cognitive unconscious. In R. J. Sternberg & S. B. Kaufman (Eds.), *The Cambridge handbook of intelligence* (pp. 442-467). New York: CUP.

Kaufman, S. B., DeYoung, C. G., Gray, J. R., et al. (2009). Associative learning predicts intelligence above and beyond working memory and processing speed. *Intelligence*, *37*, 374-382.

Kempthorne, O. (1978). A Biometrics Invited Paper: Logical, Epistemological and Statistical Aspects of Nature-Nurture Data Interpretation. *Biometrics, 34*(1), 1-23.

Kempthorne, O. (1997). Heritability: Uses and abuses. *Genetica*, *99*(2), 109-112.

Kendler, K. S. (2005). Psychiatric genetics: a methodologic critique. *American Journal of Psychiatry*, *162*(1), 3-11.

Kendler, K. S. (2012). Levels of explanation in psychiatric and substance use disorders: implications for the development of an etiologically based nosology. *Molecular psychiatry*, *17*(1), 11-21.

Kendler, K. S., Zachar, P., & Craver, C. (2011). What kinds of things are psychiatric disorders? *Psychological Medicine, 41*, 1143-1150.

Khalidi, M. A. (2014). Naturalizing Kinds. In B. Bashour & H. D. Muller (Eds.), *Contemporary Philosophical Naturalism and Its Implications* (pp. 115-135), New York: Routledge.

Khalidi, M. A. (2015). Innateness as a Natural Cognitive Kind. *Philosophical Psychology, 29*(3), 319-333.

Kincaid, H., & Sullivan, J. A. (Eds.). (2014). *Classifying Psychopathology. Mental Kinds and Natural Kinds.* Cambridge (US): MIT Press.

Kitcher, P. (2001). Battling the Undead. *How (and How Not) to Resist Genetic Determinism*. In R. S. Singh, C. B. Krimbas, D. B. Paul & J. Beatty (Eds.), *Thinking About*

Evolution: Historical, Philosophical, and Political Perspectives (pp. 396-414). New York: CUP.

Kraft, P., & Hunter, D. J. (2009). Genetic risk prediction--are we there yet?. *The New England journal of medicine*, *360*(17), 1701-1703.

Kray, J., & Frensch, P. (2002). A View From Cognitive Psychology: *g*–(G)host in the Correlation Matrix? In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 183-220). Mahwah: Lawrence Erlbaum.

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?!. *Intelligence*, *14*(4), 389-433.

Lander, E. S. (2011). Initial impact of the sequencing of the human genome. *Nature*, *470*(7333), 187-197.

Lautrey, J. (2002). Is there a general factor of cognitive development?. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 117-148). Mahwah: Lawrence Erlbaum.

Layzer, D. (1974). Heritability Analyses of IQ Scores: Science or Numerology?. *Science, 183*(131), 1259-1266.

Lee, J. J., & Chow, C. C. (2014). Conditions for the validity of SNP-based heritability estimation. *Human genetics*, *133*(8), 1011-1022.

Lehrman, D. S. (1953). A critique of Konrad Lorenz's theory of instinctive behavior. *The Quarterly Review of Biology*, *28*(4), 337-363.

Lehrman, D. S. (1970). Semantic and conceptual issues in the nature-nurture problem. In L. Tobach, E. Aronson, D. S. Lehrman & J. Rosenblat (Eds.), *Development and Evolution of Behavior* (pp. 17-50). San Francisco: Freeman and Company.

Lettre, G., Jackson, A. U., Gieger, C., et al. (2008). Identification of ten loci associated with height highlights new biological pathways in human growth. *Nature Genetics*, *5*, 584-591.

Lewontin, R. (1974). Annotation: the analysis of variance and the analysis of causes. *American journal of human genetics*, *26*(3), 400-411.

Lewontin, R. (1982). *Human Diversity*. New York: Freeman & Co.

Lewontin, R. (2001). *The triple helix: Gene, organism, and environment*. Harvard University Press.

Ludwig, D. (2015). *A Pluralist theory of the mind*. Dordrecht: Springer.

Luria, A. R. (1966). *Human brain and psychological processes*. New York: Harper & Row.

Luria, A. R. (1970). The functional organization of the brain. *Scientific American*, *222*, 66-78.

