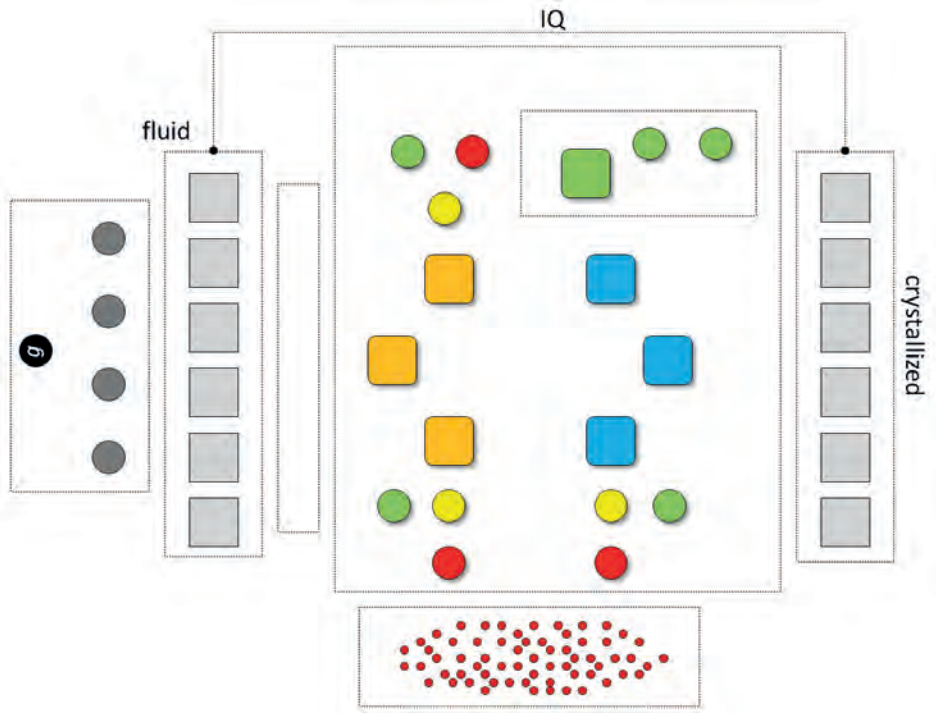
THE NATURE OF NURTURE

THE ROLE OF GENE-ENVIRONMENT INTERPLAY IN THE DEVELOPMENT OF INTELLIGENCE



Kees-Jan Kan

The Nature of Nurture: The Role of Gene-Environment Interplay in The Development of Intelligence

Kees-Jan Kan

ISBN/EAN 978-94-6191-139-1

Printed by Ipskamp Drukkers B.V., Enschede

The Nature of Nurture: The Role of Gene-Environment Interplay in The Development of Intelligence

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van Rector Magnificus
Prof. dr. D.C. van den Boom
ten overstaan van een door het college voor promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 24 januari 2012, te 12:00 uur

door

Kees-Jan Kan

geboren te Zaanstad

PROMOTIECOMMISIE:

Promotor:	Prof. dr. H.L.J. van der Maas
Copromotores:	Prof. dr. C.V. Dolan
	Dr. J.M. Wicherts
Overige leden:	Prof. dr. P.A.L. de Boeck
	Prof. dr. D. I. Boomsma
	Prof. dr. H. Kelderman
	Prof. dr. M.E.J. Raijmakers
	Dr. D. Borsboom

Faculteit der Maatschappij- en Gedragswetenschappen
Universiteit van Amsterdam
Afdeling Psychologie

To Johannes Philippus Kan and Jan Blonk sr.

CONTENTS

CHAPTER 1

INTRODUCTION

1.1 The Difficulty of Intelligence

This thesis stresses the importance of the development of better theory in intelligence research. In addition, it illustrates that taking explicit scientific philosophical standpoints (e.g. realist or nonrealist) with respect to intelligence, and the variables that relate to it, helps to understand empirical results. Without adequate theory and without researchers' philosophical standpoints, it is extremely difficult to make sense of the intelligence literature, as the author of this thesis experienced. The aim of this introduction is first to illustrate why this is so difficult.

Intelligence relates to how well systems process information, for example how well they solve complex problems, how well they can store and retrieve information, how fast they process information, but also how much information these systems contain. In addition, when we restrict ourselves to the intelligence of human information processing systems - people - it should be noted that the level of intelligence is usually not determined on a quantitative scale (interval or ratio scale), but on a relative one (ordinal scale). Without commitment to any specific theory that relates intelligence to a quantitative property (or multiple quantitative properties), intelligence is thus most appropriately expressed in terms of a rank order (Bartholomew, 2004). However, as we will show below, different rank orders of the level of information processing (henceforth cognitive functioning) can be made. This is what makes the detailed interpretation of intelligence surprisingly difficult. Consider the question whether intelligence changes during development. Based on the same data, one researcher can legitimately conclude that it grows, hence changes, while another can legitimately conclude that it is stable and does not change at all. The conclusion largely depends on the scientific perspective.

In line with a developmental perspective, we can make a rank order on the basis of a within individual comparison. Imagine a typical human newborn, call her Anne. Anne it is not yet able to speak, does not have a vocabulary, and shows no sign that she is able to solve any problem that is stated verbally. Provided Anne's cognitive development takes place normally, by the time she is 7, she has developed and improved many abilities that relate to information processing (e.g. speech), and has gained a wealth of knowledge, including a vocabulary. At this point in development, she is well able to solve certain problems that are stated verbally. Once Anne has reached adolescence, she has gained even more knowledge and is able to solve many more (and more complex) problems, both verbal and nonverbal. Given that such cognitive growth occurs in most people, the within-subjects perspective can be translated to a between group ranking: Adolescents, have a higher level of information processing than 12-year olds, for example, and 12-year olds a higher level than 7-year olds; 7-year olds a higher level than newborns. Adopting this developmental perspective, one can argue that throughout the course of development people become more intelligent, because cognitive functioning increases.

Notably, in this developmental perspective people are treated as interchangeable, so that this perspective will not provide an answer to the question of why 7-year old Anne processes information better or worse compared to other 7-year olds. This requires an inter-individual perspective. In line with this perspective, we can make a rank order based on a between individuals comparison instead of a within individual (or between age group) comparison. Theoretically, the between individuals rank order can be completely stable throughout development. Imagine three healthy newborns, Anne, John, and Neil. Anne processes information better than John, and John better than Neil. By the time they are 7, cognitive functioning has improved in all three, but Anne still processes information better than John, and John better than Neil. By the time they are adolescents, cognitive functioning has improved even further, but Anne processes information better than John, John better than Neil, etc., etc. On the basis of this inter-individuals comparison, and in contrast to the developmental perspective, one can maintain that Anne's intelligence has not changed at all. In the

light of the above, it is important to note that the cause(s) of inter-individual differences in cognitive functioning can be entirely different from the causes(s) of intra-individual differences in cognitive functioning (cognitive growth). To illustrate this, consider Figure 1.1, which shows the hypothetical developmental trajectories of a certain cognitive ability, say verbal comprehension. Within each individual, verbal comprehension grows, but the rank order among individuals stays the same. Imagine now that within each individual the growth is purely the result of learning and practice, hence of experience. Next, imagine that every individual has the exact same experience, but that the individuals differ in genetic makeup. The between subjects rank order might reflect purely these genetic differences.

In reality, the between subjects rank order on cognitive functioning is not stable throughout development. For example, it has been found that full scale IQ at age 3 correlate less than 0.5 with full scale IQ at age 12 (Sternberg, Grigorenko, & Bundy, 2001). Also, not one, but many variables, influence cognitive growth (Sameroff, Seifer, Baldwin, & Baldwin, 1993). Interventions (e.g. adoption) and other events (e.g. illness) that occur throughout development can have a profound effect on cognitive growth, and thus on the ultimate level of cognitive functioning (Sternberg, Grigorenko, & Bundy, 2001). In order to fully understand why one individual processes information better than another, we need to take into account their life histories. This requires a developmental perspective. On the other hand, cognitive growth is not unlimited, which implies that inter-individual differences in limited resources or capacities will give rise to individual differences in the developmental trajectories and the ultimate level of information processing. In order to understand why the one individual develops differently than the other, one needs to take into account inter-individual differences in these capacities.

So, in order to understand and model human intelligence, we need both the developmental perspective and the inter-individual differences perspective. The majority of the discussions, theories, and models in the field of intelligence (as measured by psychometric tests) lack the developmental perspective. Discussions and theories are mainly concerned with individual differences in hypothesized limiting capacities (e.g. Spearman, 1904; Carroll, 1993; Jensen, 1998), either indirectly, via the history of factor-analysis, or directly. The discussions have led to many different factor models of intelligence (see, e.g. Jensen, 1998). In the next section, we give a brief overview of the most well-known factor models and their history.

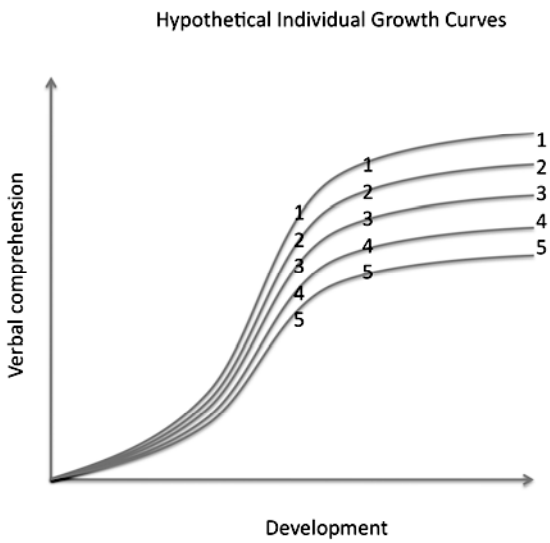


Figure 1.1 Hypothetical developmental trajectories of verbal comprehension. Within each of 5 individuals, verbal comprehension grows, but the rank order among the individuals stays the same. Individual growth can be purely the result of experience, while the between subjects rank order can reflect purely genetic differences.

1.2 Factor Modeling of Individual differences in Intelligence

Factor modeling of individual differences in intelligence started in the early twentieth century, when Spearman (1904) aimed to explain the finding that 22 pupils' ratings on pitch discrimination and 5 different school subjects (English, French, classics, mathematics, and music) were positively intercorrelated. He hypothesized that the individual differences in these ratings reflected in large part individual differences in exactly one (unobserved) variable that the scores had in common. Spearman called this common variable the general factor of intelligence (general intelligence or *g*, for short). Furthermore, he considered the correlations among the ratings to be imperfect due to the influence of other (unobserved) variables that were unique to each rating. These specific variables (*s*, for short) were assumed to be uncorrelated across the observed variables (i.e., the ratings). They were also assumed to be uncorrelated with the common variable. Spearman was aware that correlations among variables are attenuated by measurement error (*e*).

We can depict Spearman's ideas graphically, as in the path diagram in Figure 1.2. Here, the observed ratings (or test scores in general, *x*) are symbolized by squares; the unobserved, hypothetical variables by circles. As mentioned, without any commitment to a theory, intelligence, as measured by intelligence tests, is most appropriately expressed in terms of a rank order. The path diagram illustrates that a higher position on the unobserved variable *g* results (somehow) in a higher position on all observed *x* variables, while a higher position on a specific variable *s* results (somehow) in a higher position on only 1 observed *x* variable. Obviously, the diagram itself does not explain how these individual differences arise, nor if or how intra-individual differences in cognition (cognitive growth) arise, nor whether the relations among the variables are linear or nonlinear. This requires theory, for example Spearman's theory.

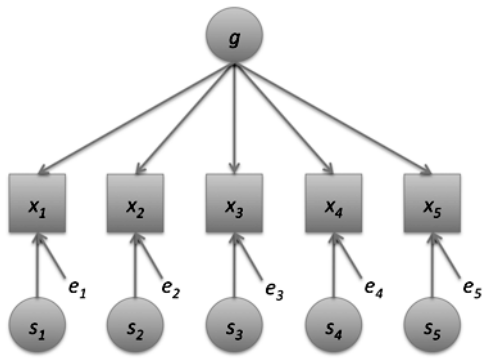


Figure 1.2 Spearman's (1904) account for the finding that cognitive test scores (*x*) are positively intercorrelated. Individual differences in these scores reflect in individual differences in exactly 1 (unobserved) common variable (the quantitative variable *g*) and other (unobserved) unique variables that were unique to each test *i* (of which the total is s_i). Correlations among variables are attenuated by measurement error (*e*).

Spearman's theory appears specific enough about the nature of the common variable *g* to draw certain conclusions in the light of a developmental perspective: He maintained that *g* is a single quantitative variable that is innate and immalleable, hence fixed. It implies that *g* is solely a source of inter-individual differences and not any cause of cognitive development within a person. Cognitive growth must be due to other influences. One way to conceive *g* is as a limiting capacity that constrains this growth. Any individual differences in the other influences cause the relation between *g* and the test scores to be imperfect.

To test whether positive correlations among people's intelligence test scores are indeed due to a single quantitative common source of individual differences, Spearman pioneered the statistical technique factor analysis. His statistical one-common factor model represented his theory formulated in linear equations. The mathematical formulation of the theory is thus stricter than the theory itself, because the relations among the variables are taken to be linear. Because the means of the unobserved quantitative variable *g* and the specific factors are unknown, Spearman's model is

usually written in the form of a regression model:

$$x_{ij} = \beta_0 + \lambda_j g_i + \varepsilon_{ij}$$

The symbol x denotes person i's test score on test j, g represents the (fixed) inter-individual quantitative variable general intelligence, and ε_j the total effect of other influences (s) on test j, including the measurement error (e). Parameters λ_j denote weights (regression coefficients). Usually, it is assumed that variable g and ε_j have means of 0 (means have been virtually subtracted). In the interpretation of this regression model as the true data generating mechanism, β_0 does thus not represent simply a scaling parameter, but includes the subtracted means of g and ε_j. In principle, it also reflects any influences that show no inter-individual differences (of which the variance is 0).

With respect to causation, the mathematical equation is conceptually weaker than the theory and the path diagram. Mathematically, provided λ_j is not zero, $x_{ij} = \beta_0 + \lambda_j g_i + \varepsilon_{ij}$ is the same as $g_i = [-\beta_0 + x_{ij} - \varepsilon_{ij}]/\lambda_j$, whereas the hypothesis 'differences in g give rise to differences in x' is different from the hypothesis 'differences in x give rise to differences in g'. In the statistical testing procedure, it is investigated, albeit indirectly, whether a higher position on the hypothesized variable g corresponds to (rather than results in) a higher position on all observed x variables, while a higher position on unobserved variable ε_j corresponds to a higher position on only one observed variable (x_j).

Spearman's theory has been criticized for a variety of reasons (see Jensen, 1998). First of all, researchers have questioned the robustness of the positive correlations (e.g. Guilford, 1964). However, positive correlations among cognitive tests have been replicated in hundreds of datasets (Carroll, 1993). The finding, which is often called the positive manifold of intelligence, is thus robust and is now regarded as an empirical fact. Second, researchers came up with valid alternative explanations of the positive manifold (Thomson, 1951; Bartholomew, Deary and Lawn, 2009; van der Maas et al., 2006; Dickens, 2008). The consensus is now that Spearman's theory is in line with the positive manifold, but that factor analysis cannot prove a theory, including Spearman's, to be correct. Factor analytical results are statistical summaries of the data; this because although factor analysis does involve goodness-of-fit testing, different mechanisms can lead to the same factorial structure. Therefore theory is required in order to attach a meaning to the factors, for example as representing a true, causal underlying variable (e.g. Borsboom, Mellenbergh, & van Heerden, 2003).

Although factor analysis cannot prove a theory is correct, it can falsify a theory, namely when the theory is formulated as a statistical model, and this model does not fit the data adequately. Eventually this happened to Spearman's theory. It was rejected because one-factor models are usually too simple to explain the data: After partialling out the variable g, certain test scores remain positively intercorrelated, which means that certain specific factors in Spearman's model are not statistically independent across tests, as was hypothesized.

Because one-factor models do not give adequate explanations for the patterns of the positive intercorrelations among IQ test scores, scientists developed factor models that include multiple factors. The pioneering researchers often published in all the fields of statistics, test theory, and intelligence, so that these fields were closely connected. Later these fields started to diverge, with the consequence that the connection between theories of intelligence and statistical models of intelligence is nowadays less clear. The consequence is that it is also unclear whether factors should be taken as mere statistical variables or as representing true causal sources of variance (e.g. limiting capacities).

One of the pioneers of (psychometric) intelligence, Thurstone (1938), advanced a model that describes intelligence as existing of multiple (7) statistically independent cognitive abilities. These comprised Word Fluency, Verbal Comprehension, Spatial Visualization, Number Facility, Associative Memory, Reasoning, and Perceptual Speed. Each cognitive ability can be measured by a number of tests (see Figure 1.3); test specific factors attenuate the relation between the cognitive ability and the test scores. Thurstone's model explains the (strong) intercorrelations among tests of (say) Verbal Comprehension on the one hand, and those among tests of Associative Memory, on the other. However, it does not explain (the weaker) positive intercorrelations among tests of Verbal Comprehension and tests of Associative Memory. Multiple solutions have been proposed (see Jensen, 1998). One can propose that IQ tests are not uni-dimensional, for example, by assuming that tests

always tap multiple of the statistically independent cognitive abilities, but to different extents (consistent with the model displayed in Figure 1.4), or by assuming that each test measures a general ability in addition to a specific cognitive ability (see Figure 1.5). Assuming that test constructors aim (and succeed) to devise uni-dimensional tests, a theoretically more satisfactory solution is to allow for positive intercorrelations among the cognitive abilities (see Figure 1.6). Of course, it can be hypothesized these intercorrelations are due to a common variable (see Figure 1.7).

More precisely than before, intelligence can now be defined as a weighted average of the cognitive abilities measured by IQ tests. In addition to the problem of how to determine the weights, this definition suffers from the same problems as those mentioned above. Without any further theorizing about the cognitive abilities, intelligence still refers to a rank order, and its interpretation is difficult because the cognitive abilities are not observed directly. Again, on the basis of the same factor analytical results, the one researcher can maintain that there is growth in intelligence, while the other can maintain there is not. We can theorize, for example, that there is growth in the cognitive abilities, which would explain why people's test scores increase during development. Alternatively, we can theorize that the observed knowledge and skills to solve the tests items increase, but that the cognitive abilities represent fixed capacities that constrain the growth of the knowledge and skills, so that performance, rather than underlying intelligence, changes. Factor analysis of individual differences cannot discriminate between those alternatives (to do so latent growth modeling can be used). In any case, in order to interpret psychometric intelligence and individual differences in full scale IQ, one needs theory.

One such theory is Cattell's (1963, 1987) investment theory of fluid and crystallized intelligence. This is one of the few theories of intelligence that aims to account for both the development of cognitive abilities and the factorial structure of intelligence. In the theory, one factor of intelligence, fluid intelligence (Gf), represents an underlying reasoning capacity, which is connected to the maturation of the brain. Individual differences in this capacity are largely due to individual differences in genetic makeup. The acquisition of knowledge and skills, called crystallized abilities, depends on this capacity. Hence individual differences in crystallized abilities reflect in part individual differences in fluid intelligence. Another factor of intelligence, crystallized intelligence (Gc), summarizes the common variance among these crystallized abilities. Because Gc depends on Gf, factors Gf and Gc are modeled as positively correlated.

Like Spearman's *g* model, Cattell's Gf-Gc model was shown to be too simple (e.g. Carroll, 1993). Nowadays, it is assumed that human intelligence comprises about 70 positively inter-correlated specific cognitive abilities (represented by first order factors) (McGrew & Flanagan, 1998). If one would factor analyze in turn the intercorrelations among these abilities, factor solutions including about 8 to 10 correlated factors would give statistically satisfactory results (Cattell, 1987; Carroll; McGrew, 2009; Horn & Blankson, 2005). At this (second-order) level, the two most widely accepted models of intelligence, the extended Gf-Gc factor model of Cattell and Horn and the three-stratum model of Carroll, are nearly identical (McGrew, 2009). The interpretation of the second order factors in these models is often casted in terms of separate cognitive or biological systems (e.g. Carroll). Also, individual differences in the factors of intelligence are considered to be due to both genetic and environmental influences (Plomin et al. 2008).

Because the second-order factors of intelligence are all positively intercorrelated, it is possible to factor analyze their intercorrelations. Carroll (1993) did so, and posited a (third-order) general factor. The model can be regarded as a model in which Spearman's *g* is reintroduced, while allowing for correlations among certain specific factors. Carroll hypothesized the general factor to represent individual differences in a cognitive system separate from the systems represented by the second order factors. Horn and Cattell did not extract a general factor. Horn (Horn & Noll, 1997) regarded such factor as nothing more than a statistical summary of the second order factors; Cattell (1987) believed that the general factor and fluid intelligence represented the same variable (hence cognitive system). Over the last decades, researchers have been trying to synthesize the Gf-Gc and the three-stratum model into one (Cattell-Horn-Carroll) model of intelligence (see McGrew, 2009). The question whether or not to extract a general factor remains a topic of discussion. This discussion is based on theoretical arguments rather than statistical ones. In this light, alternative explanations of the positive manifold are important.

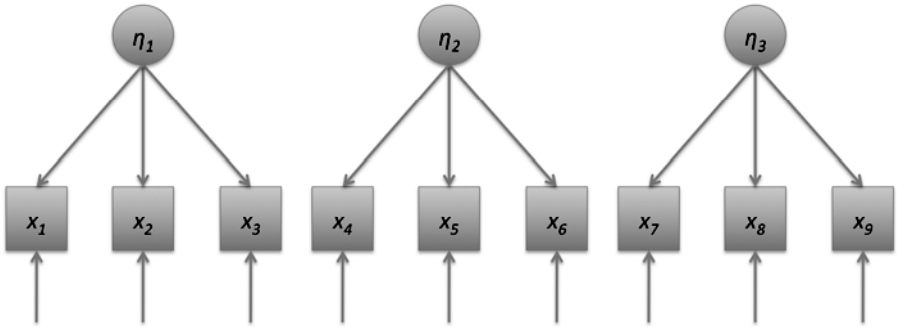


Figure 1.3 Multiple factor model of psychometric intelligence. Each human cognitive ability (η) can be measured by a number of tests (x). Test specific factors attenuate the relation between the cognitive ability and the test scores. The model explains the (strong) intercorrelations among tests that measure the same ability, but does not explain the (weaker) correlation among tests that measure different cognitive abilities.

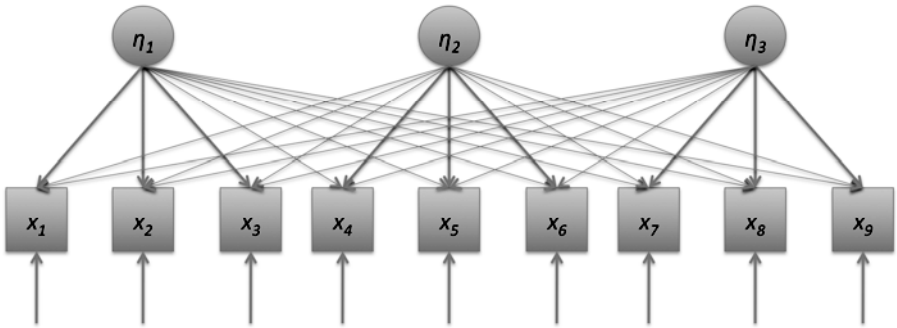


Figure 1.4 Multiple factor model of psychometric intelligence. Intelligence tests always tap from multiple statistically independent cognitive abilities, but to different extents. The model explains strong and weaker intercorrelations among intelligence tests. Test specific factors attenuate the correlations.

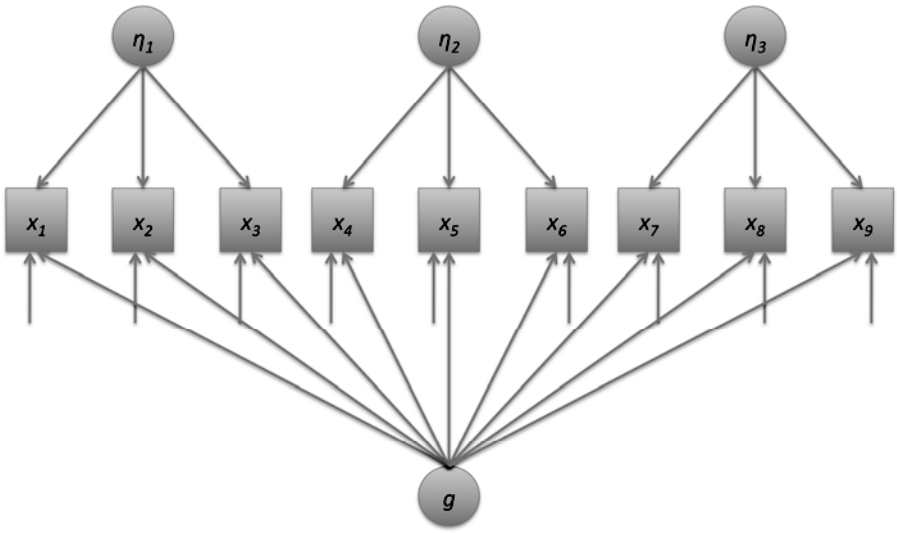


Figure 1.5 Multiple factor model of psychometric intelligence. Each intelligence test does not only measure a specific cognitive ability but also a general ability. The model explains strong and weaker intercorrelations among intelligence tests. Test specific factors attenuate the correlations.

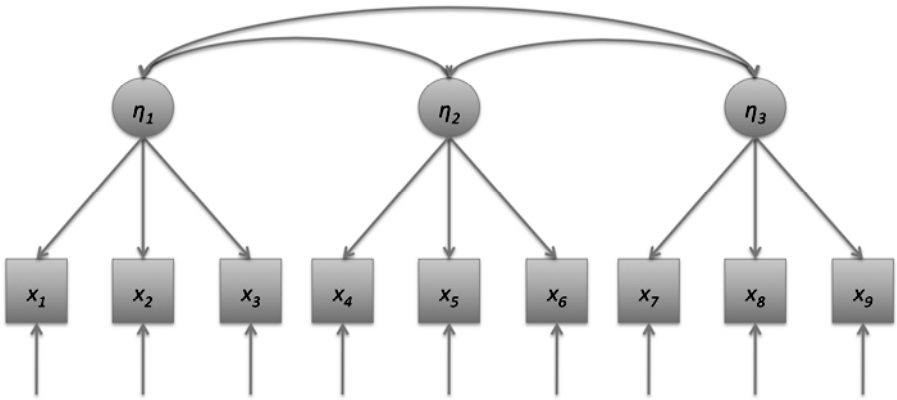


Figure 1.6 Multiple factor model of psychometric intelligence. Each human cognitive ability can be measured by a number of tests. Test specific factors attenuate the relation between the cognitive ability and the test scores. The model explains the (strong) intercorrelations among tests that measure the same ability as well as the (weaker) correlation among tests that measure different cognitive abilities. The causes of the latter are unknown. Tests are uni-dimensional.