Luria, A. R. (1973). *The working brain: An introduction to neuropsychology*. London: Penguin.

Lush, J. L. (1945). *Animal breeding plans.* Ames, IA: Collegiate Press.

Lush, J. L. (1949). Heritability of quantitative characters in farm animals. *Hereditas*, *35*(S1), 356-375.

Lynch, K. E., & Bourrat, P. (2017). Interpreting heritability causally. *Philosophy of Science*, *84*(1), 14-34.

Machamer, P. K., Darden, L., & Craver, C. (2000). Thinking about mechanisms. *Philosophy of Science, 67*, 1-25.

Mackie, J. L. (1974). *The cement of the universe: a study of causation*. Clarendon.

Mackintosh, N. J. (2011). *IQ and human intelligence* (2nd ed.). Oxford: OUP.

MacLullich, A. M. J., Ferguson, K. J., Deary, I. J., Seckl, J. R., Starr, J. M. & Wardlaw, J. M. (2002). Intracranial capacity and brain volumes are associated with cognition in healthy elderly men. *Neurology, 59,* 169-174.

Magnus, P. D. (2012). *Scientific Enquiry and Natural Kinds. From Planets to Mallards*. Basingstoke: Palgrave Macmillan.

Maher, B. (2008). Personal genomes: The case of the missing heritability. *Nature News*, *456*(7218), 18-21.

Manolio, T. A., Collins, F. S., Cox, N. J., et al. (2009). Finding the missing heritability of complex diseases. *Nature*, *461*(7265), 747.

Marcus, G. (2004). *The birth of the mind: How a tiny number of genes creates the complexities of human thought*. New York: Basic Books.

Mather, K. (1941). Variation and selection of polygenic characters. *Journal of Genetics*, *41*(1), 159-193.

Mather, K. (1943). Polygenic Inheritance and Natural Selection. *Biological Reviews, 18*(1), 32-64.

Mather, K. (1964). *Human Diversity. The Nature and Significance of Differences Among Men*. Edinburgh: Oliver & Boyd.

Mayr, E. (1963) *Animal Species and Evolution.* Cambridge, MA: Harvard University Press.

McCarthy, M. I., & Hirschhorn, J. N. (2008). Genome-wide association studies: potential next steps on a genetic journey. *Human molecular genetics*, *17*(R2), R156-R165.

McClellan, J., & King, M. C. (2010). Genetic heterogeneity in human disease. *Cell*, *141*(2), 210-217.

McDaniel, M. A. (2005). Big-brained people are smarter: A meta-analysis of the relationship between in vivo brain volume and intelligence. *Intelligence*, *33*(4), 337-346.

McDonald, J. H. (2012). Myths of Human Genetics. Introduction to the Myths. <http://udel.edu/~mcdonald/mythintro.html>

McGue, M. (1989). Nature-nurture and intelligence. *Nature*, *340*, 507-508.

McGue, M., & Lykken, D. T. (1992). Genetic influence on risk of divorce. *Psychological Science*, *3*, 368-373.

McGuire, T. R., & Hirsch, J. (1977). General intelligence (g) and heritability (H2, h2). In I. Uzgiris & F. Weitzmann (Eds.), *The structuring of experience* (pp. 25-72). New York: Plenum Press.

McRae, A. F., Matigian, N. A., Vadlamudi, L., et al. (2007). Replicated effects of sex and genotype on gene expression in human lymphoblastoid cell lines. *Human Molecular Genetics, 16,* 364-373.

Medin, D. L. (1989). Concepts and Conceptual Structure. *American Psychologist, 44*(12), 1469-1481.

Mendel, G. (1866). Versuche über Pflanzen-Hybriden. Verhandlungen des Naturforschenden Vereines in Brünn, IV.

Michel, G. F. (2010). Experience, and Learning in Behavioral Development. In K. E. Hood, C. T. Halpern, G. Greenberg & R. M. Lerner (Eds.), *Handbook of Developmental Science, Behavior and Genetics* (pp. 123-165). Wiley-Blackwell.

Michell, J. (2009). Invalidity in validity. In R. W. Lissitz (Ed.), *The concept of validity. Revisions, new directions and applications* (pp. 111-133). Charlotte, NC: Information Age Publishing.

Michell, J. (2011). Qualitative research meets the ghost of Pythagoras. *Theory & Psychology, 21,* 241-259.

Monks, S. A., Leonardson, A., Zhu, H., et al. (2004). Genetic inheritance of gene expression in human cell lines. *American Journal of Human Genetics, 15*,1094-1105.

Morrison, M. (2007). The development of population genetics. In M. Matthen & C. Stephens (Eds.), *Philosophy of Biology* (pp. 309-333). Amsterdam: Elsevier.

Naghavi, H. R., & Nyberg, L. (2005). Common fronto-parietal activity in attention, memory, and consciousness: Shared demands on integration?. *Consciousness and Cognition, 14*(2), 390-425.

Naglieri, J. A. (2015). Hundred years of intelligence testing: Moving from traditional IQ to second-generation intelligence tests. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 295-316). New York: Springer.

Naglieri, J. A., & Das, J. P. (1997). *Cognitive assessment system*. Itasca: Riverside.

Naglieri, J. A., & Das, J. P. (2002). Practical Implications of General Intelligence and PASS Cognitive Processes. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 55-84). Mahwah: Lawrence Erlbaum.

Neisser, U., Boodoo, G., Bouchard Jr, T. J., et al. (1996). Intelligence: Knowns and unknowns. *American psychologist*, *51*(2), 77-101.

Nelson, D. L., & Cox, M. M. (2002). *Lehningen Principles of Biochemistry* (6th ed.). Freeman & Company.

Nelson, R. M., Pettersson, M. E., & Carlborg, O. (2013). A century after Fisher: time for a new paradigm in quantitative genetics. *Trends in Genetics*, *29*(12), 669-676.

Newman, H. H., Freeman, F. N., & Holzinger, K. J. (1937). *Twins: a study of heredity and environment*. Chicago: Chicago University Press.

Nicholson, D. J. (2012). The concept of mechanism in biology. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *43*(1), 152-163.

Nilsson-Ehle, H. (1909). *Kreuzungsuntersuchungen an hafer und weizen* (Vol. 20, No. 2). H. Ohlssons buchdruckerei.

Noble, D. (2008). Genes and causation. *Philosophical Transactions of the Royal Society of London A: Mathematical, Physical and Engineering Sciences*, *366*(1878), 3001-3015.

Northcott, R. (2006). Causal efficacy and the analysis of variance. *Biology and Philosophy*, *21*(2), 253-276.

Norton, B. (1975). Metaphysics and population genetics: Karl Pearson and the background to Fisher's multi-factorial theory of inheritance. *Annals of Science*, *32*(6), 537-553.

Oftedal, G. (2005). Heritability and genetic causation. *Philosophy of Science, 72*(5), 699-709.

Ortiz, S. O. (2015). CHC theory of intelligence. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 209-227). New York: Springer.

Otero, T. M. (2015). Intelligence: Defined as Neurocognitive Processing. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 193-208). New York: Springer.

Owen, M. J., Craddock, N., & Jablensky, A. (2007). The genetic deconstruction of psychosis. *Schizophrenia bulletin*, *33*(4), 905-911.

Oyama, S. (1985). *The ontogeny of information. Developmental systems and evolution.* Cambridge: CUP.

Page, G. P., George, V., Go, R. C., et al. (2003). "Are we there yet?": Deciding when one has demonstrated specific genetic causation in complex diseases and quantitative traits. *American Journal of Human Genetics, 73,* 711-719.

Panofsky, A. (2014). *Misbehaving science: Controversy and the development of behavior genetics*. Chicago: University of Chicago Press.

Panse, V. G. (1940). A statistical study of quantitative inheritance. *Annals of Human Genetics*, *10*(1), 76-105.

Park, J. H., Wacholder, S., Gail, M. H., et al. (2010). Estimation of effect size distribution from genome-wide association studies and implications for future discoveries. *Nature Genetics*, *42*, 570-575.

Pearl, J. (2009). Causal inference in statistics: An overview. *Statistics surveys*, *3*, 96-146.

Pearson, C. H. (2007). Is heritability explanatorily useful?. *Studies in History and Philosophy of Science Part C: Studies in History and Philosophy of Biological and Biomedical Sciences*, *38*(1), 270-288.