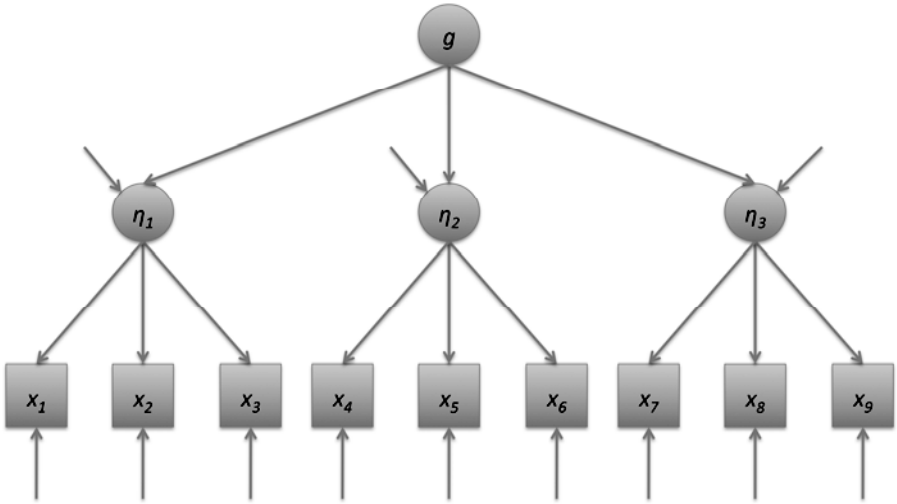


Figure 1.7 Hierarchical factor model of psychometric intelligence. Each cognitive ability can be measured by a number of tests. Test specific factors attenuate the relation between the cognitive ability and the test scores. Each cognitive ability is influenced by variable *g*. Influences specific to cognitive abilities attenuate the relation between *g* and the cognitive abilities. The model explains the (strong) intercorrelations among tests that measure the same ability as well as the (weaker) correlations among tests that measure different cognitive abilities. Tests are uni-dimensional.

1.3 A Dynamical Systems Model of Intelligence

When variables are positively intercorrelated, a statistical general factor can be extracted, but it does not necessarily imply that a true, causal general variable is present. The positive manifold of intelligence, hence the presence of a general factor of intelligence, can be the result of the development itself for example, as in the mutualism model of general intelligence of van der Maas et al. (2006). In this thesis the ideas behind the mutualism model play an important role in the

conceptualization of intelligence. Here we give a brief summary of these ideas. More detailed descriptions of the mutualism model and the underlying assumptions are provided and discussed elsewhere in this thesis (in Chapter 6 in particular).

In the mutualism model it is assumed that growth in cognitive processing is largely autonomous and self-regulating, but that it is constrained by heritable limited resources or capacities. However, different cognitive processes exist, each of which is constrained by unique limited resources. Another assumption is that the cognitive processes influence each other during their development and that the effects are largely mutually beneficial. That is, the growth of one cognitive process stimulates the growth of the other cognitive process and vice versa. As a result, the limiting capacity that constrains the growth of the one process, say working memory, has an indirect effect on the growth of other cognitive processes. On the other hand, the limiting capacities of these other processes have an indirect effect on working memory.

Mathematically the mutualism model is formulated as follows:

$$\frac{dx_i}{dt} = a_i x_i (1 - \frac{x_i}{K_i}) + a_i \sum_{\substack{j=1 \\ j \neq i}}^{W} \frac{M_{ij} x_j x_i}{K_i}, \text{ for } i, j = 1 \ldots W.$$

Where the K's represent the (W) limited resources that constrain the growth of the (W) cognitive processes. Parameters a_i are growth parameters associated with each cognitive ability x. Weight M_{ij} determines the influence of cognitive process x_j on cognitive process x_i. Genetic and environmental influences can be introduced, for example via the limited resources. Simulations showed that even if all underlying variables (a's, and K's) are initially uncorrelated, the observed variables (x's) become intercorrelated over time. That is, individual differences in the cognitive processes eventually show a positive manifold.

1.4 Overview

The aims of the thesis are twofold. The first aim is to reinvigorate the development of an adequate theory of intelligence by providing a model that accounts for both cognitive growth and (heritable) interindividual differences in intelligence. The integrated model must be able to explain salient findings in intelligence research, such as a correlation between intelligence subtests' heritability coefficients and their loadings on the (statistical) general factor of intelligence. While developing the theory, we encountered theoretical issues that are not fully addressed in the literature. The second aim is to address these issues. They concern mostly the interpretation of the heritability of intelligence.

In this thesis, we provisionally accept the Cattell-Horn-Carroll (CHC) model as a working hypothesis. We hypothesize that the second order factors in this model (including Gf and Gc) represent individual differences in unique cognitive systems, constrained by genetically and environmentally influenced capacities. We do not posit a substantive underlying general factor, because we believe the positive intercorrelations among the cognitive systems are due to reciprocal interactions among those systems (see Figure 1.8), which occur throughout cognitive development, as in the mutualism model. We also assume that in principle individual differences in these systems can be measured by intelligence tests, although we maintain that in practice intelligence tests are likely not uni-dimensional.

Chapter 2 (published as Kan, Ploeger, Raijmakers, Dolan & van der Maas, 2011) concerns difficulties with the interpretation of the latent genetic and environmental variables in (behavior genetic) statistical models in general. Obviously, this has consequences for the interpretation of the latent genetic and environmental variables in factor models of intelligence. The genetic and environmental variables in these behavior genetic models are not measured, but inferred, using genetically informative research designs (e.g. twin studies). Furthermore, the underlying variables are modeled as acting linearly. We investigate what the estimated relative contributions of the linearly modeled genetic and environmental variables to the total variance in the observed variables are, when in reality the underlying mechanisms are non-linear. We hypothesized that the estimates are then not correct. We aimed to corroborate this with a reciprocal dynamical systems model. We could not use the mutualism model, because this model does not contain nonlinear terms.

Instead we used the two-cell model of van Oss & van Ooyen (1997) (this model is described in Appendix A). The two cells in the model are interpreted as subsystems of working memory capacity (e.g. Gsm in the CHC model). We provide results that are in line with empirical findings in intelligence research: First, the underlying causes are hard to detect, if at all. Second, the estimated relative contribution of the hypothesized genetic variables to the total variance increases (hence the estimated relative contribution of the hypothesized environmental variables decreases). We conclude that caution is required in interpreting high heritability coefficients as meaning 'highly genetically influenced'. Heritability coefficients can be overestimated due to nonlinear developmental effects (e.g. stage transitions).

Further caution is required when interpreting high heritability. As we explain in Chapter 3, a relatively high estimated heritability of a cognitive ability relative to another does not mean environmental and cultural influences are relatively less important. On the contrary, the more cultural aspects differentiate between people, the higher the subtests' heritability coefficients. We show the highly contra-intuitive finding that in IQ tests the highest heritability coefficients are generally of tests that measure highly culturally dependent knowledge and skills, rather than tests that measure less cultural dependent cognitive processing. In terms of investment theory, this means that crystallized abilities are more heritable than tests that measure fluid abilities. The original theory predicted the opposite, namely that fluid abilities are more heritable than crystallized abilities. We suggest that the explanation may lie in gene-environment correlation. Also, crystallized ability tests happen to display the highest loadings on the general factor of intelligence. 'General intelligence' (as a statistical construct) appears to be more like 'crystallized intelligence' than 'fluid intelligence'.

Chapter 4 concerns the relation between fluid and crystallized abilities and group differences. We investigate the claim that when group (e.g. racial or ethnic) differences are the most pronounced on the most heritable and the most *g* loaded subtests it implies the origin must be genetic. This claim is based on Jensen's (1998) method of correlated vectors. Even if researchers accept this (methodologically weak) method, the claim cannot be made, because it involves invalid reasoning, namely affirming the consequent. We show analytically that group differences can be the most pronounced on the most heritable and the most *g* loaded tests when the origin is entirely environmental. In addition, we show empirically that group differences are the most pronounced on the most culturally loaded subtests (crystallized abilities). These results conflict with mainstream theories, because they predict the opposite: group differences are most pronounced on the least cultural loaded tests (fluid abilities).

Chapter 5 (of which parts are published as Kan, Kievit, Dolan, & van der Maas, 2011) concerns the further interpretation of fluid and crystallized intelligence. From a review of investment theory, we conclude that crystallized intelligence must be interpreted as a constructivist variable, i.e., a statistical summary (of the amount of knowledge). This interpretation differs from the interpretation of the CHC factors as representing individual differences in underlying cognitive capacities. We propose to remove factor Gc from the CHC model. We showed that this can be done legitimately, because in a reanalysis of a representative dataset on which the CHC model is based, the factor Gc was redundant, as was predicted; it was statistically equivalent to verbal comprehension. In the dataset, *g* was also redundant, because it was statistically equivalent to fluid intelligence (Gf). We propose that *g* can also be removed from the CHC model as an explanatory variable. The magnitude of the correlation between Gf and *g* is likely a function of sample heterogeneity due to developmental differences.

In Chapter 6 we address whether current theories of general intelligence (*g* theory, fluid-crystallized theory, sampling theory, and reciprocal interaction theories) explain the intriguing finding that the most cultural dependent cognitive abilities (crystallized abilities) are the most *g* loaded and most heritable. It is concluded that (in isolation) they cannot. By implication, the reviewed theories do not explain how group differences become the most pronounced on the most culturally dependent, most heritable, most *g* loaded subtests.

The thesis ends with a general discussion (Chapter 7), in which we present an integrated model of general intelligence. It is a mutualism model (van der Maas et al., 2006) that incorporates the main idea of investment theory (individual differences in cognitive processes - fluid abilities - give rise to differences in knowledge and skills - crystallized abilities) and Dickens & Flynn's (2001;

Dickens, 2008) social multiplier. In line with the mutualism theory (van der Maas et al., 2006), we assume that cognitive processing benefits from knowledge acquisition. In the integrated theory, an underlying *g* (Spearman, 1904; Carroll, 1993; Jensen, 1998) is absent. Genetic correlations among limiting capacities can be present, but are taken to be the result of what we denote genetic sampling (Thompson, 1951; Bartholomew et al. 2009; Anderson, 2001; Penke et al., 2007; see Chapter 6) and not as due to general genetic effects (Kovas & Plomin, 2006). The integrated theory accounts for the fact that individual differences (hence group differences) are the most pronounced on the most culturally dependent subtests, which are the most heritable and the most *g* loaded. The effect is due to differences in gene-environment effects across cognitive abilities.

The main points of this thesis are as follows. First, although it is still not possible to determine whether a realistic, underlying *g* is present or not, we can conclude that current *g* theories are inadequate in explaining certain salient empirical findings. Next to the individual differences perspective they have, they need a developmental perspective. The role of the dynamic interplay between genetic and environmental variables that occurs during development needs to be explicated. Second, formal modeling is important in intelligence research. Using the mutualism model can help researchers to study combined effects, such as the investment hypothesis of fluid and crystallized intelligence and the correlation among subtests' *g* loadings, cultural loadings and heritability coefficients. The concluding chapter provides an example of this kind of formal modeling. The main advantage of the mutualism model is that it can combine the developmental and individual differences scientific perspectives on intelligence. Formal modeling is challenging and may seem difficult, but, as the author of this thesis has experienced, it is considerably less difficult than trying to make sense of the intelligence literature while adequate theory is lacking.

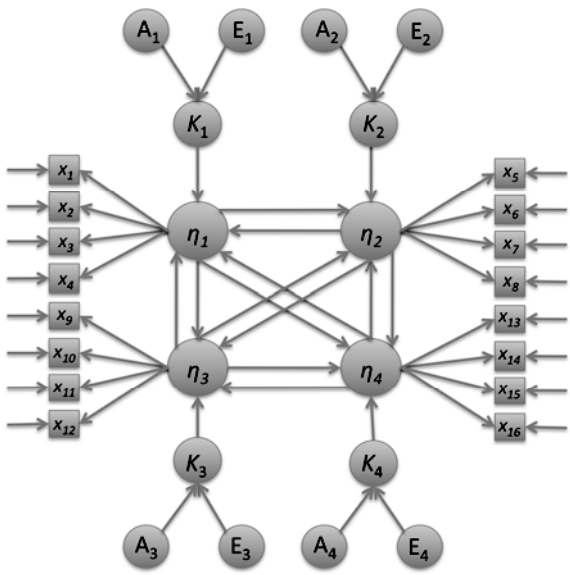


Figure 1.8 Interpretation of the broad factors in the Cattell-Horn-Carroll model of intelligence. These factors represent information processing systems. The development of the functioning of each system is constraint by genetically (A) and environmentally (E) influenced limiting capacities (*K*), which are assumed to be uncorrelated. Individual differences in the functioning of the systems (cognitive abilities, η) can be measured by a number of tests (x). Test specific factors attenuate the relation between the cognitive abilities and the test scores. Correlations among test scores that measure different cognitive abilities are due to mutual beneficial interactions among the systems, which are assumed to take place throughout development. Tests are modeled as uni-dimensional.

CHAPTER 2

NONLINEAR EPIGENETIC VARIANCE: LITERATURE REVIEW AND ILLUSTRATIVE SIMULATION STUDY

Abstract

We present a review of empirical evidence that suggests that a substantial portion of phenotypic variance is due to non-linear (epigenetic) processes during ontogenesis. The role of such processes as a source of phenotypic variance in human behavior genetic studies is not fully appreciated. In addition to our review, we present simulation studies of nonlinear epigenetic variance using a computational model of neuronal network development. In each simulation study, time series for monozygotic (MZ) and dizygotic (DZ) twins were generated and analyzed using conventional behavior genetic modeling. In the results of these analyses, the nonlinear epigenetic variance was subsumed under the nonshared environmental component. As is commonly found in behavior genetic studies, observed heritabilities and unique environmentabilities increased with time, whereas common environmentabilities decreased. The fact that the phenotypic effects of nonlinear epigenetic processes appear as unsystematic variance in conventional twin analyses complicates the identification and quantification of the ultimate genetic and environmental causes of individual differences. We believe nonlinear dynamical system theories provide a challenging perspective on the development of individual differences, which may enrich behavior genetic studies.

2.1 Nonlinear Epigenetic Variance: Literature Review and Illustrative Simulation Study

The nature-nurture question, as addressed in (behavior) genetic studies, concerns the decomposition in a linear statistical model of phenotypic (observed) variance into components attributable to latent factors. In humans, this decomposition is carried out in genetically informative designs, such as the twin design (e.g. Plomin, DeFries, McClearn, and McGuffin, 2008). Generally two general latent sources of variance are distinguished, namely the genetic and the environmental. Phenotypic variance attributable to the former may be decomposed into additive, dominance, and epistatic genetic variance, whereas phenotypic variance attributable to the latter can be further decomposed into shared environmental variance, and nonshared environmental variance, where shared influences are attributable to the environment cohabiting individuals (both human and infra-human) share.

Regardless of the (genetically informative) design employed, it is assumed that genotypic and environmental factors reflect the underlying mechanisms causing phenotypic individual differences (e.g. Plomin et al., 2008).[1] An important characteristic of the behavior genetic methodology is that the partitioning of variance and the attendant causal interpretation pertains to phenotypic individual differences, and not to phenotypes themselves (Dolan & Molenaar, 1995; Lewontin, 1974; Oyama, 1985).[2] However, if we remain strictly within the domain of individual differences, a causal interpretation still requires caution. First, the validity of the interpretation may be undermined by the limitations of the statistical model employed to carry out the decomposition of phenotypic variance. These may concern the failure to take into account genotype-environment interaction, genotypic-environment covariance, assortative mating, etc. However, it should be noted that in principle, i.e., given a suitable design or given sufficient information, these effects can be modeled (e.g. Plomin et al.).

Indeed, there is an increasing effort to devise statistical models that can accommodate these complicating effects (Dick & Rose, 2002; Purcell, 2002; Eaves & Erkanli, 2003). At present, the issue

[1] It is in this sense that we take the term "explained" seriously, i.e., genetic and environmental differences cause phenotypic differences.

[2] What causes (say) cognitive functioning, and what causes individual differences in cognitive functioning are different questions (Oyama, 1985).

of genotype-environment interaction is the focus of much attention (Turkheimer, D'Onofrio, Maes, & Eaves, 2005; Moffitt, Caspi & Rutter, 2005). Increasingly, specific - measured - genetic and environmental variables are incorporated in the model, thus replacing in part the latent factors (Plomin, et. al, 2008), and creating important opportunities for studying genotype-environment interaction. Error in the measurement of the phenotype remains hard to pin down, as it is due, by definition, to transitory effects (e.g. Molenaar, Boomsma, & Dolan, 1993; Turkheimer & Waldron, 2000).

The achievements within behavior genetics have transformed present day psychology and the standard behavior genetic model, including the interpretation of the roles of genetic and environmental variables as causes of individual differences, has been successful enough to convince most psychologists that genetic and environmental factors do indeed contribute to phenotypic variance, and to encourage further research into the role of specific environmental agents and actual genes (i.e., possible ranging in effect from major loci to quantitative trait loci). This research is based on the premise that the phenotypic variance is ultimately traceable to identifiable variables. However, as we will argue, the search for such specific genetic and environmental variables may be complicated in that a substantial portion of phenotypic variance may be due to non-linear (epigenetic) processes (Molenaar et al., 1993). As explained below, these effects will appear unsystematic, and are therefore difficult to distinguish from specific environmental effects or measurement error.

The aim of this chapter is to study the role of non-linear epigenetic processes as a source of phenotypic variance. To this end, we review the relevant literature and demonstrate that there is ample evidence in support of these processes. In addition, we present the results of computer simulations. The implications and consequences of the presence of nonlinear epigenetic variance are particularly interesting in the light of the attempts to identify specific nonshared, environmental influences (e.g. see Turkheimer & Waldron, 2000; Plomin et al., 2008).

The chapter is organized as follows. We first consider non-linear epigenetic processes in more detail. Second, we present a review of relevant studies involving inbred and isogenic animals, animal and human studies of developmental instability in several biological traits, and behavior genetic studies of human behavior. Third, we present a computational model of neuronal developmental processes, i.e., the two-cell model of neurite outgrowth of van Oss and van Ooyen (1997). This non-linear model displays the characteristics that are hypothesized to underlie nonlinear epigenetic variance. We present the results of computer simulation studies based on this model. In these studies we simulated phenotypic twin data, which allowed us to study the effects of nonlinear epigenetic variance in the standard twin design (e.g. Plomin et al. 2008). We end this chapter with a discussion concerning the role of such processes in development, the implications for behavior genetic studies, and possible limitations of our simulations.

2.2 Nonlinear Epigenetic Processes During Development

In the present section we discuss the hypothesis, that nonlinear epigenetic processes may constitute a source of phenotypic variance. In the subsequent section, we review empirical support for this hypothesis, stemming from both animal and human studies, and we discuss nonlinear epigenetic processes as a source of nonshared environmental variance in human psychological traits.

All developmental processes can be conceived of as the outcome of some dynamical system (Guckenheimer & Holmes, 2002; van der Maas & Molenaar, 1992). Viewing the development of an organism as the outcome of a nonlinear dynamical system, we accept the following characteristics. In contrast to linear systems, nonlinear systems are characterized by a disproportional relationship between cause and effect (Arnold, Afrajmovich, Il'yashenko, & Shil'nikov, 1994). This implies that large influences may have small or limited effects, whereas small causes, e.g. initial changes or differences, may have large effects. Hence, near-indistinguishable sets of initial conditions in the same system may produce different outcomes. In addition, if the system in question is sensitive to initial conditions, the outcome of the system may be hard to predict even if it is deterministic, i.e., any form of randomness is absent. This unpredictability arises from the lack of precise knowledge concerning the initial conditions. Hence, the outcome will appear stochastic despite the fact that the

system is fully deterministic. We refer to such unpredictable (either real stochastic or seemingly stochastic) relationships between cause and effect as *unsystematic*.

Sensitivity to initial conditions can refer to two related, but distinct, forms of unpredictability (Arnold et al., 1994). In some nonlinear systems the result is *seemingly random behavior*. Such behavior is referred to as chaotic. In other nonlinear systems the result is *a sudden qualitative change in behavior* when a small smooth change is made to a parameter. In mathematics, such a change is termed a bifurcation. Bifurcations are a general characteristic of nonlinear systems in physics (Prigogine, 1980), biology (Meinhardt, 1980), and psychology (e.g. van der Maas & Molenaar, 1992; Kelso, 1995; Ploeger, van der Maas, & Raijmakers, 2008). In physics they are called phase transitions; in biology and psychology they are often called stage transitions. Henceforth, we use the term bifurcation as a general term to refer to these transitions.

Non-linear systems can attain certain levels of order, structure, and stability by a process known as self-organization, i.e., an autonomous and self-regulating process (e.g. Camazine, 2001; Meinhardt, 1982). Numerous instances of self-organization have been found in both non-living and living systems. These include the formation of stripes in sand dunes and patterns on skins, coats, and shells (Camazine; Meinhardt). Self-organization has also been established in the process of morphogenesis that underlies the structured branching in organs, such as lungs, the cardiac muscle network, and the blood circulatory system. In addition, the brain may be viewed as a highly structured neural network. The notion that the formation of these structures involves self-organization is supported by the fact that the total amount of information stored in the genome is too small to prescribe these structures in any detail (Benno, 1990; Molenaar et al, 1993; Stent, 1978). In other words, self-organizing processes are required to explain the process of ontogenesis.

In this context, epigenetics and epigenesis are important concepts. In general terms, there is a close correspondence between self-organization and epigenesis, as Molenaar & Raijmakers (2000) emphasized: ”[E]pigenesis constitutes an instance of a self-organizing developmental process" (p. 45). Or in the words of Beloussov (2006): “Epigenesis may be regarded as the theory of self-organization as applied to ontogenetic phenomena” (p.1165). Traditionally, epigenetics refers to the study of such processes, i.e., “the way genes and their products bring the phenotype into being” (Jablonka & Lamb, 2002, p.82; Waddington, 1957). Nowadays, with the greater understanding of the molecular mechanisms that control gene activity during embryonic development and cell differentiation, epigenetics is defined as “the study of mitotically and/or meiotically heritable changes in gene function that cannot be explained by changes in DNA sequence” (Jablonka & Lamb, p.87).

Despite differences in meaning and explanation, both traditional and molecular biological epigenetics focus on alternative developmental pathways, and on the influence of environmental conditions and its consequences for the organism. After fertilization, non-linear mechanisms[3] initiate autonomous growth processes that give rise to structure and pattern (e.g. brain structure). However, the environment, the epigenetic process itself, and their interactions with both each other and with the genetic effects tend to perturb development, possibly resulting in variations in developmental pathway. That is, at critical points (i.e., bifurcation points, Arnold et al., 1995; Guckenheimer & Holmes, 2002), these perturbations may cause development to follow a different trajectory, resulting in variations in structure or pattern (Waddington, 1957). Given reciprocal influences between the development of neuronal form and function, on the one hand, and neural activity, on the other (e.g. van Oss & van Ooyen, 1997), differences in neuronal structure can result in differences in activity and, thus, can become manifest at the behavioral level (See also Benno, 1990). In the next sections we consider the plausibility of sensitivity to initial conditions and the related non-linear (epigenetic) processes as a source of unsystematic phenotypic variance.

[3] Rather than needing to code each detail of an organism explicitly, the genes only require a minimal amount of genetic encoding specifying the rules that constitute these mechanisms.

2.3 Empirical Evidence Of Sources of Unsystematic Phenotypic Variance

Empirical evidence, from both animal and human studies, suggests that phenotypic variance cannot be explained completely by genetic and (external) environmental factors. For example, Gärtner (1990) reviewed the decades-long efforts in his own laboratory to minimize the variance of biological traits in laboratory animals by standardization of environmental and genetic conditions. The variance in a number of traits of highly inbred rats held under strict environmental control was compared with the variance in the same traits of rats living in a natural wild setting. These traits included morphological, biochemical, and other quantitative traits, such as blood parameters and kidney weight. The variance in inbred rats was not appreciably lower than the variance in wild rats. Hence, according to Gärtner, neither the postnatal environment nor genes appeared to constitute a major source of phenotypic variance in the inbred strains.

Subsequent research was performed to assess environmental variance directly. To this end, eight-cell stage mice embryos were divided, thereby creating monozygotic twin pairs. Each twin pair was transplanted into the uterus of the same foster mother, which raised both twins. Several physical characteristics, comparable to those described above, were measured after birth and compared with those of control mice. In genetically identical individuals, genetic variation is assumed to be eliminated, leaving only environmental stimuli to account for phenotypic differences, whereas in control (outbred) mice, phenotypic differences are due to both genetic and environmental differences. Hence, if both groups are held under identical conditions, the total variance within the group of control mice is expected to be greater than the variance within the group of genetically identical mice. However, the coefficients of variation appeared to be similar in both groups. Furthermore, measured environmental influences explained only 3 to 30% of the phenotypic variance.

In a similar experiment with Friesian cattle, an additional comparison was made with a group of divided embryos, which were transferred into, and raised by different uterine foster mothers. A large amount (70 to 97%) of the phenotypic variance remained unexplained. From these studies, Gärtner (1990) concluded that the remaining variance was due to influences other than genetic and (external) environmental influences.

Results obtained with cloned animals are consistent with the results of Gärtner (1990). For example, Archer, Dindot et al. (2003) compared cloned pigs with naturally bred controls on several phenotypic bodily traits as well as blood parameters. Controls were matched for age, breed, and sex, and were kept under identical conditions. Analysis of the phenotypic variance indicated that, compared to the controls, the cloned animals displayed as much or even increased variance in several traits. Although environmental conditions cannot be controlled fully, Archer, Dindot et al. argued that it is highly unlikely that (external) environmental effects could account for physical traits, such as hair growth pattern or skin type. They suggested that these differences in traits are caused by micro-environmental influences, minimal initial differences in uterine conditions, or small deviations that are introduced during cloning.

Besides phenotypic variation in genetically identical animals, studies of symmetry also reveal unexpected variation. Bilateral organisms often show intra-individual variation, which is known as fluctuating asymmetry. For example, Stige, Slagsvold, and Vøllestad (2005) repeatedly measured feather length and color patterns of wings and tail of the pied flycatcher. The degree of feather asymmetry persisted from nestling stage to adulthood, and even across moults. Genetic analyses revealed that the heritability of the asymmetry was almost zero, and that shared environmental factors had little influence. Moreover, differences in within-nest conditions did not explain the random variance in fluctuating asymmetry. The researchers concluded that asymmetry is possibly determined by (stochastic) events during early stages of development, permanently affecting the features of the feathers.

In humans, researchers have established small morphological asymmetries in various bilateral traits, such as bodily and facial features (Fink et al., 2004; Kowner, 2001). These variations emerge at a very young age and remain present throughout the lifespan. Furthermore, the direction of these asymmetries is not under genetic control. Fink et al. argued that facial asymmetries might be the results of perturbations in uterine conditions (e.g. hormone levels) in the first stages of embryonic development.