Peterson, E. L. (2011). The excluded philosophy of evo-devo? Revisiting CH Waddington's failed attempt to embed Alfred North Whitehead's" organicism" in evolutionary biology. *History and philosophy of the life sciences*, 301-320.

Petronis, A. (2006). Epigenetics and twins: Three variations on the theme. *Trends in Genetics, 22,* 347-350.

Phillips, P. C. (2008). Epistasis—the essential role of gene interactions in the structure and evolution of genetic systems. *Nature reviews. Genetics*, *9*(11), 855-867.

Pigliucci, M. (2001). *Phenotypic Plasticity. Beyond Nature and Nurture*. Baltimore & London: The Johns Hopkins University Press.

Pintner, R. (1923). *Intelligence testing*. New York: Henry Holt.

Plomin, R. (1983). Introduction: Developmental behavioral genetics. *Child Development, 54* (2), 253-259.

Plomin, R. (1986). *Development, Genetics, and Psychology*. Hillsdale, NJ: Lawrence Erlbaum.

Plomin, R., Corley, R., DeFries, J. C., & Fulker, D. W. (1990). Individual differences in television viewing in early childhood: Nature as well as nurture. *Psychological Science*, *6*, 371-377.

Plomin, R., DeFries, J. C., Knopik, V. S., & Neiderhiser, J. N. (2013). *Behavioral Genetics* (6th ed.). New York: Worth Publishers.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2001). *Behavioral Genetics* (4th ed.). New York: Worth Publishers.

Plomin, R., DeFries, J. C., McClearn, G. E., & Rutter, M. (1997). *Behavioral genetics* (3rd ed.). New York: W. H. Freeman and Company.

Plomin, R., Haworth, C. M. A., & Davis, O. S. P. (2009). Common disorders are quantitative traits. *Nature Reviews Genetics, 10*, 872-878.

Plucker, J. A. (2016) L. L. Thurstone. <http://www.intelltheory.com/lthurstone.shtml>

Plucker, J. A., & Shelton, A. L. (2015). General intelligence (g): overview of a complex construct and its implications for genetics research. *Hastings Center Report*, *45*(S1), S21-S24.

Pretz, J. E., & Sternberg, R. J. (2005). Unifying the Field: Cognition and Intelligence. In R. J. Sternberg & J. E. Pretz (Eds.), *Cognition and Intelligence* (pp. 306-318). Cambridge, UK: CUP.

Princiotta, D., & Goldstein, S. (2015). AR Luria and Intelligence Defined as a Neuropsychological Construct. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 181-192). New York: Springer.

Purcell, S. (2013). Appendix. Statistical Methods in Behavioral Genetics. In R. Plomin, J. C. DeFries, V. S. Knopkin & J. N. Neiderhiser, *Behavioral Genetics* (6th ed.) (pp. 357-411). New York: Worth Publishers.

Radick, G. (2012). Should "Heredity" and "Inheritance" Be Biological Terms? William Bateson's Change of Mind as a Historical and Philosophical Problem. *Philosophy of Science*, *79*(5), 714-724.

Ramus, F. (2006). Genes, brain, and cognition: A roadmap for the cognitive scientist. *Cognition*, *101*(2), 247-269.

Ramus, F. (2017). General intelligence is an emerging property, not an evolutionary puzzle. *Behavioral and Brain Sciences*, *40*, 43-44.

Ratner, C. (2004). Genes and psychology in the news. *New Ideas in Psychology*, *22*(1), 29-47.

Reich, D. E., & Lander, E. S. (2001). On the allelic spectrum of human disease. *Trends in Genetics*, *17*(9), 502-510.

Rheinberger, Hans-Jörg, Müller-Wille, Staffan and Meunier, Robert, "Gene", *The Stanford Encyclopedia of Philosophy* (Spring 2015 Edition), Edward N. Zalta (ed.), URL = <https://plato.stanford.edu/archives/spr2015/entries/gene/>.

Richardson, K. (2000). *The Making of Intelligence*. New York: Columbia University Press.

Rietveld, C. A., Esko, T., Davies, G., et al. (2014). Common genetic variants associated with cognitive performance identified using the proxy-phenotype method. *Proceedings of the National Academy of Sciences*, *111*(38), 13790-13794.

Roll-Hansen, N. (1978). The genotype theory of Wilhelm Johannsen and its relation to plant breeding and the study of evolution. *Centaurus*, *22*(3), 201-235.

Rose, S. (1997). *Lifelines: Biology Beyond Determinism*. New York: OUP.

Rose, S. (2006). Commentary: heritability estimates—long past their sell-by date. *International Journal of Epidemiology*, *35*(3), 525-527.

Rose, S., Lewontin, R., & Kamin, L. (1984). *Not in Our Genes. Biology, Ideology and Human Nature*. New York: Pantheon Books.

Rusak, B., Robertson, H. A., Wisden, W., & Hunt, S. P. (1990). Light pulses that shift rhythms induce gene expression in the suprachiasmatic nucleus. *Science, 248,* 1237-1240.

Russo, F. & Williamson, J. (2007). Interpreting Causality in the Health Sciences. *International Studies in the Philosophy of Science, 21*(2), 157-170.

Rutter, M., Silberg, J., & Simonoff, E. (1993). Wither behavioral genetics? A developmental psychopathological perspective. In R. Plomin & G. McClearn (Eds.), *Nature, nurture, and psychology* (pp. 433-456). Washington, DC: APA.

Salthouse, T. A. (1996). The processing-speed theory of adult age differences in cognition. *Psychological Review, 103*, 403-428.

Samuels, R. (2004). Innateness in Cognitive Science. *Trends in Cognitive Sciences, 8*(3), 136-141.

Samuels, R. (2007). Is innateness a confused notion? In P. Carruthers, S. Laurence, & S. Stich (Eds.), *The innate mind: Foundations and the future* (Vol. 3, pp. 17-36). Oxford: OUP.

Sarkar, S. (1998). *Genetics and reductionism*. CUP.

Sarton, A. (1969). *L'Intelligence efficace*. Centre d'étude et de promotion de la lecture.

Scarr, S. (1992). Developmental theories for the 1990s: Development and individual differences. *Child Development*, *63*(1), 1-19.

Schaffner, K. F. (2016). *Behaving: What's Genetic, What's Not, and Why Should We Care?*. Oxford: OUP.

Schneider, W. J. (2013). Principles of assessment of aptitude and achievement. In D. Saklofske, C. Reynolds & V. Schwean (Eds.), *Oxford handbook of psychological assessment of children and adolescents* (pp. 286-330). New York: OUP.

Schneider, W. J., & Flanagan, D. P. (2015). The relationship between theories of intelligence and intelligence tests. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 317-340). New York: Springer.

Schoups, A. A., & Black, I. B. (1991). Visual experience specifically regulates synaptic molecules in rat visual cortex. *Journal of Cognitive Neuroscience, 3,* 252-257.

Scriver, C. R. (2007). The PAH gene, phenylketonuria, and a paradigm shift. *Human Mutation, 28*, 831-845.

Serpico, D. (2017). What Kind of Kind is Intelligence?. *Philosophical Psychology*, 1-21.

Sesardic, N. (1993). Heritability and causality. *Philosophy of Science*, *60*(3), 396-418.

Sharma, A., Sharma, V. K., Horn-Saban, S., et al. (2005). Assessing natural variations in gene expression in humans by comparing with monozygotic twins using microarrays. *Physiological Genomics, 21*, 117-123.

Shields, J. (1962). *Monozygotic twins. Brought up apart and brought up together.* Oxford, UK: OUP.

Slater, M. H., & Borghini, A. (2013). Introduction: Lessons from the scientific butchery. In J. K. Campbell, M. O'Rourke & M. H. Slater (Eds.), *Carving Nature at Its Joints. Natural Kinds in Metaphysics and Science* (pp. 1-31). Cambridge, MA, & London: MIT Press.

Slatkin, M. (2009). Epigenetic inheritance and the missing heritability problem. *Genetics*, *182*(3), 845-850.

Snustad, D. P., & Simmons, M. J. (2012). *Genetics* (6th ed.). New York: Wiley & Sons.

Sober, E. (1988). Apportioning Causal Responsibility. *Journal of Philosophy, 85*, 303-318.