A related phenomenon, which counts as another indication of developmental instability, is the development of minor physical anomalies. Townsend et al. (2005), examining the dental records and radiographs of 278 monozygotic twin pairs, illustrated the contribution of developmental instability to phenotypic variability. They found that 24 twin pairs had developed missing or extra teeth. However, 21 of these pairs showed within-pair differences in patterns of expression (for instance, a mirrored effect). These researchers suggested that the observed differences in dental features of monozygotic twins stem from molecular interactions, leading to the initiation and, ultimately, the differentiation in the development of teeth.

Physical anomalies are also found in human brain organization and structure. Several studies have demonstrated the existence of atypical, non-genetically determined, asymmetries in human brain anatomy. For example, Eckert et al. (2002) examined the heritability of asymmetry in the planum temporale in 27 monozygotic and 13 dizygotic male twins. Magnetic Resonance Imaging (MRI) measurements revealed significant dissimilarities in gyral and sulcal features between monozygotic twins. The authors suggested that intra-uterine effects may lead to such variation in morphological development. Several other studies have replicated these results (e.g. Steinmetz, 1996; Thoma et al., 2002; Wright et al., 2002).

The unresolved variance originating in the process of development is not limited to biological or structural traits. Archer, Friend et al. (2003) compared the variance in several behavioral traits within litters of cloned swines and control (outbred) swines that were held under similar conditions. The traits included food preference, temperament, and time budgets such as feeding, standing, lying on concrete, lying in bedded area, and playing or fighting. The variance within the cloned group was as great as, or even greater than, the variance in naturally bred group. As we discuss in the next section, such unresolved variance may well characterize human psychological traits as well.

2.4 The Role of Nonshared Environment in Behavior Genetic Psychological Research

Researchers often distinguish five major domains of individual differences in psychological traits: cognitive abilities, personality, social attitudes, psychological interests, and psychopathology (McGue & Bouchard, 1998). Behavior genetic research within these domains has revealed a number of consistent, common findings (e.g. Turkheimer, 2000): First, a substantial amount of phenotypic variance can be attributed to genetic differences. Second, shared environmental variance is substantially smaller than the genetic variance. Third, a large portion of the phenotypic variance is attributed to nonshared environment. Moreover, over time heritabilities commonly increase, whereas environmentabilities decrease. However, the relative contributions of the shared component and nonshared component commonly change: If present, the contribution of the shared component decreases, often down to values close to zero. Hence, ultimately, most of the environmental variance is attributed to nonshared environment. Thus once genetic influences have been taken into consideration, siblings are ultimately hardly more similar than unrelated individuals drawn randomly from the population. Finally, MZ twins are often more than twice as similar as DZ twins, which often is explained by invoking dominance or epistasis (Turkheimer & Gottesman, 1996; Bouchard & McGue, 2003).

The amount of nonshared environmental variance ranges from about 20% in cognitive abilities in adults to about 60% in personality, phobias, and some social attitudes (Bouchard, 2004; Bouchard & McGue, 2003; McGue & Bouchard, 1998; Plomin et al., 2008). A meta-analysis of specific effects of objectively defined nonshared environmental variables (e.g. differences in sibling, teacher, and peer relationships) indicated that in behavioral genetic studies a considerable portion of nonshared environmental variance remains unexplained (Turkheimer & Waldron, 2000). The median percentage nonshared variance accounted for by these influences was no more than 2%. However, these effects are largely independent. So, "incorporating all of these measures of differential environment accounts for about 13 percent of the total variance of the outcome measures." (Plomin et al, p.312). Thus, for personality, for example, a large portion of nonshared environmental variance remains unresolved.

How can these findings be explained? Firstly, estimates of nonshared environmental variance often include error variance. However, if we take into account reliabilities of many psychological tests (i.e., 0.80 or better), nonshared environmental variance still is relatively large. Secondly, a distinction should be made between environmental *events* and environmental *effects* (e.g. Turkheimer & Waldron, 2000; Rutter et al, 1997; Turkheimer, 2000; Plomin et al. 2008; Harris, 1995). Shared environmental events may affect children differently for the simple reason that individual children may react to a shared experience in different ways. Thirdly, it has been argued that specific nonshared environmental factors become increasingly important during development. Harris, for example, theorizes that over time the influences of children's peers become relatively more important than parental influences in shaping personality. However, as systematic studies have met with only limited success (Turkheimer & Waldron; Plomin et al., 2008), such specific nonshared effects are difficult to identify (Plomin, Asbury, & Dunn, 2001).

Why it is so difficult to identify nonshared influences? According to Turkheimer and Waldron (2000) both the underlying process and its effects appear unsystematic. Hence, one source of such environmental effects may be (molecular) chance processes (Finch & Kirkwood, 2000). Another possible explanation is that a significant part of unsystematic variance may result from deterministic non-linear epigenetic processes (Molenaar et al., 1993). Third, a possible class of nonshared environmental influences on psychological differences may be pre- and perinatal microenvironmental influences (McGue & Bouchard, 1998), which include, *inter alia*, molecular and cellular processes. Clearly, any combination between these sources of variance should be possible. For example, Smith (1993) argues that the intrinsic dynamics of the brain, e.g. indirect influences of spontaneous synaptic changes in activity, partially explain developmental variability.

In sum, experiments with both inbred and isogenic animals and studies of developmental instability in both humans and animals have produced converging evidence of unsystematic sources of phenotypic variance and dependency on initial conditions. In part, the ubiquitous nonshared environmental variance in psychological traits may in fact be due to nonlinear influences. Since the outcome of such processes is unsystematic (i.e., appears stochastic), these influences are hard to detect. This may explain why attempts to identify such influences have met with limited success (Turkheimer & Waldron, 2000). As suggested by Molenaar et al (1993), a part of the unsystematic phenotypic variance may actually be the result of nonlinear self-organizing epigenetic processes. Simulation studies of nonlinear epigenetic processes have been used to study the role and effects of such processes on behavior genetic analyses (Eaves et al., 1999; Molenaar & Raijmakers, 1999. The implications of these studies are discussed in the next section.

2.5 Computer Simulations of Non-linear Epigenetic variance

Both Eaves et al. (1999) and Molenaar and Raijmakers (1999) simulated nonlinear epigenetic variance in development using the discrete logistic equation as a simple model of development. This model can give rise to chaotic behavior, hence, these researchers were able to study the effects of sensitivity to initial conditions on twin correlations. In short, the main effect was a lowering of initial twin correlations. Specifically, intermediate or low (i.e., DZ) within pair correlations decreased at higher rates and to lower levels than high (i.e., MZ) within pair correlations (Eaves et al.). Genetic analyses revealed that, with time, the genetically induced correlations between the parameters had been concealed in the observed output: Observed heritabilities decreased to zero (Molenaar & Raijmakers).

From a theoretical and practical biological point of view, the model considered by Molenaar and Raijmakers (1999) and by Eaves et al. (1999) is limited. First, the discrete logistic equation prescribes that the current state depends on the previous state, but what biological or psychological state it denoted, was not specified. Second, the real time interval between the subsequent time steps was undefined. More importantly, in reality development does not unfold in discrete time. Third, the model displays chaotic behavior. According to Eaves et al., traits affected by chaotic processes should display DZ correlations close to zero while MZ correlations are high; a finding that is not frequently observed in human developmental research. In addition, over time chaotic processes will yield

observed heritabilities of 0 (Molenaar & Raijmakers), which is not observed in behavior genetic research. Hence, as Eaves et al. argued, the role of chaotic epigenetic processes in development may be quite limited. The possibility remains, however, that other types of non-linear processes, e.g. bifurcating systems, give rise to a different picture.

In this light, the computer simulations of Turkheimer and Gottesman (1996) are of interest. To study the dynamics of genes and environment in development, they considered a non-linear model that does not give rise to chaotic behavior, but that does display bifurcating behavior. Genotype, environment, and phenotype were conceived of as locations (dots) in two-dimensional space. Whereas genotype was simulated as a fixed location, environment and phenotype changed dynamically according to a set of simple rules, such as *phenotype is dynamically attracted to genotype*, *phenotype is dynamically attracted to environment*, and *environment is dynamically attracted to phenotype*. The relationship between locations of genotype and environment on the one hand, and of phenotype on the other hand, appeared highly unsystematic while the ultimate locations depended heavily on starting conditions.

Turkheimer (2000) conducted a twin simulation on the basis of this model. A significant part of phenotypic variance could be accounted for by genotypic variance, but none of it could be explained by the variance in environmentally determined starting locations. Furthermore, phenotypic DZ within-pair correlations were substantially smaller than MZ within-pair correlations. However, the DZ correlations were not close to zero as in the studies of Eaves et al (1999), and Molenaar and Raijmakers (1999). Although Turkheimer did not present his model explicitly as a model of nonlinear epigenetic variance, one may interpret it as such, especially in the light of his conclusion:

> "[P]henotype at any moment in development is the cumulative result of an organism's developmental history, encompassing genotype, environment, and all the complexities of their epigenetic interactions" (Turkheimer, 2000, p. 184).

Our present aim is to study the role of nonlinear epigenetic processes using a model of neurite outgrowth. The model shows bifurcating, but not chaotic, behavior and is biologically realistic: Time is modeled as continuous, the time-scale is defined, and the system's output (the state of an organism) represents concrete characteristics (e.g. membrane potentials). Furthermore, previous computer simulations employing this model have demonstrated results that are consistent with empirically observed phenomena, such as a temporarily overproduction of synapses (see van Oss & van Ooyen, 1997). With our simulation study, we reconsider the plausibility of nonlinear epigenetic processes as a source of unsystematic phenotypic variance. Furthermore, by linear behavior genetic modeling, we estimate the contributions of genetic and environmental components, as well as their interactions. Below, we discuss the consequences of epigenetic processes for behavior genetic analyses, and the search for specific genes and environmental influences.

2.6 Modeling Network Development

The brain may be viewed as a highly organized neural network, comprising numerous structures and connections. As mentioned, the organization of the brain represents more information than is encoded in the genes, so this suggests that brain development involves self-organizing processes. As argued above, such processes may constitute a potential important source of individual differences.

During development neurons are assembled into functional network structures (for a review see, e.g. van Ooyen, 1994). One important factor determining neuronal morphology and network formation is intrinsic electrical activity; a mutual influence exists between network activity on the one hand, and neuronal form, connectivity, and function, on the other. That is, a network is able to generate patterns of activity, thereby changing the organization of the network, which in turn leads to an alteration of activity patterns. These can further modify the network's structure or function, and so on. One example of a model of such activity dependent neural network development is the two-cell model of van Oss and van Ooyen (1997) that models neurite outgrowth.

Neurite outgrowth concerns the development of axons and dendrites, i.e., the development of connections between neuronal cells through which electrical activity is transmitted from one cell to

another. Van Oss and van Ooyen (1997) used their two-cell model to investigate the effects of the combination between activity dependent neurite outgrowth and inhibition. They showed that this combination can account for multistability, which they associated with normal and pathological end states of network development. This model has the characteristics that are considered to underlie epigenetic variance. A mathematical description of the model and its behavior is given in Appendix A. Here we limit ourselves to an informal description.

The two-cell model contains one excitatory and one inhibitory unit or cell (set out graphically in Figure 2.1). More complicated (network) versions of this model have been used in simulating biological, structural properties of the developing nervous system (van Ooyen et al., 1995; van Ooyen & van Pelt, 1996). The two-cell model is based on the behavior of populations of neurons, and was developed in order to replace the network model with its numerous differential equations with a simpler model that still displays the characteristics and behavior of this network model. A unit may thus represent a population of neurons (see also, Wilson & Cowan, 1972; Ghosh, Chang, & Liano, 1998). Comparable two-neuron models have been used to study working memory. Kirillov, Myre, and Woodward (1991) examined the functioning of inhibitory-feedback networks as programmable memory devices using a 1000-neuron model and observed rich dynamic behavior and multi-stability. In order to focus on collective behavior rather than on the behavior of individual neurons they developed a simplified two-neuron model of this network model. In subsequent research (Kirillov, Myre, & Woodward, 1993) they studied the behavior of this two-neuron model and concluded that "[m]any of the […] results generalize to N-neuron interconnected models" (p. 449).

The interpretation of a unit representing a population of neurons can be justified on biological grounds; in many areas of the brain, neurons are organized in populations of neurons with similar properties. For example, this is the case in the somatosensory cortex (Mountcastle, 1957), the visual cortex (Hubel & Wiesel, 1962), and in pools of motor neurons (Kandel, Schwartz, & Jessel, 2000).

As mentioned, the two-cell model of van Oss and van Ooyen describes activity-dependent neurite outgrowth, one of the many dynamical processes involved in shaping neuronal morphology. Excitatory and inhibitory inputs take the membrane potential towards a finite maximum and minimum potential, respectively. The excitatory unit is connected to itself (with weight w_{xx}), which can be taken to represent reciprocal excitation between excitatory neurons. For simplicity, reciprocal inhibition between inhibitory neurons is not modeled (weight $w_{yy} = 0$), i.e., the inhibitory unit is not connected to itself. In its most simple form, the connection between the inhibitory unit and the excitatory unit is modeled as symmetric (weight $w_{xy} = w_{yx}$).

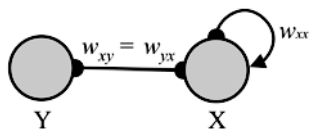


Figure 2.1 The two-cell model, where unit Y is interpreted as a population of inhibitory neurons and unit X as a population of excitatory neurons. X is connected to itself with weight w_{xx}. In contrast, Y is not connected to itself (mutual inhibition is not modeled). The connection between X and Y is modeled as symmetric ($w_{xy} = w_{yx}$). In the simulations, $w = w_{xx} = w_{xy} / p$.

A unit's output is the mean firing rate, which is modeled as a sigmoidal function of the membrane potential. Parameters determine the steepness of the function as well as a low sub-threshold firing rate. The latter can be interpreted as spontaneous activity as a result of synaptic noise or fluctuations in membrane potential. Furthermore, increasing neurite outgrowth is thought to imply increasing connection strength. In turn, as connection strength increases, the amount of input that a cell receives through this connection also increases. Hence, the system is characterized by feedback. Finally, the dynamics of the connection strength are considerably slower than those of neuronal activity. Therefore, connection strength can be considered as a slowly varying parameter.

Using their model, van Oss and van Ooyen (1997) investigated occurrences of bifurcations in neuronal development. Such occurrences imply that individual networks may develop via different

pathways, dependent on initial conditions. Van Oss & van Ooyen studied the behavior of the network as a function of the parameters that determine the membrane potential at which neurite outgrowth is 0 (ε), and the relative strength of the inhibitory connection (p) (see Appendix A for further details). They found bifurcations, multistability, and rich dynamic behavior, including oscillations and transient jumps.

Empirical studies (see e.g. van Ooyen & van Pelt, 1994) have shown that, under normal conditions, developing cultures of dissociated cortex cells display a transient overproduction, or overshoot, of synaptic connections. An initial stage of neurite outgrowth and synapse formation is followed by a period of a considerable retraction of synapses towards a final stable state. In contrast, when a developing culture is deprived of electrical activity for a certain amount of time, neurite outgrowth is enhanced. In this case retraction of synapses does not take place, even if electrical activity is restored. This suggests there may exist a critical period after which electrical activity no longer results in retraction of synaptic connections. These results imply that multistability exists, because differences in initial conditions (i.e., differences in exposure to electrical activity) result in different stable developmental pathways (i.e., differences in neurite outgrowth).

Due to the relative simplicity of the model, van Oss and van Ooyen (1997) were able construct a bifurcation diagram, in which areas in the parameter plain (ε, p) are associated with qualitatively different dynamic behavior of the model. These areas are demarcated by bifurcations. The realistic nonlinear characteristics of this model of neuronal development and the knowledge of its dynamics render it very suitable to study the generation of nonlinear epigenetic variance in a simulation experiment within a behavioral genetic design. Moreover, the systematic bifurcation analysis gives us good insight in the dynamics of the model that are relevant to the interpretation of the results of behavioral genetic analyses. We retained all model assumptions, so as to ensure that the bifurcation diagram described the model.

The model has several characteristics, which help to illustrate the theoretical point that epigenetic processes constitute a source of variance. First, due to the reciprocal interactions and the sigmoid activation functions, the model is non-linear. The specific properties of the model that may give rise to epigenetic variance are the occurrence of bifurcations (the possibility that small differences in parameter values, e.g. small differences in enzyme concentrations, result in qualitatively different phenotypes), and the presence of multistability (initial differences can result in different developmental pathways). The existence of qualitatively different possible phenotypes given one genotype implies that members of a MZ twin pair may follow different developmental pathways. Second, the model displays properties of self-organization. Third, the model describes development as a deterministic process. This is important theoretically, because it means that the sensitivity to initial conditions does not arise from any randomness, i.e., the epigenetic variance is not attributable to an external random process that introduces noise into the system.[4].

Using the two-cell model of van Oss & van Ooyen (1997), we aimed to simulate twin data in the form of time series. These time series were generated under various assumptions concerning starting values, parameter values, and degree of genetic control in order to create nonlinear epigenetic variance. In addition, we addressed another question of interest in this context: what consequences do these processes have for the estimated heritabilities and environmentabilities in the standard twin model?

2.6.1 Twin Simulations Applying The Two-Cell Model

Method Parameter values, the sigmoid function (see equation 3 in Appendix A), and the form of the differential equations constituting the two-cell model (see equations 1, 2, and 4 in Appendix A) give rise to the non-linear epigenetic process of the development of phenotypic traits. These traits are x, the time averaged membrane potential of the excitatory unit, y, the time averaged membrane potential of the inhibitory cell, and w, the connection strength between excitatory neurons. The variances of the parameters ε and p are attributable to genetic and environmental individual differences within the population. That is, individual differences in ε and p satisfy the standard behavior genetic model (see

[4] This is not to say that developmental processes are noise-free.

above). Thus, individual differences in ε and p are the causes of individual differences in the development of phenotypic traits x, y, and w. Starting values of x, y, and w were set at 0.

Phenotypic time series were generated for 200 MZ and 200 DZ twin pairs using a simple additive model: $\varepsilon_i = \varepsilon_{g,i} \sqrt{a} + \varepsilon_{c,i} \sqrt{b} + \varepsilon_{e,i} \sqrt{(1 - a - b)} + m_\varepsilon$, and $p_i = p_{g,i} \sqrt{a} + p_{c,i} \sqrt{b} + p_{e,i} \sqrt{(1 - a - b)} + m_p$, with $0 \leq a \leq 1$, $0 \leq b \leq 1$, and $a + b \leq 1$, where ε_i and p_i are the individual values of the parameters ε and p, respectively, and m_ε and m_p are the sample means of the parameters. The subscripts g, c, and e refer to the additive genetic, shared environmental, and nonshared environmental components of a parameter, with regression coefficients $\sqrt{a}$, $\sqrt{b}$, and $\sqrt{(1 - a - b)}$, respectively.[5] Assuming that the additive genetic and environmental components of a parameter are uncorrelated and sampled randomly from the standard normal distributions, heritability h^2 equals a, common environmentability c^2 equals b, and unique environmentability e^2 equals $(1 - a - b)$. MZ twin correlations of the additive genetic, shared, and nonshared environmental components equal 1 (100% additive genetic resemblance), 1, and 0, respectively. The same procedure was followed to generate 200 DZ twins, except the DZ twin correlations of the additive genetic components equals 0.5.

In matrix notation (e.g. for the parameter p) this can be written as:

$$\mathbf{p} = \begin{bmatrix} \sqrt{a} & \sqrt{b} & \sqrt{(1-a-b)} & 0 & 0 & 0 \\ 0 & 0 & 0 & \sqrt{a} & \sqrt{b} & \sqrt{(1-a-b)} \end{bmatrix} \cdot \begin{bmatrix} p_{g,1} \\ p_{c,1} \\ p_{e,1} \\ p_{g,2} \\ p_{c,2} \\ p_{e,2} \end{bmatrix} + \begin{bmatrix} m_p \\ m_p \end{bmatrix}$$

With covariance matrix

$$\boldsymbol{\Sigma}_p = \begin{bmatrix} a \cdot \sigma^2_{p_g} + b \cdot \sigma^2_{p_c} + c \cdot \sigma^2_{p_e} & r_g \cdot a \cdot \sigma^2_{p_g} + b \cdot \sigma^2_{p_c} \\ r_g \cdot a \cdot \sigma^2_{p_g} + b \cdot \sigma^2_{p_c} & a \cdot \sigma^2_{p_g} + b \cdot \sigma^2_{p_c} + c \cdot \sigma^2_{p_e} \end{bmatrix}$$

Where subscripts 1 and 2 refer to each of the members of a twin pair and r_s denotes the genetic resemblance.

In simulating the data, we chose a heritability of 0.5 for both parameters p and ε. Furthermore, referring to the epigenetic picture and to non-linearity, initial conditions may be very similar. However, in real biological systems exactly equal initial conditions are impossible. This would require an infinite amount of information (or energy), which is biologically implausible (Molenaar & Raijmakers, 1999). Therefore, we set unique environmentability to a low, but nonzero, value (i.e., to 0.01). As a consequence, common environmentability was set to a value of 0.49. This procedure implies that in the absence of nonlinearity, standard behavior genetic analyses of the phenotypic measures should be expected to yield a heritability estimate of 0.5, a common environmentability of 0.49, and a unique environmentability of 0.01. However, given the occurrence of bifurcations and sensitivity to initial conditions, additional (i.e., epigenetic) variance is expected to arise.

The differential equations were numerically solved using the variable time step integrator Lsoda (Petzoldt, 2003). Subsequently, phenotypic twin correlations of x, y, and w were obtained at timepoints $t = 0$ to 20000 (by t = 20000 all networks have reached their final stable states). Regular oscillations can and do occur in these measures, which are a consequence of switches between two states. Latency shifts in these oscillations complicate statistical analyses: At certain points of time some twin members may show behavior that is out of phase or even in anti-phase, whereas the signals

[5] One usually would implement environmental influences in a network by input of external activity. However, in doing so, the bifurcation diagrams would no longer apply to the model. In addition, this point is not relevant to the hypothesis that non-linear epigenetic processes constitute a source of variance.

over time may be highly similar. Such occurrences could lower twin correlations a bit when one would sample the raw measures at an arbitrary time point. We consider this as a kind of measurement error. In order to minimize such error, we derived alternative phenotypic measures. These are the mean of the signal over a certain period of time and the relative power. The mean signal is thus a smoothed signal, where high frequency oscillations have been filtered out. The power of a signal is a time average of energy (energy per unit time), where energy is defined by the area under the signal. Relative power is the power of one frequency band relative to that of other frequencies. In this manner we could distinguish (almost) non-oscillating signals (frequencies < 0.1) from oscillating ones (frequencies ≥ 0.1). In order to calculate the two alternative phenotypic measures, time was divided into 40 equal intervals: [0, 500), [500, 1000), …., [19500, 20000). Twin correlations were calculated in each time window.

In simulation 1 we chose region 5b within the (ε, p)-plane as region in which variance was to be created in parameters ε and p. This region is shown in Figure A.2 (see Appendix A). In this region one point attractor and two stable limit cycles exist. Such a limit cycle can be viewed as a switching between two states. With respect to epigenetics, we note that the system may follow multiple developmental pathways. However, which path is followed depends on the parameter values. A second reason for choosing region 5b was that connection strength grows to finite values, which is biologically more plausible than infinite growth. A third reason was that in this region the occurrence of overshoot is possible. As noted above, overshoot is a widespread phenomenon in empirical studies.

The means of parameters ε and p were set at 0.52, and 0.42, respectively. This parameter combination is located approximately in the centre of region 5b. In this way, the variance of the parameters could be taken as large as possible (0.0005 for both parameters). In the interest of biological plausibility, non-varying parameter values were set identical to those in the study of van Oss and van Ooyen (1997), i.e., $h = 0.1$, $\theta = 0.5$, $\alpha = 0.1$, and $b = 0.00005$. An exception was made for the value of q, which was set at 0.05.[6] This parameter determines the outgrowth rate (see Appendix A). The use of lower values would slow down outgrowth and, thus, greatly increase the number of computations. Figure 2.2 depicts the behavior of x (left) and y (middle) for a number of parameter values for $t = 19800$ to $t = 20000$ as well as the development of w from $t = 0$ to $t = 20000$. Distinct behavior can be discerned, which is due to the existence of different types of attractors within one region.

In simulation 2 we chose region 5a as parameter region. Region 5a differs only from region 5b in that the limit cycles are absent. Only the point attractor exists, hence multistability is not present. We set the parameter means of ε and p to values of 0.8, and 0.3, respectively. The variances of parameters ε and p as well as other parameter values were identical to the values used in simulation 1. Figure 2.3 depicts the behaviors of x (left), y (middle), and w (right) from $t = 0$ to $t = 20000$ for these parameter values. We carried out simulation 2 to compare the differences between the amount of unexplained variance in case of multistability with the amount of unexplained variance in case of monostability. This difference can be regarded as the amount of nonlinear epigenetic variance.

Results To determine discordance between twins, we used the relative power of x. A twin pair shows discordant behavior if the one member of a twin pair displays oscillating behavior, whereas the other member does not. Twin correlations were based on the filtered (smoothed) signals. The filtered signals of x and y still showed some oscillating behavior. As explained above, oscillations may lead to latency shifts, which yield slight overestimates of unexplained variance (i.e., due to what we above have denoted measurement error). In order to minimize the effects of such error, we here mainly concentrate on the analyses of the filtered signal of w, but it should be mentioned that the other phenotypic measures gave similar results.

As is common in behavior genetic research, we decomposed phenotypic variance into latent factors by fitting an ACE twin model, where A stands for additive genetic effects, C for common, or shared, environmental effects, and E for nonshared environmental effects. We did so over each of the 40 time intervals. In addition, in each simulation two stepwise regression analyses were carried out on

[6] As compared to the results of simulations using the original value of q (0.005), using the value of $q = 0.05$ did not give rise to qualitative different behavior of the system, or to computational artifacts.

the filtered signal of w, over the last time interval ($19500 \leq t < 20000$). To describe the amount of variance explained by the model, we calculated the adjusted squared multiple correlation coefficient ($R^2_{adjusted}$) between all of the predictors and the dependent variable. We calculated the amount of unexplained variance as ($1 - R^2_{adjusted}$) * 100%. In the regression model, the predictors were the variables p_g, p_c, p_e, ε_g, ε_c and ε_e. As above the dependent variable was the filtered signal of w. In the first analysis, the linear regression model (henceforth, referred to as 'G+E model') includes solely main effects for all independent variables. In addition to main effects, the second regression model (henceforth, referred to as 'G×E model') includes interaction terms for all possible combinations for all independent variables.