Sober, E. (2001). Separating nature and nurture. In D. Waddermann & R. Wachbroit *(Eds.), Genetics and Criminal Behaviour* (pp. 47-78). Cambridge: CUP.

Spearman, C. (1904). 'General Intelligence', Objectively Determined and Measured. *The American Journal of Psychology*, *15*(2), 201-292.

Spearman, C. (1923). *The Nature of Intelligence and the principles of cognition*. London: Macmillan.

Spearman, C., & Wynn Jones, L. (1950). *Human Ability*. London: Macmillan.

Spuhler, J. N. (1954). Some problems in the study of quantitative inheritance in man. *American journal of human genetics*, *6*(1), 130-155.

Stadler, M., Becker, N., Gödker, M., Leutner, D., & Greiff, S. (2015). Complex problem solving and intelligence: A meta-analysis. *Intelligence*, *53*, 92-101.

Stankoff, L., & Roberts, R. D. (1997). Mental speed is not the ''basic'' process of intelligence. *Personality and Individual Differences*, *22,* 69-84.

Stankov, L. (2002). g: A Diminutive General. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 19-38). Mahwah: Lawrence Erlbaum.

Stearns, S. C., & Hoekstra, R. F. (2000). *Evolution, an introduction*. Oxford University Press.

Stein, D. J. (2014). Psychopharmacology and Natural Kinds: A Conceptual Framework. In H. Kincaid & J.A. Sullivan (Eds.), *Classifying Psychopathology. Mental kinds and natural kinds* (pp. 65-73). Cambridge, US: mit Press.

Stent, G. S. (1981). Strength and weakness of the genetic approach to the development of the nervous system. *Annual Review of Neuroscience*, *4*(1), 163-194.

Stephens, C. (2008). Population Genetics. In S Sarkar & A. Plutynski (Eds.), *A Companion to the Philosophy of Biology* (pp. 119-137). Blackwell.

Sternberg R. J., & Grigorenko, E. (Eds.). (2002). *The General Factor of Intelligence. How General Is It?* (2nd ed.). London: Psychology Press.

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: CUP.

Sternberg, R. J. (1997). *Successful intelligence*. New York: Plume.

Sternberg, R. J. (2002). Beyond g: The theory of successful intelligence. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 447-479). Mahwah: Lawrence Erlbaum.

Sternberg, R. J. (2015). Multiple intelligences in the new age of thinking. In S. Goldstein, D. Princiotta & J. A. Naglieri (Eds.), *Handbook of Intelligence. Evolutionary Theory, Historical Perspective and Current Concepts* (pp. 229-241). New York: Springer.

Sternberg, R. J., & Davidson, J. E. (1983). Insight in the gifted. *Educational Psychologist*, *18*(1), 51-57.

Sternberg, R. J., & Frensch, P. A. (1990). Intelligence and cognition. *Cognitive psychology. An international review*, 57-103.

Sternberg, R. J., & Pretz, J. E. (Eds.). (2005). *Cognition and intelligence: identifying the mechanisms of the mind*. Cambridge University Press.

Stoltenberg, S. F. (1997). Coming to terms with heritability. *Genetica, 99,* 89-96.

Stranger, B. E., Stahl, E. A., & Raj, T. (2011). Progress and promise of genome-wide association studies for human complex trait genetics. *Genetics*, *187*(2), 367-383.

Stringer, S., Derks, E. M., Kahn, R. S., Hill, W. G., & Wray, N. R. (2013). Assumptions and properties of limiting pathway models for analysis of epistasis in complex traits. *PLoS One*, *8*(7), e68913.

Sturm, R. A., & Frudakis, T. N. (2004). Eye colour: portals into pigmentation genes and ancestry. *Trends in Genetics*, *20*(8), 327-332.

Tabb, K. (2015). Psychiatric progress and the assumption of diagnostic discrimination. *Philosophy of Science*, *82*(5), 1047-1058.

Tabery, J. (2014). *Nature vs nurture*. <http://eugenicsarchive.ca/discover/encyclopedia/>

Tabery, J., & Griffiths, P. E. (2010). Historical and Philosophical Perspectives on Behavioral Genetics and Developmental Science. In K. E. Hood, C. T. Halpern, G. Greenberg & R. M. Lerner (Eds.), *Handbook of Developmental Science, Behavior and Genetics* (pp. 41-60). Wiley-Blackwell.

Tal, O. (2009). From heritability to probability. *Biology & Philosophy*, *24*(1), 81-105.

Tal, O., Kisdi, E., & Jablonka, E. (2010). Epigenetic contribution to covariance between relatives. *Genetics*, *184*(4), 1037-1050.

Tamez, E., Myerson, J., & Hale, S. (2008). Learning, working memory, and intelligence revisited. *Behavioural Processes, 8*, 240-245.

Thomson, G. H. (1939). *The factorial analysis of human ability*. London: University of London Press.

Thurstone, L. L. (1935). *The vectors of mind: Multiple-factor analysis for the isolation of primary traits*. Chicago: University of Chicago Press.

Thurstone, L. L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Tsou, J. Y. (2015). DSM-5 and psychiatry's second revolution: Descriptive vs. theoretical approaches to psychiatric classification. In *The DSM-5 in Perspective* (pp. 43-62). Springer Netherlands.

Tsuang, M. T., Lyons, M. J., & Faraone, S. V. (1990). Heterogeneity of schizophrenia. Conceptual models and analytic strategies. *The British Journal of Psychiatry*, *156*(1), 17-26.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science*, *9*(5), 160-164.

Turkheimer, E. (2011). Still missing. *Research in Human Development*, *8*(3-4), 227-241.

Unsworth, N., & Engle, R. W. (2005). Working memory capacity and fluid abilities: Examining the correlation between Operation Span and Raven. *Intelligence*, *33*(1), 67-81.

Van der Maas, H. J. L., Dolan, C. V., Grasman, R. P. P. P., et al. (2006). A dynamical model of general intelligence: the positive manifold of intelligence by mutualism. *Psychological review, 113*(4), 842-861.

Van der Maas, H. J. L., Kan, K. J., & Borsboom, D. (2014). Intelligence Is What the Intelligence Test Measure. Seriously. *Journal of Intelligence, 2*, 12-15.

Vernon, P. A., Wickett, J. C., Bazana, P. G., & Stelmack, R. M. (2000). The neuropsychology and psychophysiology of human intelligence. In R. J. Sternberg (Ed.), *Handbook of intelligence* (pp. 245-264). Cambridge, UK: CUP.

Vernon, P. E. (1961). *The structure of human abilities* (2nd ed.). London: Methuen.

Visscher, P. M. (2008). Sizing up human height variation. *Nature Genetics, 40*(5), 489-490.

Visscher, P. M., Brown, M. A., McCarthy, M. I., & Yang, J. (2012). Five years of GWAS discovery. *The American Journal of Human Genetics*, *90*(1), 7-24.

Visscher, P. M., Hill, W. G., & Wray, N. R. (2008). Heritability in the genomics era--concepts and misconceptions. *Nature reviews. Genetics*, *9*(4), 255-266.

Visscher, P. M., Yang, J., & Goddard, M. E. (2010). A commentary on 'Common SNPs explain a large proportion of the heritability for human height' by Yang et al. (2010). *Twin Research and Human Genetics*, *13*(6), 517-524.

Waddington, C. H. (1968). The basic ideas of biology. Reprinted in *Biological Theory*, *3*(3), 2008, 238-253.

Wahlsten, D. (1990). Insensitivity of the analysis of variance to heredity-environment interaction. *Behavioral and Brain Sciences*, *13*(01), 109-120.

Wahlsten, D. (1994). The intelligence of heritability. *Canadian Psychology/Psychologie canadienne*, *35*(3), 244.

Wahlsten, D. (2002). The Theory of Biological Intelligence: History and a Critical Appraisal. In R. Sternberg & E. Grigorenko (Eds.), *The General Factor of Intelligence. How General is it?* (pp. 245-277). Mahwah: Lawrence Erlbaum.

Wahlsten, D. (2010). Probabilistic Epigenesis and Modern Behavioral and Neural Genetics. In K. E. Hood, C. T. Halpern, G. Greenberg & R. M. Lerner (Eds.), *Handbook of Developmental Science, Behavior and Genetics* (pp. 110-122). Wiley-Blackwell.

Waters, C. K. (2007). Causes that make a difference. *The Journal of Philosophy*, *104*(11), 551-579.