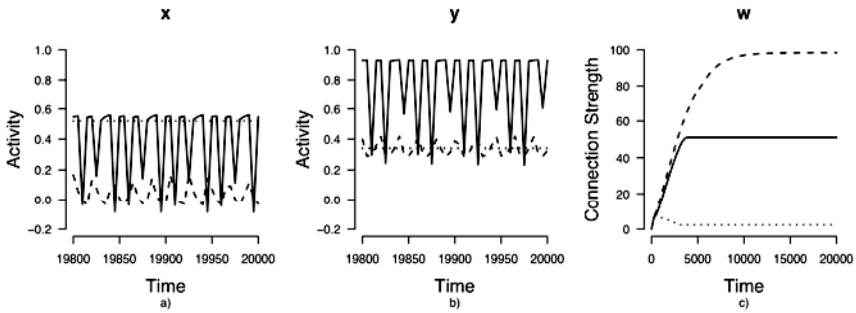


Figure 2.2 The behavior of x (a), y (b), and w (c) of developing networks in case of multistability. Parameter values are $\varepsilon = 0.52$, and $p = 0.435$ (straight lines), $p = 0.420$ (dashed lines), and $p = 0.402$ (dotted lines).

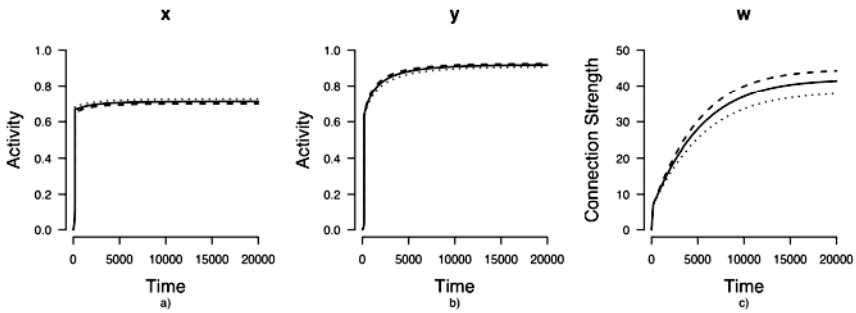


Figure 2.3 The behavior of x (a), y (b), and w (c) of developing networks in case of monostability. Parameter values are $\varepsilon = 0.8$, and $p = 0.315$ (straight lines), $p = 0.300$ (dashed lines), and $p = 0.282$ (dotted lines).

Simulation 1: subregion 5b In simulation 1, multiple stable attractors were present. Simulations using the two-cell model gave rise to bifurcations and multistability. The presence of bifurcations implies that small differences in parameter values could result in qualitatively distinct behavior; differences in developmental pathways and differences in final stable states (i.e., oscillating versus non-oscillating behavior). Twin members could thus follow discordant developmental pathways. At the end state, 27.5% of the MZ twin pairs, and 33% of the DZ twin pairs showed discordant behavior. In addition, even in the absence of noise, with time a substantial and increasing amount of unsystematic (nonlinear epigenetic) variance arose, which was attributed to the nonshared component in a linear behavior genetic analysis.

The presence of bifurcating processes resulted in a lowering of twin correlations over time. Figure 2.4 illustrates the development of these correlations with respect to the filtered signal of w, together with the development of observed heritabilities and environmentabilities. As one can see, relatively high initial correlations (MZ correlations) are affected less than relatively low initial correlations (DZ correlations). In other words, compared to DZ twin similarity, MZ twin similarity was better preserved. As a consequence, observed common environmentabilities decreased to zero,

whereas heritabilities and unique environmentabilities increased over time. Although not depicted, the development of twin correlations of other phenotypic measures yielded qualitatively the same picture.

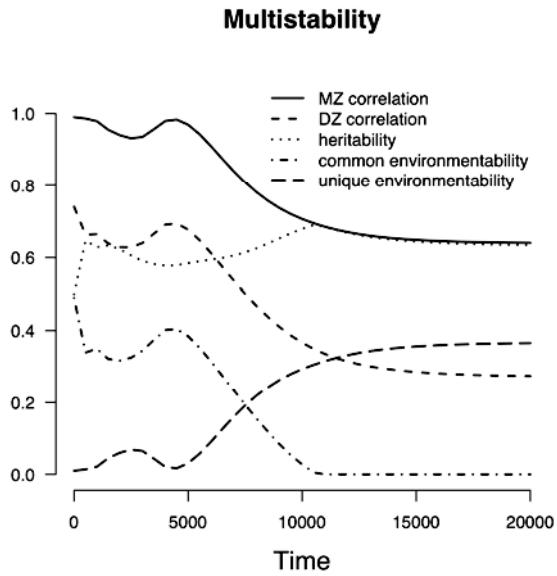


Figure 2.4 The development of MZ and DZ twin correlations, observed heritabilities and observed environmentabilities of the filtered signal of *w* in case of multistability.

Fitting an ACE-model to the twin covariance matrices of the filtered signal of *w* at the end state yielded good fits (χ^2 (3) = 0.495, p = 0.91). The variance attributed to the nonshared environmental component increased substantially (to 36.6%). The variance attributed to genetic effects also increased (to 63.4%), whereas estimated shared environmental variance had dropped to 0.00 %. Dropping the shared environment component (C) from the model did not alter the model fit.

These results are in concordance with those of the stepwise regression analyses: In the G+E regression model on the filtered signal of *w* at the end state, a large part of the variance remained unexplained (58.1% in MZ twins and 63.3% in DZ twins). The model that included interaction terms (GxE model) did little to decrease the unexplained variance (53.1% in the MZ twins, and 62.5% in the DZ twins). In other words, GxE interaction did little to account for the residual variance. Thus, nonlinear epigenetic variance was not explained by genotype-environment interaction.

As in latent variable modeling, in the regression analyses the genetic components (p_g and ε_g,) explained a significant part of the variance in the dependent variable, the filtered signal of *w*. However, in contrast to latent variable modeling, in the regression analyses the shared environmental components could explain a significant part of the variance. Hence, we may conclude that the effects of shared environmental influences were not detectable in a standard twin design.

Simulation 2: subregion 5a In simulation 2, one stable (point) attractor was present. Simulations using the two-cell model revealed the conseuqences of the absence of bifurcations and multistability. Differences in parameter values did not result in qualitatively different phenotypes or different developmental pathways. As a consequence, discordance between twins was 0. Furthermore, nonlinear epigenetic variance did not arise. Behavior genetic models yielded good fits and (nearly) all variance was explained.[7].

[7] Since we described just a single simulation in each of the series of experiment, we could not quantitatively test the amount of non-linear epigenetic variance against the null-hypothesis of an unexplained variance of 0. To do so, each simulation should be repeated applying a full Monte Carlo design with sufficient numbers of replications.

Figure 2.5 illustrates the development of phenotypic twin correlations of the filtered signal of *w*, together with the observed heritabilities and environmentabilities. Twin correlations remained stable over time. As a result, observed heritabilities and environmentabilities accurately reflect the initial genetic and environmental structure. The results in terms of twin correlations obtained with the other phenotypic measures were largely the same.

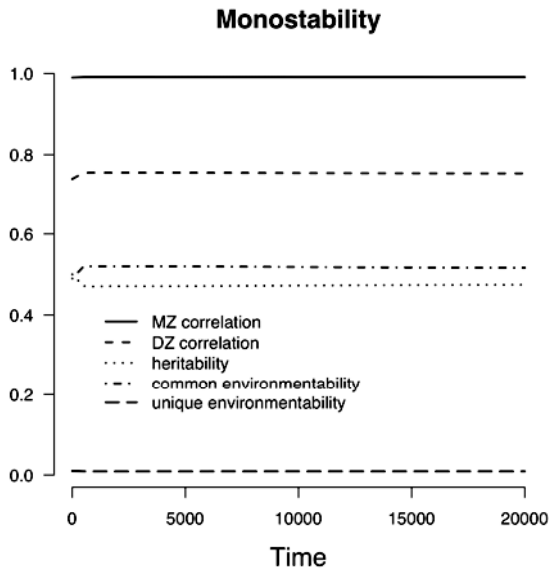


Figure 2.5 The development of MZ and DZ twin correlations, observed heritabilities and observed environmentabilities of the filtered signal of *w* in case of monostability.

Fitting an ACE-model on the filtered signal of *w* at the end state yielded perfect model fit (χ^2 (3) = 0.00, p = 0.99). The percentage of variance attributed to genetic effects was 47.5%, whereas the percentages of variance attributed to the shared and nonshared environmental components were 51.7% and 0.90%, respectively. These results are in concordance with the results from the regression analyses: The G+E regression model on the filtered signal of *w* at the end state explained nearly all of the variance (99.95% for the MZ twin data, 99.94% for the DZ twin data). The G×E regression models yielded about the same results: 99.99% of both the MZ and DZ twin data was explained. As expected, both genetic components (p_g and ε_g,) and shared environmental components (p_c and ε_c,) of the parameters were good predictors of the filtered signal of *w*.

Conclusion In simulations 1, we found that nearly all the variance that was attributed to unshared environmental effects in the twin model, was actually due to nonlinear epigenetic variance. This variance was due specifically to the disproportional relationships between causes and effects within the nonlinear process. Furthermore, we established that nonlinear epigenetic variance is distinct from variance due to genotype-environment interaction. Specifically, the interaction in our regression analyses did little to account for any of the variance. In conclusion, ultimate genetic and environmental causes of phenotypic variance were not detected in the standard twin design.

However, we carried out our simulations more than once and found comparable results. Hence, we consider our corresponding qualitative comparisons as most indicative.

2.7 Discussion

In the present chapter, we reviewed the literature on empirical evidence that suggests that non-linear (epigenetic) processes during ontogenesis account for a substantial portion of phenotypic variance. Such processes may account for a part of the ubiquitous nonshared environmental variance in twin studies of psychological traits. Second, we simulated twin data using a biological realistic model of neurite outgrowth, which displayed the characteristics that were hypothesized to underlie nonlinear epigenetic variance. The results of our simulations demonstrate that the existence of multiple attractors due to bifurcations gives rise to phenotypic variance in simulated neuronal development. This variance cannot be attributed to genetic or environmental components of the parameters, to measurement error, nor to any other external random process that introduced noise into the system. Hence, the variance can be interpreted as nonlinear epigenetic variance. In the standard behavior genetic (twin) model, this variance is subsumed under the nonshared environmental component.

In simulation 1, the existence of multistability and the related sensitivity to initial conditions gave rise to nonlinear epigenetic variance, which caused the twin correlations to decrease over time. The variance was unsystematic, thus remained unexplained in the regression analyses. The actual contributions of the genetic and environmental factors to the parameters in the model were not recovered in the standard twin model. The depressing effects of nonlinear epigenetic variance were greater on DZ than on the MZ twin correlation. As a consequence, over time heritabilities appeared to increase, and shared environmentabilities appeared to decrease to zero. In simulation 2, the absence of multistability ensured that we could recover the genetic and environmental contributions to the parameters well.

Our results are in concordance with the empirical evidence we reviewed, in which a substantial portion of phenotypic variance, provisionally attributed to unshared (nonshared) environmental influences, actually remains unexplained. This variance may be due to nonlinear sources of variance, which will include endogenous molecular or cellular processes. The hypothesis that non-linear epigenetic processes represent an appreciable source of variance is consistent with the fact that nonshared effects have proven quite hard to identify (Turkheimer & Waldron, 2000).

The results of the simulations are also consistent with the common findings in behavior genetic studies in psychology. That is, over time the estimated relative contribution of nonshared variance to phenotypic variance increases, whereas shared variance decreases to zero. As a consequence, estimated heritabilities increased, which is often observed in longitudinal twin studies (e.g. Bouchard & McGue, 2003). Finally, monozygotic twins were more than twice as similar as dizygotic twins, which is also frequently observed (Turkheimer, 2004).

The consequences of the presence of nonlinear epigenetic variance may be far reaching. The exact influences on initial conditions will be generally untraceable in standard behavior genetic modeling (of twin data, say), because the effects are unsystematic. Moreover, because the effects of nonlinear epigenetic processes may result in changes in the environmental variance over time, investigators might be inclined to seek explanations in terms of environmental factors. The same applies to the search for genetic factors. Again, due to the effects of nonlinear epigenetic processes, genetic variance may change over time. The standard interpretation of changes in heritability (say in terms of the ACE model) may be sought in specific genetic hypotheses, such as the switching on/off of genes (see also Eaves et al, 1999), while the ultimate causes lie elsewhere. Or, given DZ correlations which are markedly low compared to the MZ correlations, investigators may assume the presence of genetic interaction effects such as dominance or epistasis, whereas these may be absent (see also Turkheimer & Gottesman, 1996).

In the traditional behavior genetic model, genetic effects, shared and nonshared environmental influences are subject to a linear model, which yields the usual decomposition of phenotypic variance. A further decomposition of nonshared environmental variance into linear nonshared environmental variance and non-linear (epigenetic) variance would further increase our understanding of what we currently denote as nonshared environmental variance. Figure 2.6 depicts a schematic representation of this. Since their effects will be traceable in principle, measured nonshared influences can be incorporated in the model. In contrast, nonlinear influences act unsystematically and, as a consequence, these will difficult to trace back.

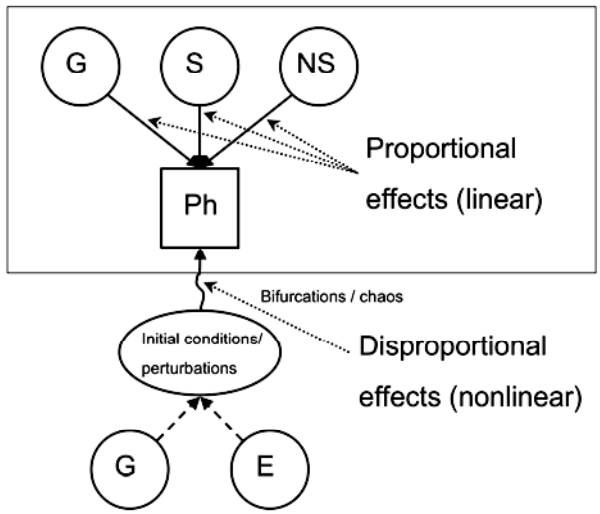


Figure 2.6 Phenotype (Ph) as a function of linear and nonlinear influences. In a linear behavior genetic model, the environmental component is decomposed into shared environment (S) and nonshared environment (NS). Error variance is subsumed under the nonshared component. In contrast to the modeled linear (proportional) effects, chaotic and bifurcating (epigenetic) processes will have unsystematic (disproportional) effects. These will be difficult to distinguish from error. The genetic (G) and environmental (E) influences on initial conditions or perturbations will be hard to detect. Although the ultimate influences on phenotype are only genetic and environmental, nonlinear (epigenetic) processes constitute a distinct and independent source of variance.

One solution to this problem, within the context of the present model, is to categorize the networks into classes according to their qualitative behaviors, which depend on the network's attractor. One can then fit a linear behavior genetic model to the data within each class. Within each class, at each point of time, the genetic and environmental contributions to the parameters can be recovered quite well. In theory, this can be achieved by fitting mixture models, which allows one to identify latent classes and fit the standard twin models within these classes. Whether this is actually viable in twin data is an open question, but this does provide an interesting perspective on the recent interest in mixture modeling of twin data (Gillespie & Neale, 2006; Muthén, Asparouhov, & Rebollo, 2006).

Our interpretation of nonshared environmental variance as consisting in part of non-linear epigenetic variance is consistent with Molenaar et al. (1993), who consider nonlinear epigenetic processes as a third source of variance, alongside genetic and environmental influences. According to the definition of nonshared variance in behavior genetics, non-linear epigenetic variance is subsumed under the nonshared environmental component even though initial conditions may be determined mainly by genetic effects and shared environmental influences. However, the ultimate causes (sources) of phenotypic differences are thus only genetic and environmental. Hence, to refer to nonlinear epigenetic variance as a third source is a matter of definition.

One important issue concerns the scalability of nonlinear epigenetic processes. Since the processes we simulated were low-level processes, it is important to consider how nonlinear effects may accumulate and combine. Such effects can average out to produce a kind of statistical macro-level behavior (that may show stage/phase transitions), or they can amplify each other as they do in chaotic systems (e.g. in a system such as the weather system). In chaotic epigenetic systems the genetic structure initially present will be destroyed, i.e., twin correlations will approach zero, a result not frequently observed in empirical studies (Molenaar & Raijmakers, 1999). The role of chaos in development may thus indeed be limited, as Eaves et al. (1999) argue. However, if within a nonlinear system only a limited number of bifurcating processes are present, e.g. a limited number of stage transitions, critical periods, or developmental pathways, a different picture may arise. In that case,

differences in (initial) conditions may only have an effect at critical points (bifurcation points), even when they are very small. At other points, differences in conditions may not affect development, even when they are very large. Our review of the literature indicated that initial conditions are of great importance, whereas phenotypes are often stable over time. These findings would support such an account.

One possible step researchers may take to investigate whether nonlinear (epigenetic) processes contribute to nonshared variance in a trait is to study the developmental trajectory of that trait. Ontogenesis invariably involves self-organizing processes, and self-organization invariably arises in nonlinear systems. Given that bifurcations (stage/phase transitions) are a common characteristic of nonlinear dynamics, the presence of critical periods or stage transitions strongly suggest that the process is nonlinear. Van der Maas and Molenaar (1992; see also Gilmore, 1981) discuss in considerable detail the detection, using catastrophe flags, and the classification of stage transitions and the ways to distinguish between transitions due to nonlinear dynamics and mere acceleration in growth (e.g. growth spurts). We believe that these can be used to study complex human behavior, such as Piagetian cognitive development, motor development, and language development. The use of a longitudinal twin design, would allow one to study dynamics of development by means of the catastrophe flags, and relate these to changes in the nonshared environmental covariance structure.

Another issue concerns the generalizability of our simulation experiments. We studied the effects of nonlinear epigenetic processes in an unsupervised, self-organizing network. Within neural network modeling, many networks are supervised systems, which are generally not self-organizing systems (but see van der Maas, Verschure & Molenaar, 1990; Raijmakers & Molenaar, 2004; Pollack, 1991; Rodriguez, Wiles, & Elman, 1999; Daucé, 2007). Since, non-linear epigenetic processes were considered to involve self-organization, nonlinear epigenetic variance may be expected to be less important in supervised systems. However, to conceive the whole organism as fully supervised systems is implausible. Indeed, in real biological systems (fully) supervised networks may exist, but they will be coupled with and thus dependent on unsupervised or mixed supervised/unsupervised networks. There is no reason to expect that nonlinear epigenetic variance will not arise in these systems. Hence, we may find nonlinear epigenetic variance at all levels, biological as well as psychological, just as bifurcations (phase transitions and stage transitions) are found at all these levels. Kelso stresses repeatedly:

> "[O]ver and over again nature uses the same principles of self-organization to produce dynamic patterns on all scales. The precise patterns that form may differ from one scale of observation to another, but the basic principles are the same." (Kelso, 1995, p. 24).

Indeed, he described bifurcations at molecular levels in the activity of ion channels, at psychological levels, in perception, motor coordination, speech, language acquisition, learning, working memory, and cognition, and even at the level of interindividual behavior, for instance in case of social coordination.

Furthermore, (critical) periods may exist, before the inception of any supervision, during which sensitivity to supervision is developed. This suggestion is compatible with van Oss & van Ooyen's (1997) account of their simulation experiments, in which they study the existence of critical periods. As mentioned above, under normal conditions, developing cells display an overproduction of connections followed by a period of retraction. When a developing culture is deprived of electrical activity for a certain amount of time, retraction does not take place, even if electrical activity is restored. Hence, in such a case supervision may have a different effect. Further simulation studies addressing these issues will elucidate the role of nonlinear epigenetic variance in supervised versus unsupervised learning.

We consider it possible that genetic influences and shared environmental events are causes of nonshared effects. For example, if shared environmental events have influence on a nonlinear system, they may act as nonshared influences and thus might have nonshared effects. Perhaps shared events have a greater chance to result in shared effects in MZ twins (pushing them towards to identical or similar attractors) whereas shared events have a greater chance to result in nonshared effects in DZ twins (pushing them towards nonidentical or different attractors). MZ correlations may

increase, while DZ correlations may increase less or may even decrease. This would further increase heritability and nonshared environmentability, and decrease shared environmentability. In that case, shared environmental factors that do not differ between children growing up in the same family may explain why children growing up in the same family are different and why pairs of identical twins become different.

In sum, we demonstrated the emergence of nonlinear variance in a bifurcating developmental process. We interpreted this variance as epigenetic variance. The presence of nonlinear epigenetic variance had clear effects on the twin correlations, which in turn affected the estimated heritabilities and environmentabilities. These were found to be comparable to those commonly reported in the literature. In conclusion, we believe that it is important to be aware of the role of nonlinear (epigenetic) processes in psychological development. Theoretically, these provide a challenging perspective on the sources of individual differences, and the interpretation of variance components. We believe that this perspective will ultimately enrich behavior genetic studies, by raising new research questions, and by reducing the gap between (linear) statistical modeling and developmental (nonlinear systems) theories of individual differences.

CHAPTER 3

THE NATURE OF NURTURE: ON THE HIGH HERITABILITY OF CULTURAL DEPENDENT COGNITIVE ABILITIES

Abstract

Heritabilitity coefficients differ across cognitive abilities, but it is not clear why. We propose that these differences are in part the result of differences in gene-environment correlation. Specifically, we hypothesize that during development such correlations become larger for societally valued, hence culturally influenced, knowledge and skills ('crystallized abilities') than cognitive processing abilities ('fluid abilities'). In support of this hypothesis, we present the results from 23 twin studies (total N = 7852). In adults, but not in children, heritability coefficient correlated with cultural load; cultural loaded crystallized tests showing higher heritability coefficients than culture reduced fluid tests. The contra-intuitive finding that more heritable cognitive abilities are more culturally dependent sheds a new light on the longstanding nature-nurture debate.

3.1 The Nature-Nurture Debate in Intelligence

Whether individual differences in intelligence are more due to nature than to nurture is a longstanding philosophical debate, dating back to seventeenth century rationalists, who favoured nature, and empiricists, who favoured nurture (Fancher, 1996). The emergence of psychometrics and behavioural genetics in the twentieth century has made possible an empirical approach to this issue (in terms of individual differences) through the decomposition of variance in (psychometric) intelligence into genetic and environmental components (e.g. Plomin et al, 2008). On first sight, the results seem to favour nature: Numerous studies have shown that the genetic components of full scale IQ and general intelligence (*g*) explain more than half (up to 80%) of the variance (Plomin et al.). However, it is well known that genotype-environment interaction and covariance affect heritability coefficients (Purcell, 2002). While such effects are receiving increasing attention, there are few definite results. This is due in part to the lack of success in identifying the actual genes underlying intelligence (Deary, Johnson, & Houlihan, 2009; Plomin and Spinath, 2004). Thus from this perspective, the nature and nurture of intelligence remains poorly understood. We aim to elucidate the issue of genotype-environment covariance by considering the differential heritabilities of specific cognitive abilities. Our results disconfirm predictions from the mainstream theories of intelligence, i.e., Cattell's fluid-crystallized theory (Cattell, 1963; 1987) and Spearman and Jensen's *g* theory (Spearman, 1904; 1927; Jensen 1998).

Both fluid-crystallized and *g* theory account for differential heritabilities of cognitive abilities by positing a major source of individual differences in IQ, i.e., Gf (fluid intelligence) in fluid-crystallized theory, and *g* (general intelligence) in *g* theory, which is related to heritable biological capacities of the brain. Individual differences in cognitive processes, as measured by IQ tests, depend on this source, and thus on its underlying genetic factors. The more complex a cognitive processing test is, the larger its dependency on this source (as indicated by its factor loading on Gf or *g*), hence on genetic variable factors (as indicated by its heritability coefficient, h^2). Individual differences in acquired knowledge are heritable as well, because they are in part the result of individual differences in cognitive processing. However, as the acquisition of knowledge is susceptible to environmental influences, heritability of individual differences in knowledge is lower than individual differences in the underlying cognitive processes.

On the basis of fluid-crystallized or *g* theory, researchers categorize IQ subtests as fluid versus crystallized (e.g. Cattell, 1963; Ackerman, 1996) or culture-reduced versus culture-loaded (e.g. Rushton & Jensen, 2010a; Georgas, van de Vijver, Weiss, & Saklofske, 2003). Fluid tests minimize

the role of individual differences in prior knowledge, whereas crystallized tests maximize it. Individual differences in fluid test scores reflect primarily individual differences in cognitive processes, such as reasoning and (working) memory. Individual differences in crystallized test scores reflect differences in knowledge and skills that are acquired during the lifespan, especially in educational settings. Because crystallized tests depend more strongly on (culture dependent) prior knowledge, they usually require more adjustments to adapt a test from one language group or culture to the next (Georgas et al.). In this sense, crystallized tests are typically more cultural loaded than fluid tests.

Fluid-crystallized and *g* theory predict relatively low heritabilities of cognitive undemanding, highly culture loaded, crystallized tests. Here, we formulate an alternative hypothesis based on the idea of dynamic gene-environment interplay (Dickens & Flynn, 2001; van der Maas et al., 2006; Dickens 2008). Consider the following. Heritability of IQ increases over the life span. Behaviour geneticists attribute this to increasing G-E covariance, which arises because more intelligent people are exposed to and seek out more cognitively stimulating environments (Scarr & McCartney, 1983; Dickens & Flynn, 2001). Assuming these environments foster societally valued knowledge and skills (i.e., crystallized abilities) rather than cognitive processing per se (fluid abilities) (Dickens & Flynn 2001; Dickens, 2008), we expect the increase in gene-environment covariance to be higher for crystallized abilities than fluid abilities. This ultimately results in higher heritabilities of crystallized abilities.

3.2 Empirical Data

3.2.1 Method

Consistent with Gf-Gc and *g* theories, several studies using the Wechsler scales (e.g. Wechsler, 1991, 1997) have shown that the subtests' *g* loadings and heritabilities correlate positively (Jensen, 1987; Rijsdijk et al., 2002, Pedersen et al., 1992). However, the degree to which these subtests were cognitive demanding, crystallized, or cultural loaded was not considered, nor were the effects of genotype-environment covariance. We re-examined these data in the light of the opposing predictions from Gf-Gc and *g* theories (crystallized abilities show lower heritability than fluid abilities) and the alternative hypothesis mentioned above (crystallized abilities show higher heritabilities than fluid abilities). Additionally, we conducted a comprehensive search of studies to locate all behaviour genetic studies that involved either Wechlser's Intelligence Scale for Children (WISC or its revisions) or Wechlser's Adult Intelligence Scale (WAIS or its revisions), and that contained sufficient information to compute subtests' heritabilities (h^2). These subtests comprise Vocabulary, Information, Arithmetic, Similarities, Verbal Comprehension, Digit Span, Picture Completion, Picture Arrangement, Block Design, Coding, and Object Assembly. Of these 11 subtests, the first 5 are often considered to be crystallized tests. We determined the degree to which these subtests are cultural loaded by considering alterations made to each subtest in adapting it to different cultures (languages). This degree was based on the average proportion of the number of items that had to be adjusted in each subtest of the WISC-III when it was adapted for use in thirteen different countries (Georgas et al., 2003). The 5 crystallized subtests were indeed the 5 most cultural loaded subtests (see Table 3.1). Since the Wechsler Verbal IQ (VIQ) and Performance IQ (PIQ) subscales also mapped well on cultural load (see Table 3.1), we also located studies that reported heritabilities of the Wechsler's VIQ and PIQ.

Below, we collate the results of 23 independent twin studies that used representative subject samples (combined N = 7852). Figure 3.1 shows the rank orders of the subtests according to their heritabilities. In all statistical analyses, we corrected mean heritabilities (see Table 3.2 and Table 3.3) and *g* loadings (see Table 3.4 and Table 3.5) for attenuation due to reliability (see Table 3.6 and 3.7).

Table 3.1, showing the cultural load of (both crystallized and fluid) cognitive abilities as measured by the subtests on the verbal intelligence scale (VIQ) and performance intelligence scale (PIQ) of the Wechsler Intelligence tests (e.g., Wechsler, 1991, 1997)

Subtest	Cultural load	Category	Scale
Vocabulary	0.35[1]	Crystallized	VIQ
Information	0.22[1]	Crystallized	VIQ
Comprehension	0.15[1]	Crystallized	VIQ
Similarities	0.09[1]	Crystallized	VIQ
Arithmetic	0.08[1]	Crystallized	VIQ
Picture Completion	0.03[1]	Fluid	PIQ
Picture Arrangement	0.02[1]	Fluid	PIQ
Block Design	0.01[1]	Fluid	PIQ
Coding	0.00[2]	Fluid	PIQ
Digit Span	0.00[1]	Fluid	VIQ
Object Assembly	0.00[1]	Fluid	PIQ

[1] see Table 18.1 in Georgas, van de Vijver, Weiss, & Saklofske (2003), [2] via e-mail correspondence with prof. dr. van de Vijver

3.2.2 Analyses

Wechsler Tests Analysis 1: Subtest scores

Our first analysis involved 6 child samples (WISC) and 4 adult samples (WAIS). We computed average heritabilities per subtests over the samples by weighting the heritabilities by the square root of the studies' sample size (twin pairs). In line with the finding that the heritability of IQ increases throughout development, a paired t-test showed that the WAIS subtests displayed higher heritabilities (mean = .56) than the WISC subtests (see Table 3.3) (mean = .43, t = 3.56, df = 10, p = .005). In line with the alternative hypothesis, in adults, but not in children, crystallized abilities were the most heritable (see Table 3.2): Two tailed tests showed that the WAIS heritabilities of crystallized tests (mean = .65) were higher than those of fluid tests (mean = .49) (t = -2.55, df = 8.38, p = .033) and that the correlation between heritabilities and cultural load was positive (r = .61, df = 9, p = .044; ρ = 0.60). In the WISC these effects were non-significant. Noteworthy, the most cultural loaded subtests had the highest *g* loadings (see Table 3.3) (WAIS: r = .82, p < 0.01; WISC: r = .81, p < 0.01). The correlations between *g* loading and heritability were not significant (see Table 3.8 and Table 3.9).

Wechsler tests Analysis 2: VIQ vs. PIQ scores

Our second analysis involved 12 samples (apart from the samples from the first analysis), of which 5 comprised children (WISC) and 7 comprised adults (WAIS). Corroborating the findings above, a one-sided paired Wilcoxon test showed that the WAIS, but not the WISC, displayed higher heritability for VIQ than PIQ (median of the WAIS VIQ-PIQ differences = 0.07, V = 19, p = 0.037; see Table 3.10 and Figure 3.2). Thus, the Wechsler data show that in adults, but not in children, highly cultural loaded, crystallized abilities are more heritable than cultural reduced, fluid abilities.

Minnesota Twin study

Our third analysis involved data from the Minnesota Twin study, in which 42 cognitive subtests from diverse batteries (see Table 3.11) were administered to 126 adult twin pairs raised apart. We computed intercorrelations among the subtests' *g* loadings, heritabilities, and two alternative operationalizations of cultural loading (see Table 3.12). We obtained a positive correlation between *g* loadings and heritabilities (r = 0.53, t = 3.93, df = 40, p < 0.001), but again the highest heritability coefficients and *g* loadings were clearly for the most cultural loaded subtests (see Figure 3.3).

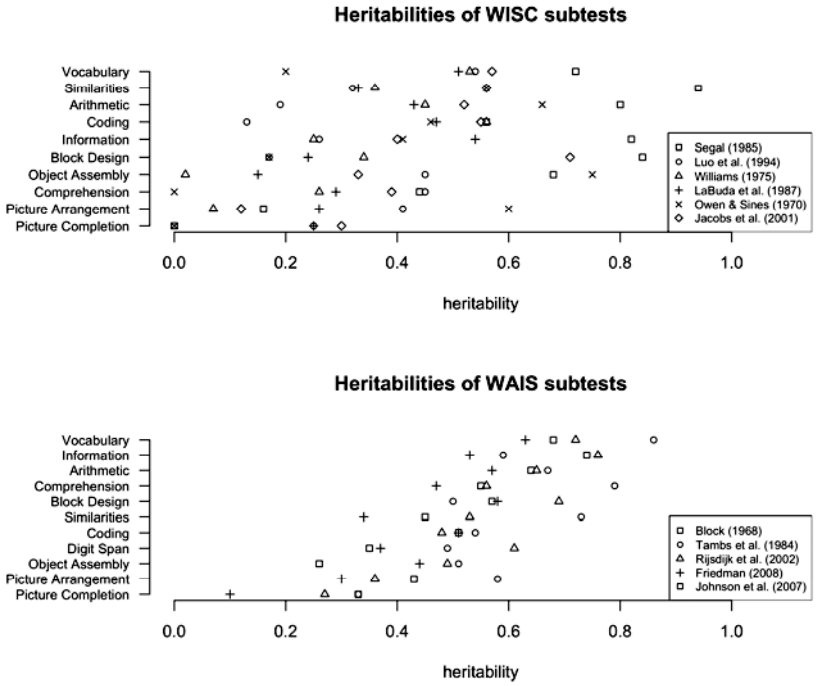


Figure 3.1 Heritabilities of the subscales of the Wechsler Intelligence Scale for Children (WISC; top) and the Wechsler Adult Intelligence Scale (WAIS; bottom)

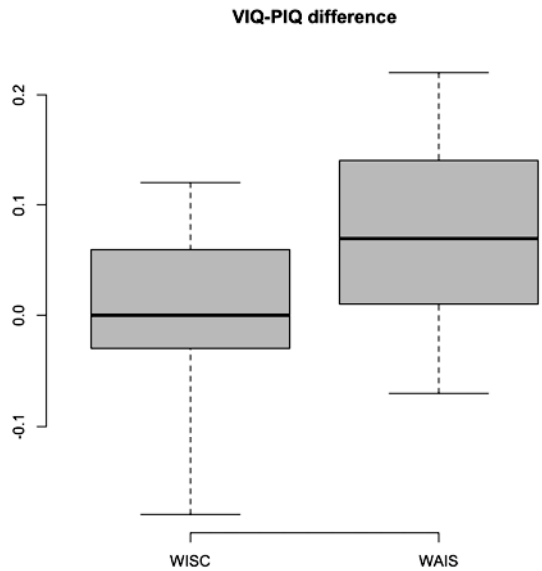


Figure 3.2 Boxplots showing that in the Wechsler adult scale, but not in de children's scale, Verbal IQ is more heritable than Performance IQ

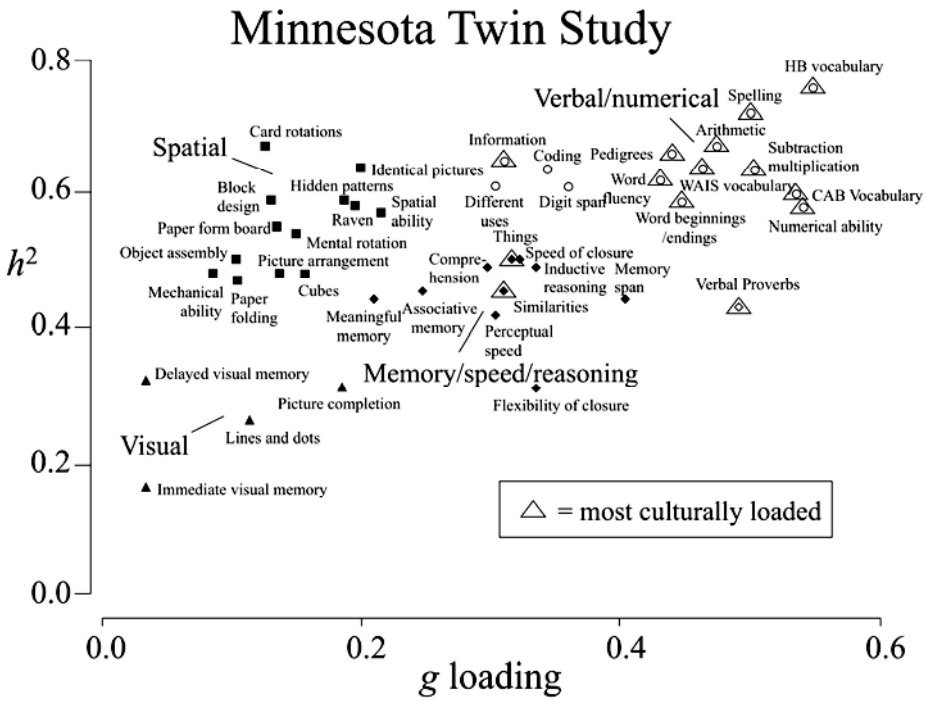


Figure 3.3 The relation between *g* loading, heritabilities (h^2) and cultural load in the Minnesota Twin Study, showing that culture loaded test are the most *g* loaded and the most heritable. An additional cluster analysis revealed that culture loaded verbal-scholastic tests were relatively highly heritable and *g* loaded. Complex, culture reduced spatial tests were also relatively highly heritable, but less *g* loaded. Culture reduced, memory, processing speed and reasoning tests tended to be moderately heritable, and varied substantially in *g* loading. Their *g* loadings were lower than those of the verbal-scholastic tests. Culture reduced visual processing tests were lowly heritable and *g* loaded.

3.3 Conclusions and Discussion

In sum, in contrast to the prediction from fluid-crystallized and *g* theory, we found in a wide range of samples that in adults, but not in children, highly culturally loaded tests displayed the largest heritability. They were also the most highly *g* loaded, which does not sit well with the notion that *g* loading reflects mainly cognitive demand. Our results, which support the differential increase in gene-environment correlation hypothesis, are hard to understand without considering the role education and experience. We hypothesize the following. Because the acquisition of knowledge depends on cognitive processing, individuals, who develop relatively high levels of cognitive processing, tend to achieve relatively high levels of knowledge. High achievers are more likely to end up in cognitive demanding environments that encourage and facilitate the further development of a wide range of knowledge and skills. These environments, e.g. higher educational systems, are largely determined by societal demands. These demands thus influence the magnitude of interactions among cognitive processes and knowledge, hence their intercorrelations, and so their *g* loadings. With respect to the nature-nurture debate, we conclude that the G-E covariance implies - contra-intuitively - that a high heritability of one cognitive ability relative to another implies that it depends relatively more on nurture.

Table 3.2 Heritabilities of the Wechsler's Intelligence Scale for Children (WISC) subtests

Subtest	Segal (1985) n = 103	Luo et al. (1994) n = 283	Williams (1975) n = 100	LaBuda et al. (1987) n = 143	Owen & Sines (1970) n = 42	Jacobs et al. (2001) n = 451	Weighted Average	Corrected for attenuation
Vocabulary	0.72	0.54	0.53	0.51	0.20	0.57	0.5511	0.5855
Information	0.82	0.26	0.25	0.54	0.41	0.40	0.4081	0.4421
Comprehension	0.44	0.45	0.26	0.29	0.00	0.39	0.3708	0.4209
Similarities	0.94	0.32	0.36	0.33	0.56	0.56	0.4872	0.5374
Arithmetic	0.80	0.19	0.45	0.43	0.66	0.52	0.4500	0.5056
Picture Completion	0.00	0.25	-0.12	0.25	0.00	0.30	0.2048	0.2331
Picture Arrangement	0.16	0.41	0.07	0.26	0.60	0.12	0.2282	0.2652
Block Design	0.84	0.17	0.34	0.24	0.17	0.71	0.4726	0.5114
Coding	0.56	0.13	0.56	0.47	0.46	0.55	0.4323	0.5179
Digit Span	-	0.15	0.13	0.44	0.16	0.26	0.2378	0.2765
Object Assembly	0.68	0.45	0.02	0.15	0.75	0.33	0.3575	0.4280

Table 3.3 Heritabilities of the Wechsler's Adult Intelligence Scale (WAIS) subtests

Subtest	Block (1968) n = 120	Tambs et al. (1984) n = 80	Rijsdijk et al. (2002) n =194	Friedman (2008) n = 293	Weighted Average	Corrected for attenuation
Vocabulary	0.68	0.86	0.72	0.63	0.6909	0.7117
Information	0.74	0.59	0.76	0.53	0.6386	0.6695
Comprehension	0.55	0.79	0.56	0.47	0.5467	0.6000
Similarities	0.45	0.73	0.53	0.34	0.4583	0.4971
Arithmetic	0.64	0.67	0.65	0.57	0.6165	0.6647
Picture Completion	0.33	0.33	0.27	0.10	0.2150	0.2388
Picture Arrangement	0.43	0.58	0.36	0.30	0.3723	0.4410
Block Design	0.57	0.50	0.69	0.58	0.6000	0.6442
Coding	0.51	0.54	0.48	0.51	0.5050	0.5554
Digit Span	0.35	0.49	0.61	0.37	0.4483	0.4920
Object Assembly	0.26	0.51	0.49	0.44	0.4308	0.5122

Table 3.4 Reliabilities of the Wechsler's Intelligence Scale for Children (WISC) subtests

Subtest	WISC US	WISC-R US	WISC-III US	WISC III-UK	WISC III-NL	Average
Vocabulary	0.86	0.86	0.86	0.89	0.96	0.8860
Information	0.76	0.85	0.85	0.85	0.95	0.8520
Comprehension	0.68	0.77	0.77	0.73	0.93	0.7760
Similarities	0.75	0.81	0.81	0.81	0.93	0.8220
Arithmetic	0.75	0.77	0.77	0.74	0.93	0.7920
Picture Completion	0.64	0.77	0.77	0.81	0.87	0.7720
Picture Arrangement	0.72	0.73	0.73	0.64	0.88	0.7400
Block Design	0.86	0.85	0.85	0.77	0.94	0.8540
Coding	0.60	0.72	-	0.77	-	0.6967
Digit Span	0.56	0.78	-	0.73	0.89	0.7400
Object Assembly	0.66	0.70	0.70	0.66	0.77	0.6980

Table 3.5 Reliabilities of the Wechsler's Adult Intelligence Scale (WAIS) subtests

Subtest	WAIS US	WAIS-R US	WAIS-III US	WAIS III-NL	Average
Vocabulary	0.95	0.96	0.93	0.93	0.9425
Information	0.91	0.89	0.92	0.92	0.9100
Comprehension	0.78	0.84	0.85	0.85	0.8300
Similarities	0.86	0.84	0.84	0.86	0.8500
Arithmetic	0.82	0.84	0.88	0.90	0.8600
Picture Completion	0.83	0.81	0.82	0.78	0.8100
Picture Arrangement	0.67	0.74	0.72	0.72	0.7125
Block Design	0.84	0.87	0.89	0.87	0.8675
Coding	-	0.82	0.82	0.84	0.8267
Digit Span	0.68	0.83	0.91	0.90	0.8300
Object Assembly	0.68	0.68	0.73	0.74	0.7075

Table 3.6 Loadings on the first principal factor ('*g* loadings') of the Wechsler's Intelligence Scale for Children (WISC) subtests

Subtest	WAIS US	WAIS-R US	WAIS-III US	WISC-III UK	WISC-III NL	Average	Corrected for attenuation
Vocabulary	0.81	0.82	0.82	0.78	0.78	0.8775	0.9039
Information	0.79	0.79	0.80	0.76	0.75	0.8500	0.8910
Comprehension	0.69	0.72	0.70	0.65	0.65	0.8100	0.8891
Similarities	0.72	0.79	0.80	0.78	0.75	0.8200	0.8894
Arithmetic	0.67	0.64	0.70	0.60	0.69	0.7400	0.7980
Picture Completion	0.49	0.58	0.60	0.50	0.49	0.6300	0.7000
Picture Arrangement	0.58	0.56	0.52	0.42	0.50	0.6325	0.7493
Block Design	0.59	0.67	0.66	0.54	0.57	0.6700	0.7193
Coding	0.44	0.40	0.33	0.26	0.34	0.5500	0.6049
Digit Span	0.51	0.45	0.45	0.42	0.43	0.5775	0.6339
Object Assembly	0.47	0.56	0.58	0.40	0.47	0.5500	0.6539

Table 3.7 Loadings on the first principal factor ('*g* loadings') of the Wechsler's Adult Intelligence Scale (WAIS) subtests

Subtest	WAIS US	WAIS-R US	WAIS-III US	WAIS-III NL	Average	Corrected for attenuation
Vocabulary	0.88	0.88	0.88	0.87	0.8775	0.9039
Information	0.86	0.84	0.84	0.86	0.8500	0.8910
Comprehension	0.80	0.80	0.82	0.82	0.8100	0.8891
Similarities	0.79	0.80	0.84	0.85	0.8200	0.8894
Arithmetic	0.77	0.74	0.72	0.73	0.7400	0.7980
Picture Completion	0.73	0.67	0.60	0.52	0.6300	0.7000
Picture Arrangement	0.69	0.62	0.64	0.58	0.6325	0.7493
Block Design	0.73	0.67	0.64	0.64	0.6700	0.7193
Coding	0.58	0.58	0.51	0.53	0.5500	0.6049
Digit Span	0.63	0.60	0.52	0.56	0.5775	0.6339
Object Assembly	0.52	0.55	0.58	0.55	0.5500	0.6539

Table 3.8 Correlations among heritabilities, *g* loadings and cultural load in the WISC

	Heritability	Corrected Heritability	*g* loading	Corrected *g* loading	Cultural load
Heritability	1.00				
Corrected heritability	0.99*	1.00			
g loading	0.57	0.48	1.00		
Corrected *g* loading	0.54	0.45	0.99*	1.00	
Cultural load	0.55	0.47	0.81*	0.79*	1.00

* significant $\alpha = .05$

Table 3.9 Correlations among heritabilities, *g* loadings and cultural load in the WAIS

	Heritability	Corrected Heritability	*g* loading	Corrected *g* loading	Cultural load
Heritability	1.00				
Corrected heritability	0.99*	1.00			
g loading	0.60	0.53	1.00		
Corrected *g* loading	0.50	0.46	0.98*	1.00	
Cultural load	0.63*	0.57	0.87*	0.82*	1.00

* significant $\alpha = .05$

Table 3.10 Heritabilities of the Wechsler VIQ and PIQ scales

Study	Test	n	VIQ	PIQ	VIQ-PIQ difference
Alarcon et al. (2000)	WISC	1110	0.74	0.68	0.06
Edmonds et al. (2008)	WISC	240	0.85	0.73	0.12
Gosso (unpublished)	WISC	359	0.7	0.73	-0.03
Hoekstra et al (2007)	WISC	115*	0.46	0.64	-0.18
van der Sluis et al. (2008)	WISC	385	0.78	0.78	0,00
Betjemann et al. (2010)	WAIS	142	0.66	0.66	0,00
Gosso (unpublished)	WAIS	580	0.78	0.71	0.07
Hoekstra et al (2007)	WAIS	115*	0.84	0.74	0.10
Malykh et al. (2005)	WAIS	160	0.86	0.84	0.02
Rijsdijk et al. (1998)	WAIS	416	0.64	0.46	0.18
Vandenberg (1968)	WAIS	-	0.63	0.41	0.22
van der Sluis et al. (2008)	WAIS	360	0.68	0.75	-0.07

* longitudinal study

Table 3.11 Characteristics of the Comprehensive Ability Battery (CAB), Hawaii Battery (HB) and Wechsler Adults Intelligence Scale Revised (WAIS-R) subtests used in the Minnesota Twin Study

Subtest	Battery	Name	Origin	Cultural load 1	Cultural load 2	Reliability	Kind	*g* loading	Corrected *g* loading	Heritability	Corrected Heritability
1	CAB	Numerical Ability	Hakstian & Cattell (1978)	0	1	0.79	test-retest	0.5418	0.6096	0.58	0.6526
2	CAB	Spatial Ability	Hakstian & Cattell (1978)	0	0	0.86	test-retest	0.2160	0.2329	0.57	0.6146
3	CAB	Memory Span	Hakstian & Cattell (1978)	0	0	0.96	split-half	0.4049	0.4132	0.44	0.4491
4	CAB	Flexibility of Closure	Hakstian & Cattell (1978)	0	0	0.79	split-half	0.3364	0.3785	0.31	0.3488
5	CAB	Mechanical Ability	Hakstian & Cattell (1978)	0	0	0.72	split-half	0.0863	0.1018	0.48	0.5657
6	CAB	Speed of Closure	Hakstian & Cattell (1978)	0	0	0.71	test-retest	0.3257	0.3865	0.50	0.5934
7	CAB	Perceptual Speed	Hakstian & Cattell (1978)	0	0	0.64	test-retest	0.3058	0.3823	0.42	0.5250
8	CAB	Word Fluency	Hakstian & Cattell (1978)	0	1	0.78	separately timed parts	0.4640	0.5254	0.64	0.7247
9	CAB	Inductive Reasoning	Hakstian & Cattell (1978)	0	0	0.74	split-half	0.3348	0.3892	0.49	0.5696
10	CAB	Associative Memory	Hakstian & Cattell (1978)	0	0	0.79	split-half	0.2470	0.2779	0.46	0.5175
11	CAB	Meaningful Memory	Hakstian & Cattell (1978)	0	0	0.84	split-half	0.2127	0.2321	0.44	0.4801
12	CAB	Verbal-Vocabulary	Hakstian & Cattell (1978)	1	1	0.78	split-half	0.5366	0.6076	0.60	0.6794
13	CAB	Verbal-Proverbs	Hakstian & Cattell (1978)	1	1	0.78	split-half	0.4930	0.5582	0.43	0.4869
14	CAB	Spelling	Hakstian & Cattell (1978)	1	1	0.78	split-half	0.5006	0.5669	0.72	0.8152
15	HB	Card Rotations	Wilson et. al (1975)	0	0	0.88	internal consistency	0.1267	0.1351	0.67	0.7142
16	HB	Mental Rotation	Wilson et. al (1975)	0	0	0.88	internal consistency	0.1501	0.1600	0.54	0.5756
17	HB	Paper Form Board	Wilson et. al (1975)	0	0	0.84	internal consistency	0.1353	0.1476	0.55	0.6001
18	HB	Hidden Patterns	Wilson et. al (1975)	0	0	0.92	internal consistency	0.1875	0.1955	0.59	0.6151
19	HB	Cubes	Ekström et al. (1976)	0	0	0.84	internal consistency	0.1580	0.1724	0.48	0.5237
20	HB	Paper Folding	Ekström et al. (1976)	0	0	0.84	internal consistency	0.1055	0.1151	0.47	0.5128
21	HB	Raven	Wilson et. al (1975)	0	0	0.86	internal consistency	0.1952	0.2104	0.58	0.6254

Table 3.11 (Continued)

22	HB	HB Vocabulary	Wilson et. al (1975)	1	1	0.96	internal consistency	0.5491	0.5604	0.76	0.7757
23	HB	Subtraction/ Multiplication	Wilson et. al (1975)	0	1	0.96	internal consistency	0.5035	0.5139	0.64	0.6532
24	HB	Word Beginnings/ Endings	Wilson et. al (1975)	0	1	0.71	internal consistency	0.4471	0.5306	0.59	0.7002
25	HB	Pedigrees	Wilson et. al (1975)	1	1	0.72	internal consistency	0.4394	0.5179	0.66	0.7778
26	HB	Things Categories	Wilson et. al (1975)	0	1	0.74	internal consistency	0.3174	0.3689	0.51	0.5929
27	HB	Different Uses	Ekström et al. (1976)	1	1	0.76	internal consistency	0.3031	0.3477	0.61	0.6997
28	HB	Immediate Visual Memory	Wilson et. al (1975)	0	0	0.58	internal consistency	0.0335	0.0439	0.16	0.2101
29	HB	Delayed Visual Memory	Wilson et. al (1975)	0	0	0.62	internal consistency	0.0335	0.0425	0.32	0.4064
30	HB	Lines and Dots	Wilson et. al (1975)	0	0	0.89	internal consistency	0.1133	0.1201	0.26	0.2756
31	HB	Identical Pictures	Ekström et al. (1976)	0	0	0.87	internal consistency	0.2008	0.2153	0.64	0.6862
32	WAIS-R	Information	Wechsler (1981)	1	1	0.89	split-half	0.3120	0.3307	0.65	0.6890
33	WAIS-R	Comprehension	Wechsler (1981)	1	1	0.84	split-half	0.2960	0.3230	0.49	0.5346
34	WAIS-R	Vocabulary	Wechsler (1981)	1	1	0.96	split-half	0.4304	0.4393	0.62	0.6328
35	WAIS-R	Coding	Wechsler (1981)	0	0	0.82	split-half	0.3451	0.3810	0.64	0.7068
36	WAIS-R	Arithmetic	Wechsler (1981)	1	1	0.84	split-half	0.4754	0.5188	0.67	0.7310
37	WAIS-R	Similarities	Wechsler (1981)	1	1	0.84	split-half	0.3120	0.3404	0.45	0.4910
38	WAIS-R	Digit Span	Wechsler (1981)	0	0	0.83	test-retest	0.3615	0.3969	0.61	0.6696
39	WAIS-R	Picture Completion	Wechsler (1981)	0	0	0.81	split-half	0.1864	0.2071	0.31	0.3444
40	WAIS-R	Block Design	Wechsler (1981)	0	0	0.87	split-half	0.1311	0.1406	0.59	0.6325
41	WAIS-R	Picture Arrangement	Wechsler (1981)	0	0	0.74	split-half	0.1384	0.1608	0.48	0.5580
42	WAIS-R	Object Assembly	Wechsler (1981)	0	0	0.68	split-half	0.1039	0.1260	0.50	0.6063

Table 3.12 Correlations among heritabilities, *g* loadings and cultural load in the Minnesota Twin Study

	Heritability	Corrected Heritability	*g* loading	Corrected *g* loading	Cultural load 1	Cultural load 2
Heritability	1.00					
Corrected heritability	0.98*	1.00				
g loading	0.52*	0.52*	1.00			
Corrected *g* loading	0.50*	0.52*	0.99*	1.00		
Cultural load 1	0.37*	0.36*	0.53*	0.52*	1.00	
Cultural load 2	0.46*	0.48*	0.76*	0.75*	0.76*	1.00

* significant at $\alpha = .05$

CHAPTER 4

A GENETIC ORIGIN OF BLACK-WHITE MEAN IQ DIFFERENCES? WEAK INFERENCES BASED ON AMBIGUOUS RESULTS

Abstract

Recently, Intelligence published an editorial with the aim to clarify the relation between subtests' g *loadings, heritability coefficients, and Black-White mean IQ differences. The following conclusions were drawn: (1) There is no strictly nongenetic (e.g., environmental or cultural) explanation for positive correlations among* g *loadings, heritability coefficients, inbred-outbred differences, and mean Black-White differences; and (2) These positive correlations support the 'biological* g *theory'. We demonstrate that the reasoning underpinning the first conclusion is invalid. The correlations have no bearing on the relative contributions of genetic and environmental effects on the mean Black-White differences. In addition, we scrutinize Black-White IQ data in the light of biological* g *theory. This theory predicted that the mean differences are larger for culture reduced tests than culture loaded tests. The data show the opposite*

4.1 A Genetic Origin of Black-White Mean IQ Differences?

Group differences in intelligence, racial differences in particular, are the topic of extensive and often heated debate. In this debate, political opinions and emotions often overrule scientific arguments. However, scientific research into these differences is legitimate, and empirical results can lead to the conclusion that racial differences in mean IQ, are, like individual differences, to some extent due to genetic differences. One complication within this research is that mean group differences can be entirely of environmental origin, while individual differences can be highly heritable (Lewontin, 1970; Rushton & Jensen, 2005). To date, the relative contributions of genetic and environmental influences to phenotypic IQ group differences remain generally unknown. In a recent editorial in *Intelligence*, however, Rushton and Jensen (2010a) maintained that certain empirical results help to reveal the origin of group differences. Such results are thus potentially important in the scientific debate on the origins of racial differences.