Wechsler, D. (1939). *Measurement of adult intelligence.* Baltimore: The Williams and Wilkins Company, 1944.

Weedon, M. N., Lango, H., Lindgren, et al. (2008). Genome-wide association analysis identifies 20 loci that influence adult height. *Nature Genetics*, *40*, 575-583.

Weinberg, W. (1908). Uber den nachweis der vererbung beim menschen. *Jh. Ver. vaterl. Naturk. Wurttemb.*, *64*, 369-382.

Weiskopf, D. A. (2016). Integrative modeling and the role of neural constraints. *Philosophy of Science*, *83*(5), 674-685.

Weldon, W.F.R. (1902a). Mendel's laws of alternative inheritance in peas. *Biometrika 1*(2), 228-254.

Weldon, W.F.R. (1902b). On the ambiguity of Mendel's categories. *Biometrika 2*(1), 44-55.

Wickett, J. C., Vernon, P. A., & Lee, D. H. (2000). Relationships between factors of intelligence and brain volume. *Personality and Individual Differences*, *29*(6), 1095-1122.

Williams, B. A., & Pearlberg, S. L. (2006). Learning of three-term contingencies correlates with Raven scores, but not with measures of cognitive processing. *Intelligence, 34*, 177-191.

Williams, B., Myerson, J., & Hale, S. (2008). Individual differences, intelligence, and behavior analysis. *Journal of the Experimental Analysis of Behavior*, *90*(2), 219-231.

Wilson, R. A., Barker, M. J., & Brigandt, I. (2007). When Traditional Essentialism Fails: Biological Natural Kinds. *Philosophical Topics*, *35*(1/2), 189-215.

Wissler, C. (1901). *The correlation of mental and physical tests* (Vol. 2). Lancaster, PA: Press of the New Era Printing Company.

Woltereck, R. (1909). Weitere experimenelle Untersuchungen uber Artveranderung, speziell uber des Wesen quantitativer Artunterschiede bei Daphniden. *Ver Deutsche Zool Gesell*, *19*, 110-172.

Wood, A. R., Esko, T., Yang, J., et al. (2014). Defining the role of common variation in the genomic and biological architecture of adult human height. *Nature genetics*, *46*(11), 1173-1186.

Woodward, J. (2002). What is a mechanism? A counterfactual account. *Philosophy of Science*, *69*(S3), S366-S377.

Wray, N. R., Purcell, S. M., & Visscher, P. M. (2011). Synthetic associations created by rare variants do not explain most GWAS results. *PLoS biology*, *9*(1), e1000579.

Wright, L. (1997). *Twins: And What They Tell Us About Who We Are*. New York: Wiley & Sons.

Yang, J., Benyamin, B., McEvoy, B. P., et al. (2010). Common SNPs explain a large proportion of the heritability for human height. *Nature genetics*, *42*(7), 565-569.

Yule, G. U. (1902). Mendel's Laws and their Probable Relations to Intra-Racial Heredity. *New Phytologist*, *1*(10), 222-238.

Zachar, P. (2014). Beyond Natural Kinds: Toward a "Relevant" "Scientific" Taxonomy in Psychiatry. In H. Kincaid & J. A. Sullivan (Eds.), *Classifying Psychopathology. Mental kinds and natural kinds* (pp. 74-104). Cambridge, US: MIT Press.

Zhong, S., Chew, S. H., Set, E., et al. (2009). The heritability of attitude toward economic risk. *Twin Research and Human Genetics*, *12*, 103-107.

Zhu, M., & Zhao, S. (2007). Candidate gene identification approach: progress and challenges. *International Journal of Biological Sciences*, *3*(7), 420-427.

Zuk, O., Hechter, E., Sunyaev, S. R., & Lander, E. S. (2012). The mystery of missing heritability: Genetic interactions create phantom heritability. *Proceedings of the National Academy of Sciences*, *109*(4), 1193-1198.

Zwijnenburg, P. J. G., Meijers-Heijboer, H., & Boomsma, D. I. (2010). Identical but not the same: The value of discordant monozygotic twins in genetic research. *American Journal of Medical Genetics. B: Neuropsychiatric Genetics, 153B,* 1134-1149.