With their editorial Rushton and Jensen (2010a, henceforth R&J) set out to "clarify the relation between *g* loadings, heritabilities, Black-White differences, and the secular rise in IQ." (p. 214). In effect, they reviewed intercorrelations among these and a number of other variables, including the correlation between Black-White mean IQ differences and inbreeding effects (see R&J, p. 216, Table 1). From these correlations R&J drew various conclusions concerning the origin (environmental or genetic) of IQ group differences in general, and of Black-White IQ mean differences and secular gains in particular. In the light of these conclusions, R&J discussed implications for competing theories and explanations of Black-White differences. Although we are willing to entertain the hypothesis that mean Black-White differences (or indeed any phenotypic mean differences in valid measures) are partly genetic in origin, we demonstrate that the reasoning underpinning some of R&J's conclusions is invalid. Additionally, we demonstrate, by means of a re-analyses of R&J's data, that Black-White mean differences are inconsistent with the theory of intelligence that R&J favor.

The conclusions in the R&J's editorial can be summarized as follows: (1) the general factor of intelligence represents a real - that is, biological - largely genetically influenced variable (*g*), (2) the omnipresent secular gains in IQ are of environmental origin, and, (3) do not reflect differences (gains) on this biological variable, while (4) mean Black-White differences in IQ do reflect differences on this biological variable and (5) are, in part, of genetic origin. Conclusion 1 was based on the finding that IQ subtests' *g* loadings correlate positively with their heritability coefficients (h^2).

Conclusion 2 was based on the finding that the subtests secular gains are zero or negatively correlated with their heritability coefficients. Conclusion 3 was based on the finding that these gains are zero or negatively correlated with *g* loadings. Conclusions 4 and 5 were based on the finding that Blacks score on average lower than Whites, and the mean differences are largest on the most *g* loaded subtests (see, e.g. Jensen 1885; 1987), which are also the most heritable subtests, and on the finding that the (standardized) mean Black-White differences correlate positively with the heritability coefficients themselves.

R&J considered results pertaining to samples that are characterized by a degree of inbreeding depression to gather support for their conclusion that the Black-White mean differences are partly of genetic origin. They offered the following rationale (see R&J, p. 214). First, inbreeding depression is 'a purely genetic effect'. Second, because inbreeding depressed samples score systematically lower on IQ tests than outbred samples, and the subtests' *g* loadings and heritabilities correlate positively with the mean differences, inbreeding depression influences IQ via *g*. Third, the mean Black-White differences correlate positively with inbred-outbred differences. The upshot was that 'there is no non-genetic explanation' (R&J, p. 214) for the finding that *g* loadings, heritabilities, inbreeding effects, and mean Black-White differences intercorrelate positively.

One of the purposes of R&J was to provide empirical evidence that, in contrast to the subtests' secular gains, which are generally considered to be of environmental origin (but see Mingroni, 2007), Black-White mean differences show a pattern similar to the inbred-outbred mean differences (considered to be a purely genetic effect), and are, at least in part, of genetic origin. That is, R&J reviewed evidence that mean Black-White differences are the largest on the most heritable, most *g* loaded subtests. This purpose appears to make sense in light of their hypothesis that "if population group differences are greater on the more *g* loaded and more heritable subtests, it implies they have a genetic origin." (R&J, p. 214). From this hypothesis, which we will denote the genetic origin hypothesis, R&J inferred the following (R&J, p. 214):

> "(1) Genetic theory predicts a positive association between heritability and group differences; (2) culture theory predicts a positive association between environmentality and group differences; (3) nature+nurture models predict both genetic and environmental contributions to group differences; while (4) culture-only theories predict a zero relationship between heritability and group differences." (Rushton & Jensen, 2010a, p. 214)

Our present aims are threefold. First, we show that one cannot infer that these group differences are genetic because these differences are greater on the more *g* loaded and more heritable subtests. Secondly, we demonstrate that the predictions of the competing theories, as conveyed by R&J, are open to debate. Specifically, contrary to R&J's inference concerning culture theory and culture only theory, these theories can readily accommodate a positive correlation between heritability and *g* loading. Thirdly, in the light of Rushton and Jensen's biological *g* theory (Rushton & Jensen, 2005; 2010a, 2010b), we reexamine the original Black-White data in Jensen (1985, 1987), to which R&J refer. We show that one important association remains to be clarified, namely the relation between the subtests' Black-White mean differences and their cultural loading. We conclude that 1) the relative contributions of genes and environment to Black-White differences remain unknown and 2) biological *g* theory does not give a sufficient explanation of the relation between cultural loading and Black-White differences.

4.2 The Reasoning Underlying the Genetic Origin Hypothesis is Weak

We first demonstrate that the reasoning underlying the genetic origin hypothesis is open to criticism because it involves an *affirmation of the consequence*. Consider the (standardized) regression model in Figure 4.1, in which two groups are compared with respect to IQ. Within each group, the IQ subtest scores are regressed on the (latent, unmeasured) variable *g* (with regression weights λ_i); *g*, in turn, is regressed on latent (unmeasured) genetic and environmental influences (with regression weights *h* and *e*, respectively). We introduce a mean group difference on the genetic component of *g*. The effect is purely genetic, as we assume that the means of the environmental components over the groups are

equal. The genetic mean difference results in a group mean difference in *g*. Because there is a mean difference in *g*, the IQ mean differences are the most pronounced on the subtests with the largest the *g* loadings (i.e., the subtests' with the largest regression weights on *g*). The *g* loadings (λ_i) correlate perfectly with the subtests' regression weights on *g*'s genetic component, which take the value of $h\lambda_i$. Because in this model, *g* is the only source of the subtests genetic variance, it follows that the subtests' heritability coefficients take the value of the squared regression weights on *g*'s genetic component: $(h\lambda_i)^2$. The relation between the subtests' *g* loadings (λ_i) and their heritability coefficients ($h^2\lambda_i^2$) is thus also perfect (although quadratic). Because both the IQ mean differences and the regression weights on *g*'s genetic component ($h\lambda_i$) are collinear with the *g* loadings (λ_i), the (rank) correlation between the subtests' IQ mean differences and their heritability coefficients ($h^2\lambda_i^2$) is +1.

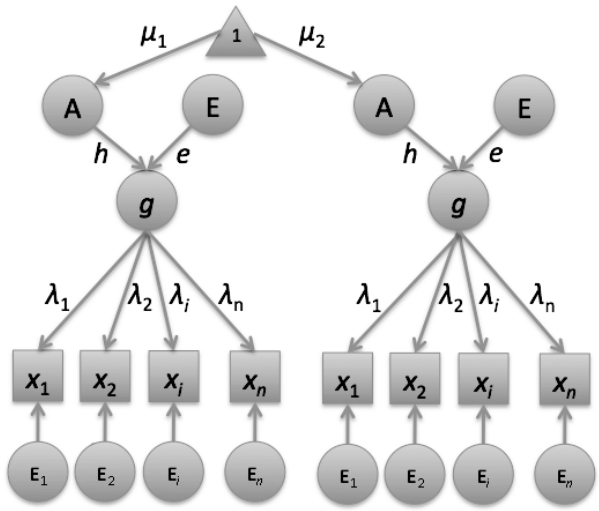


Figure 4.1 A mean difference on the genetic component of *g* will result in a mean difference on *g*. A mean difference in *g*, in turn, will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$), because both correlate perfectly with the ranks of the subtests' *g* loadings.

Thus, due to a purely genetic origin, the population group differences are greater on the more *g* loaded and more heritable subtests. We let the two groups represent an inbreeding depressed sample and an outbred white sample (we assume that the inbreeding depressed sample has a lower average level of *g*: $\eta_1 < \eta_2$).

Next, we compare the normal white sample with a normal black sample (Figure 4.2). We introduce a mean difference on *g* (but not on the more specific factors of intelligence), such that blacks have a lower average than whites ($\eta_1 < \eta_2$). Like the inbred-outbred IQ mean differences, the Black-White IQ mean differences will be the most pronounced on the most *g* loaded tests (those with the highest λ_i). So, these differences will correlate perfectly with the inbred-outbred mean differences. Because *g* loadings (λ_i) are collinear with the regression weights on *g*'s genetic component ($h\lambda_i$), the Black-White gaps are also collinear with them. As a corollary, the rank correlation between Black-White gaps and heritability coefficients ($h^2\lambda_i^2$) is +1. In short, the intercorrelations among inbred-outbred mean differences, Black-White mean differences, heritability coefficients, and *g* loadings are positive and perfect.

In view of the genetic origin hypothesis, the crucial question is: Does this finding *imply* that the Black-White differences has a genetic origin? Is there 'no non-genetic explanation' for the finding that *g* loadings, heritabilities, inbreeding effects and mean Black-White differences intercorrelate positively?

We agree that it is *possible* that the mean Black-White difference in *g* is due in part to a mean difference on *g*'s genetic component (as in Figure 4.1). If so, in studying Black-White differences, we would expect find the pattern of correlations among *g* loadings, heritabilities, and mean group differences similar to those found in inbred-outbred studies. However, inferring an underlying cause from this pattern of correlations is problematic in view of the following.

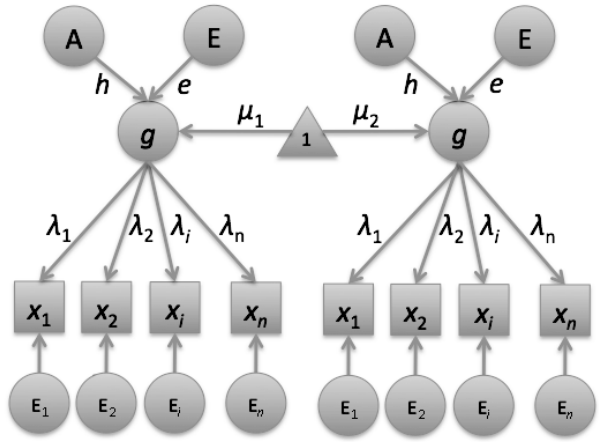


Figure 4.2 A mean difference in *g* will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$). The source (genetic or environmental) of this mean difference is unknown.

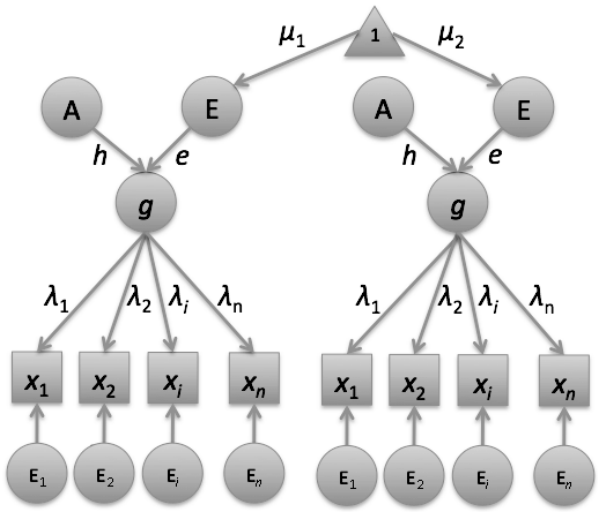


Figure 4.3 A mean difference on the environmental component of *g* will result in a mean difference on *g*. A mean difference in *g*, in turn, will result in mean differences in IQ. The ranks of the subtests' mean differences will correlate perfectly with the ranks of the subtests' heritability coefficients (which take the value of $h^2\lambda_i^2$), because both correlate perfectly with the ranks of the subtests' *g* loadings (λ_i).

Consider the situation depicted in Figure 4.3. That is, suppose that the true situation was that Blacks and Whites differed on *g*'s *environmental* component. As a result, there would be a mean difference on *g*. Again, the mean IQ differences would be greatest on the most *g* loaded tests (i.e.,

those with the highest λ_i), which happen to be most heritable tests (as shown above, the subtests' heritability coefficients will equal $h^2\lambda_i^2$). The Black-White mean IQ differences would correlate perfectly with the inbred-outbred mean IQ differences, since they are both collinear with *g* loading (λ_i). Thus, in this situation, in which the Black-White differences were purely of environmental origin, we would have found the same pattern of correlations among Black-White IQ differences, inbred-outbred IQ differences, *g* loadings, and heritability coefficients as in the previous situation, in which the origin of the group differences was genetic. So, in this *g* model, any relative contribution of genes and environment to group differences could lead to the finding that Black-White differences, inbred-outbred differences, *g* loadings, and heritabilities intercorrelate positively. Correlations among these variables have no necessary bearing on the origin of group differences, and do not furnish a basis for inferences, strong or otherwise.

4.3 Inferences are Invalid

We now examine R&J's (p. 214) inferences concerning 'genetic theory' and 'culture (only) theory'. As explained in Rushton & Jensen (2005), genetic theory "contends that a substantial part (say 50%) of both individual and group differences in human behavioral traits is genetic" (Rushton & Jensen, 2005, p. 238), while culture only theory "finds no need to posit any genetic causation" (Rushton & Jensen, 2005, p. 238) of group differences:

> "*The defining difference* [between genetic theory and culture-only theory] *is whether any significant part of the mean Black-White IQ difference is genetic rather than purely cultural or environmental in origin*." (Rushton & Jensen, 2005, p. 238, italics original)

In the light of this distinction between 'genetic theory' and 'culture-only theory, consider the first inference: "Genetic theory predicts a positive association between heritability and group differences." (R&J, p. 214). On the one hand, it is true that genetic theory can predict this association, but under certain circumstances. It is the case, for example, if *g* exists and if *g* is the sole (or main) source of subtests genetic variance, and if there is a group difference on *g*'s genetic component. On the other hand, as demonstrated above, a (possibly perfect) positive correlation between heritability and group differences is quite compatible with purely environmental origins of the phenotypic mean Black-White differences. That 'culture-only theories predict a zero relationship between heritability and group differences' is thus not necessarily true. In addition, in more complex models than the simple 1-factor model with the common factor as the only source of subtests' genetic variance, mean differences can be entirely genetic without a positive association between subtests' heritabilities and mean differences. One can only make unambiguous predictions about the correlation between heritabilities and mean group differences, if one explicitly models the causal genetic and environmental pathways, for example in a Structural Equation Model.

The third inference that "culture theory predicts a positive association between environmentality and group differences" (R&J, p.214) is not also necessarily true. In Figure 4.3, for example, the origin of mean group differences is purely environmental, but the correlation between group mean differences and environmentalities (1- $h^2\lambda_i^2$) is -1. Analogous to the argument above, one cannot make predictions about the correlation between environmentalities and mean group differences without explicitly modeling the causal pathways.

We conclude that nothing about genetic and culture-only theories can be *inferred* from the vector correlations. Thus, in order to establish with any certainty the pathways through which genetic and environmental latent variables affect the factors of intelligence, these variables and pathways must be modeled explicitly. This is not sufficient to establish the contributions of environmental and genetic sources of the mean group differences, however, because in the latent *g* model these contributions are unidentified. To establish these, environmental and genetic variables should have been measured and incorporated in the model, replacing in part or entirely the latent variables. With respect to genetic influences, this requires the identification of genes that explain (preferably in a causal sense) individual differences in intelligence. Once identified, their actual

contribution to between group, i.e., Black-White, differences can be estimated. The same applies to environmental (causal) influences.

Besides the fact that the inferences are not valid, R&J's explanation of Black and White IQ mean differences is problematic. Specifically, as we discuss in the next section, the data are not consistent with respect to R&J's position on the role of cultural loading.

4.4 Problems for Biological *g* Theory

When comparing competing theories of intelligence or explanations of Black-White differences, one needs falsifiable predictions. Here, we argue that in addition to the problem with Rushton and Jensen's inferences (as identified above), predictions stemming from their biological *g* theory are not consistent with certain empirical results concerning Black-White mean differences. That is, if we accept - for the sake of argument - Rushton and Jensen's method (essentially Jensen's method of correlated vectors; Jensen, 1998, pp. 372-374), and assuming that Black-White mean differences are in indeed due to the biological variable *g* and partly of genetic origin, we are confronted by puzzling empirical results concerning the role of cultural loading in the Black-White differences.

In biological *g* theory, the variable *g* is viewed as highly heritable and largely fixed, i.e., not readily malleable by cultural influences. Here cultural influences are contrasted with genetic influences, and correspondingly cultural loadings are contrasted with *g* loadings (see Rushton & Jensen, 2005, 2010b; Rushton, 1998). Culture reduced test are hypothesized to have higher *g* loadings and higher heritability coefficients than culture loaded tests. Hence, the hypothesis that Black and White group differences are more pronounced on the more *g* loaded tests, implies that in biological *g* theory the Black-White mean differences are relatively small on strongly cultural loaded tests ('culture loaded' tests) and relatively large on weakly cultural loaded tests ('culture reduced' tests). Rushton & Jensen (2005, p.272) expressed this as follows: "Mean Black-White differences [are] greater on *g* loaded cognitive tests than on culturally loaded cognitive tests".

Culture theory, as conveyed by R&J, on the other hand, "predicts that differences between races will be greater on those culturally malleable items on which races can grow apart as a result of dissimilar experiences." (Rushton, 1998, p. 222). We note that 'culture theory' differs from 'culture only theory' (Rushton & Jensen, 2005) in the sense that culture theory does not necessarily exclude the possibility that genetic differences play a role in the development of individual and group differences. Consider the theory of Dickens & Flynn, 2001), for example. In this theory intelligence is (highly) heritable, but also malleable by environmental (e.g. cultural) influences. These influences can increase initial (individual and mean group) differences, whether these are genetic or environmental. A culture theory can thus encompass a genetic theory (a theory that describes that a substantial part of both individual and group differences in intelligence is genetic), or a 'nature-nurture model' (R&J, p. 214).

Both biological *g* theory and Dickens' and Flynn's theory are nature-nurture models, but the former predicts that Black-White differences are greater on culture reduced tests than culture loaded test, while the latter predicts that these differences are greater on cultural reduced tests than culture loaded tests. In the following section, we first explain the concept of cultural loading in some more detail. Next, using data upon which R&J based their conclusions, we evaluate the relation between Black-White differences, *g* loading and cultural loading in the light of biological *g* theory.

4.4.1 Cultural loading

Jensen provided the following conceptualization of a test's cultural loading:

> "Tests and test items can be ordered along a continuum of culture loading, which is the specificity or generality of the informational content of the test items. The narrower or less general the culture in which the test's information could be acquired, the more cultural loaded it is. A test may contain information that could only be acquired within a particular culture. This can actually be determined simply by examination of the test items. The specificity or generality of the content corresponds to its cultural loading." (Jensen, 1976, p.340)

Culture loaded tests are thus the tests that occupy relatively high positions on the cultural loading continuum. At the other end of the continuum are located 'culture reduced' (or 'culture fair') items or tests. The theoretical distinction between culture loaded and culture reduced tests features quite often in the literature (e.g. Hunt & Sternberg, 2005; te Nijenhuis, Tolboom, Resing, & Bleichrodt, 2004; Helms-Lorenz & van de Vijver, 2003), and different operationalizations of cultural loading have been used (e.g. te Nijenhuis et al., 2004; Helms-Lorenz & van de Vijver, 2003). One operationalization involves establishing loading on the basis of ratings by experts or ratings by students, another involves the categorization of tests into 'aptitude tests' and 'achievement tests'. These operationalizations are consistent with each other: Knowledge, scholastic, and achievement test are viewed as most cultural loaded (see also Chapter 3).

In this connection, the distinction between crystallized and fluid abilities is also relevant. Specifically, we note that some researchers equate culture loaded tests with 'crystallized abilities' tests (e.g. Hunt & Sternberg, 2005). Indeed, crystallized abilities refer to knowledge and well-practiced skills, and hence are highly culturally influenced (Cattell, 1987). They are differentiated from 'fluid abilities', which concern reasoning abilities. Although the process of solving the items on fluid abilities tests always requires knowledge to some degree, hence is culturally influenced to some degree, this knowledge is considered to be equally known or equally new for the people for whom the test is valid. This implies that, in contrast to crystallized ability tests, prior knowledge is assumed not to *differentiate* between the members of the population of interest, and is therefore no source of individual differences.

The distinction between fluid and crystallized abilities tests shows a large overlap with the distinction between culture loaded versus culture reduced tests (see Chapter 3), but they do differ (Cattell, 1987; Jensen, 1998). Verbal *knowledge* tasks, for example, are crystallized ability tests, while tests of verbal *reasoning* are fluid ability tests. However, the assumption that the required knowledge in verbal reasoning is equally known is reasonably tenable only if the testees share the same linguistic and cultural background. The items of verbal reasoning tests have to be translated or adjusted before they can be administered to testees from other cultural backgrounds; otherwise the tests would be invalid. Both verbal (crystallized) knowledge and verbal (fluid) reasoning tests are culture loaded, but the latter less than the former.

Here, we prefer to retain the terms culture reduced and culture loaded, because the terms fluid and crystallized abilities are linked to a particular theory of intelligence (Cattell's investment theory of fluid and crystallized intelligence; e.g. Cattell, 1987), whereas the concept of cultural loading is more independent of theories of intelligence.

Henceforth, in our statistical analysis, we abide by Jensen's definition of culture reduced tests, and dichotomize cultural loading in the light of this definition: Culture reduced tests are "those that are nonlanguage and nonscholastic and do not call for any specific prior information for a plus-scored response" (Jensen, 1980, p. 374). Therefore, we consider tests that depend on language, that are scholastic or that require specific substantive prior information as culture loaded (Jensen, 1980, p.). Examples of culture loaded tests are the (WISC and WAIS) subtests Vocabulary, Information, and Similarities (verbal knowledge tests, calling for specific knowledge), Verbal Comprehension (a verbal reasoning test, depending on language) and Arithmetic (a scholastic test) (see also Chapter 3). Examples of tests that are considered to be culture reduced are Raven's Progressive Matrices (nonverbal, abstract reasoning), Block Design (mental manipulation in space), and elementary cognitive tasks (e.g. reaction time tasks) (see also Chapter 3).

4.4.2 Reanalyses of the Data: Black-White Differences are Largest on the Most Culturally Loaded Tests

Predictions Rushton & Jensen (2005; 2010b) implied that *g* loading and cultural loading are intrinsically inversely, i.e., negatively, related. As quoted above, on the basis of biological *g* theory, Rushton and Jensen hypothesized (and also reported, e.g. Jensen, 1985; Rushton, 1995) that mean Black-White differences in IQ are larger on *g* loaded than on culture loaded tests. One would expect the data to be in line with Figure 4.4.

We note that contrary to the biological *g* prediction, but compatible with Dickens and Flynn's theory (2001; Dickens, 2008), cultural loading and *g* loading may well be intrinsically *positively* related (see Chapter 3), for example when *g* loadings are a function of societal demands (Dickens, 2008). The finding that Black-White differences are largest on the most *g* loaded, most heritable tests, poses no problem for this theory.

We argue that empirical data are not in line with the predictions from biological *g* theory. To this end, we present the results of a re-analyses of data, to which R&J (Jensen, 1998, p. 378, 386) refer.

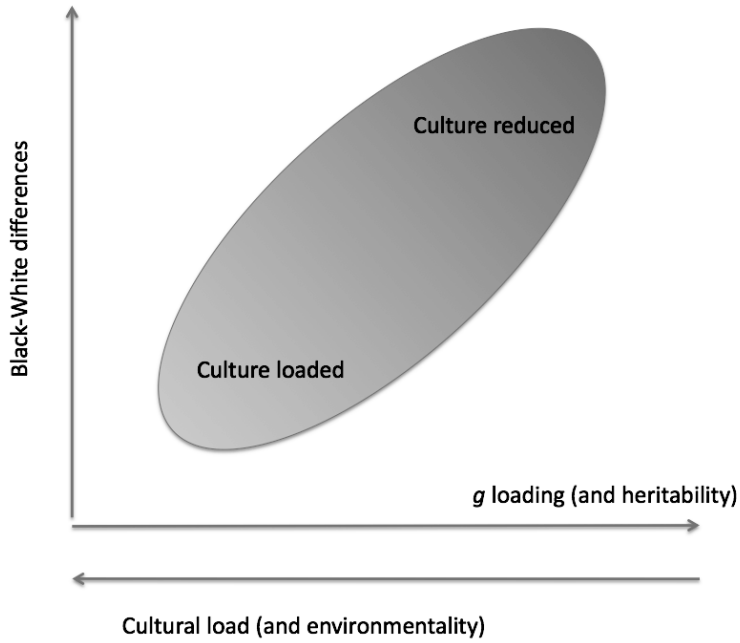


Figure 4.4 In biological *g* theory, *g* is viewed as highly heritable and not readily malleable by cultural influences. Cultural influences are contrasted with genetic influences; cultural loading with *g* loading. Culture reduced tests have higher *g* loadings and higher heritability coefficients than culture loaded tests. Black-White mean differences are relatively small on culture loaded tests and relatively large on culture reduced tests.

Data The data are summarized in a review of Black-White differences by Jensen (1985, 1987), who intended to show that the standardized mean differences are a function of *g* loading. They come from 11 data sets in Jensen's (1985) target article in Behavioral and Brain Sciences (1985), together with the additional single dataset from a later commentary on this article (Jensen, 1987). In these studies, the loadings on the first principal factor featured as the *g* loadings.

Participants The total set contained 145 Black and White standardized mean differences and *g* loadings. Study characteristics and data are provided in these papers. The subject samples were representative for the US Black and White populations (e.g. subjects participated in national studies or in standardization procedures). The total number of participants was 26464 (1560 whites and 10824 blacks).

Method The total number of unique subtests was 73 (see Table 4.1 for details). Within each study, we averaged subtests' *g* loadings in the black and the white samples (we weighted by sample size of blacks and whites). In most cases these were given separately. If not, we assumed *g* loadings were

equal in both groups. We considered this acceptable because the comparison of factor loadings revealed that they were similar (Jensen, 1985). For each unique subtest, we calculated the weighted means of *g* loading, reliability, and Black-White differences across studies (we weighted by the studies' total sample size).

Subtests' cultural load was determined by examining the subtests descriptions or the description of the factors on which they loaded. If the subtests measured scholastic abilities, achievement, crystallized intelligence, verbal comprehension, arithmetic, knowledge (or synonyms of these terms), we coded cultural load as 1, indicating that these tests are 'culture loaded' tests. The cultural load of the other tests was coded 0 to absence of appreciable cultural load. The proportion culture loaded test so defined was approximately 53%.

Results Visually inspecting the original data we noticed that within each battery the largest Black-White standardized mean differences were on culture loaded tests (most often involving vocabulary; see Jensen 1985, p.203). Also in the complete dataset (see the scatterplot in Figure 4.5), the Black-White mean differences appeared the largest on culture loaded tests.

The boxplot in Figure 4.6 (left pane) shows that the subtests' Black-White differences were indeed larger on culture loaded tests than on culture reduced tests. This is inconsistent with biological *g* theory. We also note that Rushton and Jensen's culture loaded *versus g* loaded distinction is problematic, because *g* loadings were the highest for culture loaded tests (see the boxplot in Figure 4.6, middle pane, see also Chapter 3). The intercorrelations between the subtests' cultural load, averaged *g* loadings, and averaged Black-White differences were all positive and significant at α = .01.

However, as Jensen (1987) mentioned, reliabilities of culture loaded tests were higher than reliabilities of culture reduced tests (see the boxplot in Figure 4.6, right pane), which have resulted in underestimation of factor loadings of culture reduced test. Below, we show that even if we accept, for the sake argument, R&J Jensen's method of correlated vectors (p. 372-374), we would arrive at the same conclusion: Black-White differences are largest on culture loaded tests.

The method prescribes that reliability should be taken into account out by adjusting *g* loadings and Black-White standardized mean differences using correction for attenuation. We considered biological *g* theory not to be in conflict with the data when if this is done, *g* loading correlates with mean Black-White differences, while *g* loading and cultural load do not correlate positively. We corrected *g* loadings and Black-White differences for attenuation. That is, we divided them by the square root of their reliabilities. Next, we calculated the intercorrelations among cultural load, the corrected *g* loadings and the corrected Black-White differences. These were all positive (see Table 4.2) and significant at a significant level of α = .01. Culture loaded tests still showed larger Black-White differences (mean = 0.78) than culture reduced tests (mean = 0.63) (t = -5.84, df = 57, p < 0.001).

Conclusion Black-White mean differences largest on the most culture loaded subtests whereas cultural load and *g* loading correlate positively. These findings are problematic for biological *g* theory, because in this theory cultural load and *g* loading are juxtaposed; it was predicted that Black-White mean differences are largest on cultural reduced subtests. We maintain that we need more adequate theory to explain these data. As the contributions of genes and environment to the Black-White mean IQ differences cannot be determined or inferred from the present data (nor methodology), we cannot state whether or not these differences are (partly) of genetic or (entirely) of environmental origin. That is, we cannot state whether genetic theories or culture-only theories are correct.

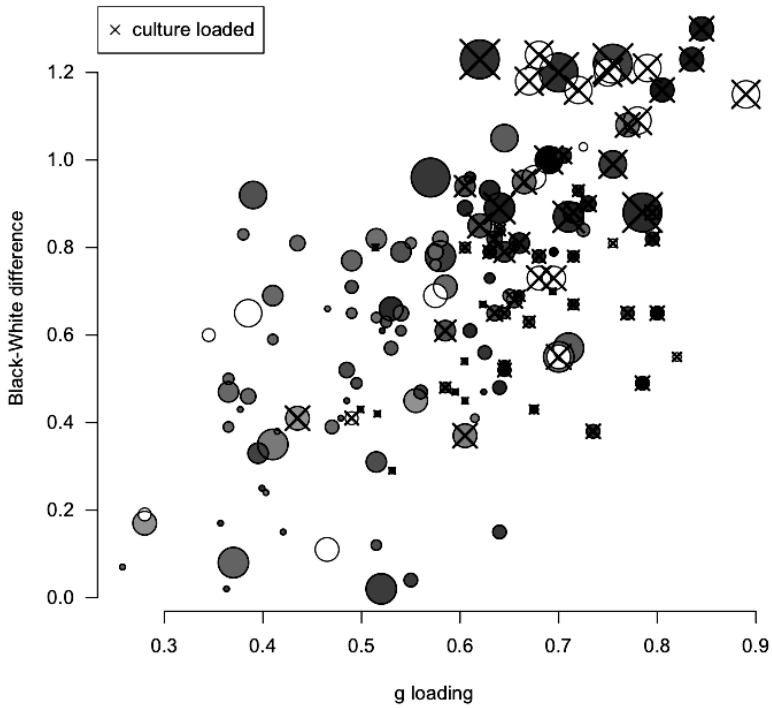


Figure 4.5 Scatterplot of Black-White Differences set out against loadings on the first principal factor for each study apart. The circles denote the data points; their size (area) represents the subject sample size; the darkness of the shade is proportional to the tests' reliabilities (circles are rendered in white if the reliability of the test is unknown).

Figure 4.6 Black-White differences are the largest on culture loaded subtests (left pane). Those tests are the most *g* loaded (middle pane) and the most reliable (right pane).

4.5 Discussion

We showed that Rushton & Jensen (2010a, p.214)'s hypothesis that "if population group differences are greater on the more *g* loaded and more heritable subtests, it implies they have a genetic origin" is invalid. Specifically we agree that a particular genetic origin of black-white differences would imply these correlations. However, we cannot infer a genetic origin from the correlations. This is because they are necessary also consistent with an environmental origin of black-white differences (hence with culture-only theory). The method on which Rushton & Jensen based their inferences is non-informative as applied in their editorial. The correlations have no necessary bearing on the origin of group differences.

Additionally, we reexamined Black-White data (Jensen, 1985; 1987) in the light of the theory that they favored, i.e., biological *g* theory (as we have denoted it). We provided evidence for an association that had not been discussed or clarified by the authors, namely the positive correlation between the subtests' Black-White mean differences and their cultural loading. Culture loaded test (language, knowledge, and achievement or scholastic tests) showed large *g* loadings and mean IQ differences, larger than culture reduced tests. The more cultural aspects of a test differentiate between people, the larger were the Black-White mean differences. This was inconsistent with biological *g* theory, because this theory predicted the opposite. Even if Black-White differences are indeed to some extend due to genetic differences, biological *g* theory does not explain why they are more pronounced on culture loaded tests than culture reduced tests.

Black-White differences were thus smaller on tests that measure cognitive processing than on tests that measure achievement ('on which races can grow apart as a result of dissimilar experiences', Rushton, 1998, p. 222). Achievement is certainly not fixed, but the result of past cognitive functioning and environmentally influenced learning experiences that take place throughout the course of development. In explaining the Black-White data adequately a developmental perspective is required; the use of cross-sectional data is insufficient. We advocate longitudinal designs and latent growth modeling, incorporating measured variables that replace in part or entirely the latent variables.

We entertain the possibility that *g* loadings are a function of environmental (culturally influenced) demands (Dickens & Flynn, 2001; Dickens 2008). Dickens simulated data using a model in which the more societal valued a (heritable) cognitive skill was, the more that skill was trained. He showed that, in this model, the most societal valued skills became the most *g* loaded and most heritable. In these simulations, interindividual differences were highly heritable, while intraindividual differences (growth) was large. The simulations provide an interesting perspective on individual and group differences. It implies that any initial environmental or genetic differences, large or small, can increase to large individual and group differences. This increase is due to a dynamical interplay between individuals' heritable characteristics and their environment. Such interplay will lead to a gene-environment correlation. We hypothesize that gene-environment correlation is higher for culture loaded ('crystallized') tests than culture reduced tests. Because gene-environment correlation can be incorporated in behavior genetic models (Purcell, 2002), this hypothesis is amenable to empirical investigation.

We conclude that the actual contributions of genes and environment to group differences cannot be determined or inferred from the method of correlated vectors. To date, whether and how genetic and environmental influences contribute to the Black-White differences remains unknown. This will remain unknown until the genetic and environmental variables that influence intelligence have been identified and the Black-White differences are regressed on them - preferably in a causal sense.

Table 4.1 Subtests' characteristics used in the statistical analyses

			Characteristic					
Test Number	Battery	Test name	Cultural load	*g* loading	Black-White Difference	Reliability	*g* loading*	Black-White Difference*
1	WISC-R	Information	1	.652	0.671	.778	.740	0.760
2	WISC-R	Similarities	1	.645	0.769	.810	.716	0.855
3	WISC-R	Arithmetic	1	.609	0.609	.770	.694	0.694
4	WISC-R	Vocabulary	1	.681	0.840	.826	.749	0.924
5	WISC-R	Comprehension	1	.547	0.648	.702	.653	0.773
6	WISC-R	Digit Span	0	.507	0.394	.712	.601	0.467
7	WISC-R	Picture Completion	0	.531	0.697	.770	.605	0.795
8	WISC-R	Picture Arrangement	0	.545	0.732	.726	.640	0.859
9	WISC-R	Block Design	0	.591	0.806	.846	.642	0.876
10	WISC-R	Object Assembly	0	.531	0.792	.700	.635	0.946
11	WISC-R	Coding	0	.334	0.343	.675	.407	0.417
12	WISC-R	Mazes	0	.404	0.729	.720	.476	0.860
13	WISC-R	Tapping Span	0	.387	0.330	.800	.432	0.369
14	CGP	Vocabulary	1	.687	0.913	.900	.724	0.962
15	CGP	Picture Number	0	.378	0.641	-	-	-
16	CGP	Reading	1	.759	0.928	.810	.844	1.031
17	CGP	Letter Groups	0	.644	0.985	.739	.749	1.145
18	CGP	Math	1	.785	1.090	-	-	-
19	CGP	Mosaic Comparisons	0	.398	0.882	.770	.453	1.005
20	SAT	Verbal	1	.886	1.150	-	-	-
21	SAT	Math	1	.792	1.210	-	-	-
22	ACT	English	1	.722	1.160	-	-	-
23	ACT	Social Studies	1	.752	1.200	-	-	-
24	ACT	Science Reading Comprehension	1	.693	1.240	-	-	-
25	ACT	Math	1	.685	1.180	-	-	-
26	Benttler Gestalt	Form Perception	0	.572	0.690	-	-	-
27	ITPA	Auditory-Vocal Association	0	.678	0.960	-	-	-
28	-	Draw a Man	0	.460	0.110	-	-	-
29	WRAT	Spelling	1	.694	0.730	-	-	-
30	WRAT	Reading	1	.678	0.730	-	-	-
31	WRAT	Arithmetic	1	.702	0.550	-	-	-
32	ASVAB	General Science	1	.836	1.230	.860	.901	1.326
33	ASVAB	Arithmetic Reasoning	1	.813	1.160	.870	.872	1.244
34	ASVAB	Vocabulary	1	.840	1.300	.860	.906	1.402
35	ASVAB	Paragraph Comprehension	1	.763	1.080	.680	.925	1.310
36	ASVAB	Numerical Operations	1	.657	0.950	.710	.779	1.060
37	ASVAB	Coding Speed	0	.559	0.960	.820	.617	1.350
38	ASVAB	Auto-Shop Information	1	.615	1.230	.830	.675	0.960
39	ASVAB	Mathematics Knowledge	1	.788	0.880	.840	.859	0.949

Table 4.1 (Continued)

			Characteristic					
Test Number	Battery	Test name	Cultural load	*g* loading	Black-White Difference	Reliability	*g* loading*	Black-White Difference*
40	ASVAB	Mechanical Comprehension	1	.707	1.200	.830	.776	0.867
41	ASVAB	Electronics Information	1	.758	1.220	.800	.847	1.364
42	GATB	Vocabulary	1	.640	0.890	.860	.690	0.960
43	GATB	Numerical	1	.710	0.870	.840	.775	1.317
44	GATB	Spatial	0	.580	0.780	.810	.644	0.867
45	GATB	Form Perception	0	.700	0.550	.730	.819	0.644
46	GATB	Name Comparison	0	.710	0.570	.750	.820	0.658
47	GATB	Motor Coordination	0	.520	0.020	.810	.578	0.022
48	GATB	Finger Dexterity	0	.410	0.350	.670	.501	0.428
49	GATB	Manual Dexterity	0	.370	0.080	.730	.433	0.094
50	K-ABC	Hands Movements	0	.512	0.549	.760	.587	0.630
51	K-ABC	Number Recall	0	.528	0.037	.810	.587	0.041
52	K-ABC	Word Order	0	.594	0.153	.820	.656	0.169
53	K-ABC	Gestalt Closure	0	.455	0.367	.710	.540	0.436
54	K-ABC	Triangles	0	.606	0.610	.840	.661	0.666
55	K-ABC	Matrix Analogies	0	.642	0.478	.850	.697	0.519
56	K-ABC	Spatial Memory	0	.552	0.467	.800	.617	0.522
57	K-ABC	Photo Series	0	.593	0.513	.820	.655	0.567
58	K-ABC	Faces and Places	1	.704	0.366	.840	.768	0.400
59	K-ABC	Arithmetic	1	.755	0.817	.870	.810	0.876
60	K-ABC	Riddles	1	.781	0.853	.860	.842	0.920
61	K-ABC	Reading (Decoding)	1	.747	0.479	.920	.779	0.500
62	K-ABC	Reading (Comprehension)	1	.772	0.620	.910	.810	0.650
63	WAIS	Digit Span	0	.627	0.410	.680	.760	0.497
64	Lorge-Thorndike	Sentence Completion	1	.818	0.550	-	-	-
65	Raven	Matrices	0	.731	1.030	-	-	-
66	Ammons	Picture Vocabulary	1	.750	0.810	-	-	-
67	WAIS	Information	1	.678	0.430	.910	.710	0.451
68	WAIS	Coding	0	.698	0.790	.920	.727	0.824
69	CGP	Sentences	1	.680	0.780	.840	.742	0.851
70	CGP	Integrative Reasoning	0	.722	0.840	.730	.845	0.983
71	CGP	Intersections	0	.285	0.190	-	-	-
72	CGP	Information About Technology	1	.479	0.410	-	-	-
73	CGP	Algebra	1	.628	0.790	.880	.670	0.670

Table 4.2

Correlation	Cultural load	*g* loading (N = 73)	Black-White Differences (N = 73)
Cultural load	1.00	0.67	0.53
g loading (N=54)	0.63	1.00	0.58
Black-White Differences (N=54)	0.49	0.57	1.00

Note: *g* loadings and Black-White standardized mean differences before (upper triangle) and after (lower triangle) correction for attenuation. N denotes the number of subtests.

CHAPTER 5

ON THE INTERPRETATION OF THE CHC FACTOR GC

Abstract

On the one hand, the factors in the Cattell-Horn-Carroll (CHC) model of intelligence (e.g., g, Gf and Gc) are hypothesized to represent individual differences in unique psychological or biological capacities. On the other hand, the CHC factors Gf and Gc are interpreted as representing the theoretical variables fluid and crystallized intelligence in investment theory. This leads to a theoretical conflict because crystallized intelligence is purely a statistical entity. We conclude that if CHC factor Gc represents a capacity it cannot represent crystallized intelligence, and if Gc represents crystallized intelligence, it does not represent a capacity. We also conclude that in investment theory, the CHC factors Gf and g *represent one and the same capacity. In support of these conclusions, we present a reanalysis (confirmatory factor analysis) of a HCA (Human Cognitive Abilities project) data set. As expected, Gc was effectively absent, and Gf and* g *were statistically equivalent. Factors Gc and* g *could be removed from the CHC model without any reduction in model fit. We contend that these factors are redundant as explanatory variables.*

5.1 Is Gc an elusive construct?

In a recent editorial in *Intelligence*, McGrew (2009, p.1) made three recommendations. First, he recommended "to begin using the [Cattell-Horn-Carroll] taxonomy as a common nomenclature for describing research findings and a theoretical framework from which to test hypotheses regarding various aspects of human cognitive abilities". Second, he recommended "to reinvigorate the investigation of the structure of human intelligence". Third, he recommended to "access the Carroll [Human Cognitive Abilities Data Set Archive] datasets to test and evaluate structural models of human intelligence with contemporary methods (confirmatory factor analysis)". Encouraged by this editorial, we hope to contribute to the theoretical framework underlying the Cattell-Horn-Carroll (CHC) model, and to reinvigorate the investigation of the structure of human intelligence. We aim to do so by discussing the interpretation of CHC factor Gc. The discussion sheds light on the relation between the factors Gf, and *g*. Additionally, we test hypotheses stemming from this discussion by applying structural equation modeling to a dataset in the Human Cognitive Abilities Data Set Archive (HCA).

The CHC taxonomy (McGrew, 1997, 2009) is considered to be a well-validated model of human intelligence (e.g. Evans, Floyd, McGrew, & Leforgee, 2002). It is a synthesis of Cattell and Horn's extended 'Gf-Gc model' (Cattell, 1987; Horn, 1968, 1991; Horn & Stankov, 1982) and Carroll's (1993) 'three-stratum model'. From the statistical point of view, the common factors in these models represent the common variance among test scores. However, substantively, they are often interpreted in terms of substantive underlying variables. Carroll, for instance, stated:

> "[U]nderlying each factor of the three-stratum theory there is a specific state or substrate that exists in the individual and that accounts for his or her ability or inability to perform tasks in which that ability is called for" (Carroll, 1996, p.15).

Although factor analysis may support such substantive, or realist, interpretation of common factors, it cannot prove that this interpretation is correct (see Borsboom & Mellenbergh. 2002; Bartholomew, 2004; van der Maas et al., 2006; Borsboom & Dolan, 2006). Hence, to attach substantive interpretation to the results of factor analysis, i.e., to interpret a given common factor as a substantive, realistic underlying variable, theory is required.

Concerning the theoretical status of the factors in CHC model, we note certain disagreements among Cattell, Horn, and Carroll, despite large agreement about the statistical structure. Consider the following two examples. First, in the CHC model the second order factors are positively intercorrelated, which opens the possibility of positing a general intelligence factor at the apex of the hierarchy. The three-stratum model includes such a factor (*g*), but the Gf-Gc model does not. This salient difference between the three-stratum model and the Gf-Gc model has a theoretical background: Whereas Carroll (1998) took a realist position concerning the general factor by interpreting it as a unique cognitive ability, Horn (e.g. Horn & Noll, 1997) rejected this realist interpretation, and viewed general intelligence as merely a statistical entity; he considered it to be as nothing more than an aggregate of various cognitive abilities.

Second, Carroll (1993) argued that it is a matter of preference whether Gc is interpreted as verbal ability or as crystallized intelligence, whereas Cattell (1987) maintained that verbal ability and crystallized intelligence are distinct. In view of these disagreements, we conclude that the theoretical status of the CHC factors is open to debate. We also subscribe to Keith and Reynolds (2010)'s conclusion, after discussing "what we have learned over the past 20 years of research on the nature and measurement of intelligence from a CHC perspective." (p. 647):

> "Gc remains an elusive construct, and researchers often talk past each other when discussing Gc, with it being referred to as crystallized intelligence, academic achievement, verbal ability, or comprehension/knowledge, to name a few [...] Clarification about the nature of Gc versus verbal ability and achievement would be useful" (Keith & Reynolds, 2010, p. 643).

Our first aim is therefore to discuss the status of CHC factor Gc as representing an underlying variable. Our second aim is to address the question whether or not Gf and *g* represent distinct underlying variables.

In accepting the CHC model, and referring to factors Gf and Gc as fluid and crystallized intelligence, researchers seem to subscribe, implicitly or explicitly, to Cattell's (1963; 1971; 1987) investment theory of intelligence. This theory includes two central hypotheses. One hypothesis states that the variable fluid intelligence (represented by CHC factor Gf) is a real, substantive cause of individual differences in intelligence test performance, e.g. a relation perceiving or reasoning capacity. The second hypothesis, the investment hypothesis, states that during the course of development people 'invest' their fluid intelligence to acquire domain specific knowledge and skills (e.g. vocabulary and arithmetic skills), which are called crystallized abilities. The conglomerate of crystallized abilities is crystallized intelligence. Individual differences in the variable crystallized intelligence are usually denoted as Gc. In some interpretations of the CHC model, this theoretical variable crystallized intelligence is supposed to be represented by the factor Gc.

We discuss, in light of Cattell's (1987) investment theory, the theoretical status of CHC factor Gc as representing (1) crystallized intelligence, and (2) a substantive underlying variable. In so doing, we hope to advance our understanding of the relations between Gf, *g*, and Gc in the CHC model, hence the theory behind the CHC model. The chapter is organized as follows. We first discuss the concept of a latent (underlying) variable, and the scientific philosophical positions that one can take with regard to this concept (i.e., realist versus nonrealist), because the explication of one's position with respect to latent variables is important in general, but indispensible in a discussion of the nature of Gc. Next, we present a brief review of investment theory, the development of the CHC taxonomy, and various interpretations of the CHC factor Gc as given in the literature.

We conclude first that crystallized intelligence is purely a statistical entity, which does not represent an underlying cause of individual differences in intelligence, and we conclude that within investment theory, fluid and general intelligence represent one and the same variable (see also Kvist & Gustafsson, 2008). Second, we conclude that within investment theory two interpretations of CHC factor Gc as representing a realistic underlying variable remain viable. Gc may represent individual differences in exposure to information via education, which is consistent with Cattell (1971, 1987). Alternatively, Gc may represent individual differences in verbal comprehension, which is consistent with Johnson and Bouchard (2005). Which interpretation is preferable depends on the exact

constitution of the (CHC) test battery, and thus on what tests are found to load on the factor designated as Gc.

Next, in support of our conclusions, we present a confirmatory factor analysis of a subset of HCA data set ALLI00 (Allison, 1960; Carroll, 1993). In this dataset, CHC factor Gc was found to be indistinguishable from a verbal comprehension factor. In addition, CHC factor Gf, the variable that represented individual differences in fluid intelligence, was indeed indistinguishable from the CHC factor *g*. Our final conclusions are that in the CHC model and investment theory (a) crystallized intelligence does not represent a true psychological or biological capacity; and (b) CHC factors Gf and *g* represent one and the same capacity. The implications for the CHC taxonomy are that Gc and Gf or *g* can be removed from the CHC factor model. We end with a brief discussion concerning the theoretical status of other CHC factors, and, following McGrew (2009), we offer some recommendations concerning the CHC taxonomy as a common nomenclature.

5.2 The Interpretation of a Latent Variable

In the literature, the term latent variable is used in more ways than one (Borsboom, Mellenbergh, & van Heerden, 2003). Firstly, the term latent variable can refer to a formal, mathematical concept. This concept is used in mathematical treatments of measurement and structural relation models, and is usually symbolized by a letter, e.g. θ in Item Response Theory, or η in Structural Equation Modeling. Beyond the mathematical treatment the formal concept has no meaning, i.e., it does not reveal anything about reality. Secondly, the term latent variable can refer to an operational concept, which is the result of an algebraic function of observed scores, e.g. a weighted sum score, like a full-scale IQ score. The factors extracted from a test battery (e.g. the CHC factors) represent instances of the operational concept. Since there is nothing latent about an algebraic function, there is nothing latent about the operational concept. The link between the formal concept (label) and the operational concept (algebraic result) is not self-evident, and requires theoretical interpretation, such as '*g* represents individual differences in general intelligence, and a part of the common variance in intelligence tests scores corresponds to individual differences in general intelligence'.

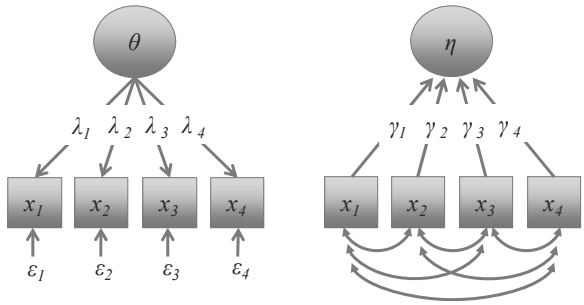
With respect to the theoretical interpretation of the link between formal concept (label) and operational concept (algebraic result), there are essentially three philosophical positions (Borsboom et al., 2003). First, one can take an operationalist position. This position holds that there is nothing beyond the operational variable (algebraic result). This implies, *inter alia*, that different sets of items necessarily measure different variables (e.g. different IQ test batteries or subsets of batteries measure different variables, or 'intelligences', so to speak; see Borsboom & Mellenbergh, 2002). We therefore maintain that "operationalism and latent variable theory are fundamentally incompatible." (Borsboom et al., p. 207). Second, one can take a constructivist position.[8] This position holds that the latent variable is a construction of the scientist. It implies that different IQ batteries may measure or estimate the same variable (or set of variables), but that this variable need not be attributed existence independent of the scientist (e.g. although a construct called general intelligence is not a real-world entity, different batteries estimate this construct; the variable *g* represents individual differences in this construct). Third, one can take a realist position. A realist position holds that the unobserved (i.e. latent) variable does exist independent of the scientist (e.g. *g* represents individual differences in a real-world entity, and the tests on IQ batteries estimate one's relative position on this variable).

With respect to causation, the constructivist and realist positions are associated with different interpretations of measurement (e.g. Borsboom et al., 2003) and structural relations (e.g. Diamantopoulos et al., 2008). The constructivist position implies that a formative model is appropriate, while the realist position implies a reflective model is appropriate (Bollen & Lennox, 1991; Edwards & Bagozzi, 2000; Edwards, 2011, Diamantopoulos et al., 2008). The latent variable in a reflective model (as depicted in Figure 5.1a) is conceptualized as *determining* the measurements,

[8] Strictly, operationalism is a kind of constructivism, but we mean the term constructivism to denote a broader class of views (for example, the view of van Fraassen, 1980).

whereas the latent variable in a formative model (as depicted in Figure 5.1b) is conceptualized as *constructed from* measured (or lower level) variables.

Reflective modeling is customary in psychometric measurement models (e.g. IRT models) and in many structural relation models (e.g. factor models of intelligence), and appears to be implicit in much psychological theory. For instance, the measurement of cognitive abilities, say, working memory capacity, usually involves reflective modeling. Here, one assumes that individual differences in the unobserved variable give rise to individual differences in the indicators. So, if persons differ in working memory capacity (the unobserved variable), these differences will give rise to individual differences, hence variance, in observed indicators of working memory capacity, e.g. digit span.



a) b)

Figure 5.1 Two measurement models. The reflective model (a), in which individual differences in the latent variable (e.g. cognitive ability) cause individual differences in the indicators. Positive correlations between the indicators are the result of the common dependence on the latent variable. The formative model (b), in which individual differences in the indicators determine the unobserved variable. It summarizes the positive correlations among the indicators.

Although formative modeling (Edwards & Bagozzi, 2000) is less customary in psychology (in contrast to economic modeling and marketing research), it is increasingly the subject of debate (Diamantopoulos et al., 2008; Kievit et al., 2011). Recently, several special issues were devoted to reflective versus formative modeling (see, e.g. Diamantopoulos, 2008). As an example of formative modeling consider the measurement model for SES (socioeconomic status). The construct SES is a linear combination of relevant variables, such as income and occupation. Here, differences in the indicators give rise to differences in the unobserved variable, rather than vice versa. For instance, individual differences in salary determine individual differences in SES, and a raise in one's salary results in an increase in one's level of SES. Constructivist variables as 'general health', 'fitness', and 'mental health' may also be conceptualized as formative variables; one does not have all kinds of health problems because of poor general health, rather one has a poor general health, because of all kinds of health problems.

We note that the interpretation of a latent variable as a realistic, underlying variable is compatible with a reflective model, but cannot be justified merely by reference to the results of factor analysis, or the application of any other measurement or structural relation model (Borsboom et al., 2003). The common factor model may be compatible with a reflective model, but the results of a factor analysis cannot prove the existence of a causal underlying variable. To elucidate this, consider the following. In intelligence research, the positive correlations amongst cognitive ability test scores and the presence of a statistical general factor are often explained by positing a common dependence on a realistic underlying variable, *g* (Spearman, 1927; Jensen, 1998; Carroll, 1993). However, several alternative explanations have been proposed that can also account for the positive correlations, hence for the statistical general factor (Thorndike, 1927; Thomson, 1951; Cattell, 1987, pp. 365-433; Anderson, 2001; Penke, Denissen, & Miller, 2007; van der Maas et al., 2006; Bartholomew et al., 2009; Dickens, 2008). In the mutualism theory of van der Maas et al., for instance, the positive

correlations are the result of mutually beneficial interactions among cognitive processes that are assumed to take place throughout development. In this case, the general factor obtained in a factor analysis does not correspond to a realistic, underlying causal variable; it simply summarizes the positive correlations among cognitive abilities. In the mutualism theory, the general factor is conceptualized as a constructivist variable; from the realist standpoint this general factor does not represent a causal underlying variable. However, a formative model of general intelligence (e.g. Diamantopoulos et al., 2008, Figure 3) would be appropriate given mutualism.

In light of the above, we maintain that the variables and relations between the variables revealed in factor analyses require theoretical interpretation, and that, in this interpretation, the scientist's philosophical position (realist versus non-realist) concerning these variables is important.

Throughout this chapter, we approach the factors in the Cattell-Horn-Carroll taxonomy from the realist position, because these factors are often interpreted as such (e.g. Carroll, 1993). We note that, within the CHC setup, Gc is considered to represent individual differences in crystallized intelligence. The present aim is to demonstrate theoretically and empirically that the interpretation of Gc as representing a substantive, causal underlying variable (e.g. a biological or psychological capacity), and the interpretation of Gc as crystallized intelligence are mutually incompatible. That is, if Gc represents a capacity it cannot represent crystallized intelligence, and if Gc represents crystallized intelligence, it does not represent a capacity. To demonstrate this theoretically, we review the investment theory of fluid and crystallized intelligence, individual differences models of intelligence that include a factor denoted as Gc (e.g. the CHC model), and interpretations of Gc as an substantive underlying variable. Subsequently, to support our main conclusions empirically, we fit the CHC model to a HCA data set.

5.3 Cattell's Investment Theory of Fluid and Crystallized Intelligence

According to Cattell (1941; 1943; 1963), intelligence tests measure two kinds of abilities, which differ greatly with respect to the role knowledge. One kind is measured by tests in which individual differences in prior knowledge, rather than individual differences in cognitive processing, determine individual differences in the subtest scores. Cattell called these abilities crystallized abilities. The other kind, which Cattell called fluid abilities, is measured by tests in which prior knowledge does not discriminate. That is, although knowledge is required to solve the items on a fluid ability test, this knowledge is considered to be equally novel or equally known to the testees. Hence, individual differences in the subtest scores are not primarily determined by differences in prior knowledge, but by differences in cognitive processing.

Furthermore, in his early work, Cattell argued that Spearman's general factor of intelligence, *g*, can be split into two distinct factors, although these factors, "[b]eing cooperative, are very difficult to separate" (Cattell, 1963, p.2). In addition, fluid and crystallized abilities were found to load on both factors, but fluid abilities loaded primarily on the one factor and crystallized abilities primarily on the other. Therefore Cattell postulated that Spearman's *g* comprised two general factors of intelligence: fluid intelligence and crystallized intelligence.[9]

Cattell (1963, 1987) entertained specific ideas about the nature and ontology of fluid and crystallized intelligence. He viewed fluid intelligence as a 'source trait' (Cattell, 1987, p.15), i.e., an underlying causal variable, or a latent variable of the reflective kind (Bollen & Lennox, 1991). He considered fluid intelligence to be a 'single relation perceiving capacity', which is related to the maturation of the cortex (Cattell, 1987, p. 138). In addition, in his investment theory, Cattell attributed a causal role to fluid intelligence in the development of other cognitive abilities. He hypothesized that people 'invest' their fluid intelligence to acquire specific skills, strategies, and knowledge in all kinds of fields (Cattell, 1963; Cattell, 1987, pp. 138-147), for example vocabulary

[9] Other observations have been cited in support of the distinction between fluid and crystallized abilities (Cattell, 1963, 1987; Hunt, 2000). For example, fluid and crystallized abilities display different developmental curves (see also Horn & Cattell, 1967, McArdle et al, 2002; cf. Schaie & Strother, 1968). In addition, brain injuries have distinct effects on crystallized abilities and fluid abilities (Cattell, 1963; Lee, Choi, & Gray, 2007).

and arithmetic skills. These acquired abilities are the crystallized abilities. The terms fluid and crystallized have their origin in investment theory: "The term crystallized is meant to imply this freezing in a specific shape of what was once fluid ability" (Cattell, 1987, p. 140). Crystallized abilities are thus domain specific; their expressions are 'tied to particular areas' (Cattell, 1987, p. 139), whereas fluid intelligence, in contrast, is hypothesized to be domain transcending, i.e., "has the fluid quality of being directable to almost any problem" (Cattell, 1987, p. 97). The investment hypothesis boils down to the hypothesis that knowledge acquisition, hence learning, requires relation perceiving, with fluid intelligence representing the capacity to do this.

It follows from the investment hypothesis that, *ceteris paribus*, people with high levels of fluid intelligence tend to acquire more and better-developed crystallized abilities than people with low levels of fluid intelligence. Or, stated more mundanely, they learn faster. As a result, provided investment occurs in several areas, crystallized abilities across these areas become positively intercorrelated, and a factor analysis of scores on tests that measure the relevant crystallized abilities will thus yield a common factor. Accordingly, crystallized intelligence is defined as a conglomerate of crystallized abilities. Individual differences in crystallized intelligence are denoted Gc or gc, and individual differences in fluid intelligence Gf or gf. Because fluid intelligence is invested to acquire crystallized abilities across a variety of domains, crystallized intelligence, and hence the theoretical variable Gc, is also associated with abilities across domains (see also Hunt, 2000).

In the literature, 'investment' is taken quite literally. Consider the following examples:

> "As the term 'investment theory' indicates, gf is liable, in its generation of gc, to all risks of an investment. Laziness may cause it scarcely to be invested at all; differences of individual interest may cause it to be invested in directions different from that in which 'traditional' intelligence tests measure it." (Cattell, 1987, p. 334).

> "Persons high in Gf tend to acquire more Gc (i.e., they reap greater returns on their initial investment) from their opportunities for learning than persons of lower Gf." (Jensen, 1998, p.123)

In psychometric modeling fluid and crystallized intelligence have become popular constructs: In the model of Cattell and Horn (Cattell, 1987; Horn; 1968), Carroll's three stratum model (Carroll, 1993), as well as in combined models (e.g. Gustafsson, 1984, Alfonso et al., 2005), the higher order factor Gf is considered to represent individual differences in fluid intelligence, whereas Gc is considered to represent individual differences in crystallized intelligence. Since crystallized intelligence is considered to be the result of investment of fluid intelligence, researchers consider *factor* Gc, in its interpretation as representing individual differences in crystallized intelligence, to arise as result of investment of Gf: "Gc […] develops out of the investment of Gf." (Ackerman, 2003).

Just as we take the reflective model seriously (as a causal model of inter-individual differences), we take investment theory seriously. That is, we take Cattell's theory as a theory of investment of peoples' fluid intelligence in a variety of areas or domains. We note, however, that investment theory concerns an intra-individual process, i.e., one at the level of the individual, whereas the CHC model of intelligence concerns the structure of interindividual differences in intelligence. The link between investment theory and the factors in the CHC model is not self-evident and has yet to be established. One difficulty in establishing this link is that investment theory has not undergone any significant changes, whereas the psychometric models of intelligence have. The relation between investment theory and the CHC model has become more complicated and less clear than the relation between investment theory and early models.

To establish the link between investment theory and the factors in the CHC model, a historical overview is useful. Therefore, below we consider the development of the CHC model in more detail. We also discuss the relation between Gf and *g*. We consider this discussion to be relevant for two reasons. First, some researchers consider crystallized intelligence to be the result of investment of general intelligence (e.g. Carroll, 1993; Jensen, 1998). Second, factors denoted as *g* and Gf are sometimes found to be indistinguishable in factor analysis (e.g. Gustafsson, 1994), and are

therefore considered to represent the same underlying variable (e.g. Jensen, 1998). This finding that Gf and *g* are indistinguishable has been associated with investment theory (Kvist & Gustafsson 2008).

5.4 The (Development of the) CHC factor model

As mentioned above, in his early work Cattell (Cattell, 1943, 1957, 1963) argued that Spearman's general factor of intelligence (*g*) can be split into two distinct factors. Because fluid abilities tended to load primarily on the one factor, and crystallized abilities on the other, he called the one factor fluid intelligence (Gf) and the other one crystallized intelligence (Gc). Originally the theory of fluid and crystallized intelligence involved a bipartite taxonomy of individual differences in cognitive abilities, and investment theory was an attempt to account for the ontology of these two factors.

This bipartite Gf-Gc model was expanded following the inclusion of a broader range of subtests measuring cognitive abilities and intellectual achievement (Horn, 1968, 1991; Horn & Stankov, 1982). Additional higher order factors were extracted from diverse batteries. These higher order factors are instances of the operational concept of a latent variable (see above). The factors were labeled Gv, Gsm, Glr, Gs, Ga, Gt, Gq, Grw, and were interpreted as visual perception, short-term memory, long-term storage and retrieval, processing speed, auditory processing ability, reaction time, quantitative ability, and broad reading and writing ability, respectively. These interpretations provide the theoretical link between the formal and operational concepts (see above). As argued above, in order to interpret a factor as representing a causal underlying variable or as merely a statistical summary, one requires an explicit scientific philosophical position, that is, realist or nonrealist.

Because the theory of fluid and crystallized intelligence provided its point of departure, the extended model is still referred to as the Gf-Gc model (Horn & Noll, 1997, Alfonso et al. 2005). This extended Gf-Gc model closely resembles Carroll's three-stratum model (Carroll, 1993; 1996; 2003; Bickley, Keith, Wolfle, 1995). The latter may be regarded as a synthesis of the Gf-Gc model and Spearman's *g* model (Alfonso et al. 2005; McGrew, 2009). Like the Gf-Gc model, the three-stratum model incorporates higher order factors, denoted Gf and Gc, as well as many other higher order factors, defined by Horn and Cattell. However in some aspects the models differ. A relatively minor difference is that several narrow second order factors in the Gf-Gc model are subsumed under broader second order factors in the three-stratum model (e.g. Grw in the Gf-Gc model is subsumed under Gc in the three-stratum model). The major difference between the models is that Carroll posited a general factor (*g*) at the highest level in order to account for the positive intercorrelations between the broad factors at the second stratum, whereas Horn and Cattell did not.

Correlated factors Gf and Gc are invariably included in the CHC model, but the inclusion of *g* remains an issue of debate (McGrew, 1997, Alfonso et al. 2005). In some representations of the CHC model *g* is included at the apex of the hierarchical factor model (e.g. Vanderwood, McGrew, Flanagan, Keith, 2001), whereas in others it is omitted (e.g. Alfonso et al., 2005). Note in this connection that

> "The exclusion of *g* does not mean that the integrated model does not subscribe to a separate general human ability or that *g* does not exist. Rather, it was omitted […] since it has little practical relevance to cross-battery assessment and interpretation." (Flanagan, Ortiz, & Alfonso, 2006, p. 6; McGrew, 1997).

Whether or not to include a general factor is related to its actual theoretical conceptualization. Horn in particular objected to the extraction of a general factor, because in his opinion such a general factor does not represent a unique cognitive ability: "[G]eneral factors represent different mixture measures, not one general intelligence." (Horn & Noll, 1997, p. 68, but see Johnson et al., 2004; Johnson et al., 2008, for a different view). We conclude that in Horn's view, *g*, as an aggregate rather than a unique cognitive ability, is better conceptualized as a constructivist variable, and that a formative model of *g* is more appropriate than a reflective (i.e., underlying variable) model (e.g. Diamantopoulos et al., 2008, Figure 2-II, p. 1207).

Related to the discussion of whether *g* represents a substantive underlying variable is the discussion whether fluid and general intelligence are distinct cognitive abilities or capacities. Some

researchers (e.g. Jensen, 1998) have argued that *g* and Gf represent one and the same variable in view of the following. In another attempt to synthesize existing models of cognitive abilities, Gustafsson (1984) extracted a third-order general factor (*g*) as well as a number of second order factors, including factors referred to as fluid intelligence (Gf) and crystallized intelligence (Gf). In this model, which he called the HILI model (HIerarchical LIsrel model) of intelligence, the correlation between *g* and Gf approached one, which renders them indistinct as common factors, i.e., as statistical entities.

This finding was replicated in some (Gustafsson, 1988, Keith, 2005; Reynolds & Keith, 2007), but not all other studies (Carroll, 1993, 2003). We conclude therefore that the relation between *g* and Gf remains puzzling, and that the theoretical status of *g* remains open to debate. With our discussion of Gc, we hope to shed light on this relation.

Cattell's (1987, p 141) proposal that *g* represents "the fluid ability of yesteryear, which fathered the present fluid ability" can be regarded as a first step towards explaining the mixed results concerning the distinctiveness of Gf and *g*, and as providing an account of the relation between these variables. However, this proposal raises questions about investment theory, and thus about the ontology of Gf and Gc, because originally Cattell argued that highest order factor *g* can be split into fluid intelligence (a reasoning factor) and crystallized intelligence (a knowledge factor). But with this proposal Cattell is reintroducing *g* as single variable at the highest order.

A second step towards explaining these mixed results is provided by Kvist and Gustafsson (2008), who argued that the relation between Gf and *g* is a function of cultural background:

> " [T]he Gf-factor would be equal to the *g*-factor in populations which are homogeneous with respect to opportunity to having learned the knowledge and skills measured but […] this relationship would not hold in heterogeneous populations where subgroups differ with respect to opportunity to learn" (Kvist & Gustafsson, 2008, p. 433).

They consider this to be consistent with Cattell's investment theory of fluid and crystallized intelligence. But this also raises questions about investment theory. As has been suggested by Jensen (1998, see below), the splitting of *g* into separate factors Gf and Gc might be due to heterogeneity of the subject sample.

To advance our understanding of the relations among the factors *g*, Gf, and Gc, it is important to establish whether the interpretation of Gc as crystallized intelligence is compatible with the interpretation of the CHC factor Gc as a substantive underlying variable.

5.5 The Interpretation of Factor Gc as a Crystallized Intelligence

In the literature, the common factor denoted Gc has been interpreted in many ways (see Table 5.1). These interpretations are associated with different scenarios concerning the ontogeny of crystallized intelligence in the investment hypothesis. In the following two sections, we discuss the interpretations of CHC factor Gc in the light of the investment hypothesis.

The first scenario is based on a literal interpretation of the investment hypothesis (see Figure 5.2). People invest their fluid intelligence, and so acquire crystallized abilities in a variety of domains. Gf represents individual differences on the psychological dimension fluid intelligence (Cattell, 1963, 1987). The investment hypothesis predicts that people who have a high level of fluid intelligence (have a high position on Gf) will acquire more and better developed crystallized abilities than people with low levels of fluid intelligence. Given that people generally spread their investments, and given they do so in similar ways within a given society, crystallized abilities become positively correlated. The conglomerate or total of a person's crystallized abilities is a person's crystallized intelligence.

We note that in this literal interpretation of the investment hypothesis, factor Gf is the ultimate source of the correlations between the crystallized abilities. In a factor analysis of a battery that includes both fluid and crystallized abilities, a distinct Gc factor should not appear. Given this scenario, we should conclude that (1) crystallized intelligence is not a latent variable of the reflective kind (i.e., an underlying variable), but of the formative kind, and (2) this scenario does not support

factor models that include two separate factors Gf (fluid intelligence) and Gc (crystallized intelligence), e.g. the CHC model.

We also note that crystallized intelligence, as measured in psychometric intelligence tests, emerges as a particular composite of crystallized abilities (namely the total of those crystallized abilities that are measured by the subtests). In the ASVAB, for example, the factor designated crystallized intelligence summarizes mathematical and verbal skills (Roberts et al. 2000); in the Woodcock-Johnson (WJ)-III achievement battery this factor summarizes general and academic knowledge (McGrew & Woodcock, 2001). Furthermore, in the view of Horn and Blankson (2005), the best estimate of crystallized intelligence is the total of all achievements in the WJ-R achievement battery. In this sense, crystallized intelligence is merely an operational variable (see above), because different intelligence tests measure different conglomerates of crystallized abilities, hence measure different 'crystallized intelligences'.

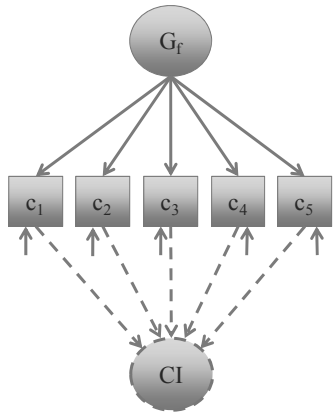


Figure 5.2 A literal interpretation of the investment hypothesis. Gf represents individual differences on the psychological dimension fluid intelligence. Fluid intelligence is invested to acquire crystallized abilities in a variety of domains. As a result, crystallized abilities are positively correlated. In a factor analysis including Gf, a distinct Gc factor does not appear; Gf accounts for the correlations between the crystallized abilities. Crystallized intelligence (CI) is not a latent variable of the reflective kind, but of the formative kind.

In the second scenario, we reverse the reasoning. We suppose crystallized intelligence is a distinct causal variable, i.e., a substantive reflective latent variable, and ask whether this ability can be reasonably considered to be the result of investment. As mentioned, in investment theory crystallized intelligence is associated with crystallized abilities across domains, so assuming Gc is a substantive reflective latent variable, it has to exert its influence across domains. This interpretation of Gc is consistent with, for example, Schweizer and Koch (2002)'s revision of investment theory, in which crystallized intelligence is defined as 'the potential to perform on the basis of' crystallized abilities (p. 58), or with Marks' (2010, p.646) contention that crystallized intelligence should not be equated with knowledge and skills itself, but that it is represents the ability to use them.

We maintain that in this second scenario Gc cannot both be the result of investment of fluid intelligence and a unique domain transcending cognitive ability (e.g. as proposed by Schweizer & Koch, 2002, and Marks, 2010), because this would imply that transfer of training occurs across domains. To illustrate this, imagine that one invests one's fluid intelligence in a particular area, say, by practicing a mathematical skill. Then, according to the investment hypothesis, this increases one's level of crystallized intelligence, i.e., Gc (the ability or potential to perform on the basis of knowledge). An increase in Gc (ability or potential), in turn, should have a direct positive effect on all

the indicators of Gc, such as vocabulary, general information, spelling, etc.[10]. So, in this scenario, training the mathematical skill should result in an increase in all other crystallized abilities. This poses a problem, because although there is empirical evidence for transfer of training *within* domains, evidence for transfer of training *across* domains is absent (e.g. Cattell, 1987; Ceci, 1990, p.185). In this light, we conclude that this scenario is not tenable.

Given the investment hypothesis, we can draw the following conclusions. First, crystallized intelligence is not a latent variable of the reflective kind, but of the formative kind; it is not a substantive latent variable. Second, crystallized intelligence should not appear as a common factor (Gc) that is distinct from Gf. Hence, the question remains what meaning to attach to the finding in factor analysis of the distinct factor commonly designated Gc. To answer this, we have to identify a second domain transcending cause of individual differences, which we denote *X* (see Figure 5.3). In fact, Cattell's *investment theory* is broader than solely *the investment hypothesis* of fluid intelligence, because it stipulates that such second cause of individual differences is necessary to explain the patterns of covariance between fluid and crystallized abilities (Cattell, 1987). Regardless of what this second cause, *X*, may represent, it must represent a variable different from crystallized intelligence. In the next section, we discuss candidates for *X*.

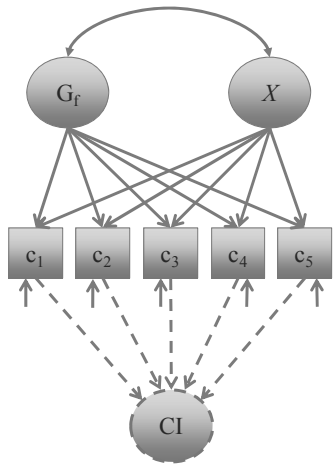


Figure 5.3 If a second factor Gc is found in factor analysis and if this factor is interpreted as a variable of the reflective kind (and as a variable different from Gf), it must represent variable *X*. Crystallized intelligence (CI) is a latent variable of the formative kind. Plausible candidates for *X* are exposure to information via education and verbal comprehension.

5.5 The Interpretation of Gc as a Causal Variable

If we find Gc in a factor analysis and want to interpret this factor as an underlying cause of individual differences (and different from Gf; i.e., the *X* in Figure 5.3), then what can it be? As Guilford (1980) pointed out, the factor Gc may well represent an environmental cause of individual differences, rather

[10] On the level of individual differences it implies the following. Given that 2 persons (person A and person B) have same level of fluid intelligence and the same history of investment, if person A improves a mathematical skill, thereby increasing the level of Gc (hence taking a higher position on the dimension Gc), person A must tend to perform better on all crystallized tasks (vocabulary, general information, spelling, etc.) as compared to person B.

than a psychological (or biological) capacity. Furthermore, assuming that it represents a psychological capacity, it can represent a cognitive capacity or some other psychological capacity.

If we interpret *X* as the potential to perform on the basis of acquired knowledge (Schweizer & Koch, 2002) - but not as the result of investment (see above) - we can regard *X* as the capacity to store, retain or retrieve information, thus as an capacity to memorize. However, the CHC model (Horn, 1991; Carroll, 1993; McGrew, 1997) already includes a short term memory (Gsm) and a broad retrieval and storage capacity (Glr) factor, and both are distinct from this substantiation of *X*. Furthermore, if *X* represents such a capacity, it is difficult to explain why *X* has additional loadings on tests in which information is new to the participants (Cattell, 1963; Carroll, 1993), i.e., tests on which better memory, or storage or retrieval capacity confers no clear advantage. We conclude that the capacity to store, retain or retrieve information is not a plausible candidate for *X*.

As discussed by Cattell (1987), other candidates for *X* include motivation, broad interest, and persistence, provided that these can feature as unitary, causal variables. In view of empirical results, variable *X* must be quite highly correlated with fluid intelligence (because Gc correlates substantially with Gf). However, fluid intelligence is not particularly highly correlated with any of these variables (e.g. Ackerman, 1996). In addition, although crystallized abilities are correlated with them, the correlations are low to moderate, which suggests that motivation, broad interest, or persistence cannot be equated with Gc. We conclude that these variables are not plausible candidates for variable *X* either.

Are there any other candidates? From the literature, we conclude that two interpretations of *X* (Gc) as representing an underlying variable are tenable. Consistent with Cattell (1971, 1987), it may represent individual differences in exposure to information via education. Consistent with Johnson and Bouchard (2005), it may represent individual differences in verbal comprehension. Below, we discuss these options in more detail.

5.5.1 Candidate 1: *X* is Exposure to Information Via Education

Exposure to information via education represents the effect of individual differences in educational attainment (especially those due to enrollment in formal educational systems), and the quality of that education. It is affected by opportunities, by social influences, such as cultural values, and by psychological influences, such as motivation, interests, and persistence.[11]

We consider the interpretation of Gc as exposure to information via education to be plausible, because it is consistent with salient findings in intelligence research. First, the high correlation between Gf (or *g*) and exposure can be explained as follows: People who have a higher level of fluid or general intelligence generally receive more (longer and better) education. Second, the educational curriculum affords many opportunities to acquire specific abilities in all kinds of domains. Hence, exposure to information via education can explain additional common variance across domains after Gf (or *g*) has been taken into account. Third, in many IQ batteries, subtests that load on the variable designated Gc typically measure knowledge and skills acquired during the formal education. Fourth, exposure to information via education is consistent with the empirical finding that delinquents often have disproportionately low levels of crystallized abilities (e.g. Blecker, 1983; Law & Faison, 1996). They are likely to have invested their fluid intelligence to acquire knowledge, skills, and solving strategies that are not measured by intelligence tests (Cattell, 1987, p. 137).[5]

The interpretation of Gc as exposure to information via education is also consistent with Cattell's investment theory. First, the distinction between fluid and crystallized abilities remains valid. Second, it is consistent with the observation that IQ subtests display different age curves as a function of knowledge (McArdle et al., 2002). The more differences in prior knowledge (rather than differences in reasoning) determine differences in the testees' test scores, the less the age curve

[11] In a society in which social economic opportunities influence access to educational institutions, SES will predict exposure to education. In a society in which such opportunities do not influence access, SES will not predict well exposure to education. Group differences in exposure to education between people from different cultures, however, may well reflect group differences in (social economic) circumstances or cultural values, even if these variables do not predict well exposure to education within a group (see also Kvist & Gustaffson, 2008).

declines. Also, the more time one invests to acquire or maintain this knowledge, the less one's age curve declines. We therefore expect the smallest decline in one's area(s) of expertise. Third, crystallized intelligence can still be defined in the original way, i.e., as a conglomeration of knowledge, skills, and solving strategies across domains (Cattell, 1963, 1987). Finally, Cattell himself proposed that the dislocation of the variable crystallized intelligence from Gf is due to the influence of education. Education itself, in turn, can be affected by influences described above, such as personality, motivation, persistence, and interests.

5.5.2 Candidate 2: *X* is Verbal Comprehension

Although exposure to information through education is a plausible candidate theoretically, we question whether in practice the factor Gc in the CHC taxonomy truly represents a domain transcending factor. Many of the tests that load on the factor denoted Gc rely, directly or indirectly, on language (Carroll, 1993; Hunt, 2000). For this reason, Carroll (1993) suggested that it is a matter of preference whether one views Gc as crystallized intelligence or verbal ability. Theoretically, however, it is important to make a distinction between crystallized intelligence and verbal ability for three reasons (Cattell, 1987; Hunt, 2000; Johnson & Bouchard, 2005; Johnson & Gottesman, 2006). First, as explained above, the investment hypothesis in itself does not explain why Gc ('crystallized intelligence') should emerge as a common factor, whereas equating Gc with verbal comprehension does (provided that verbal comprehension is the manifestation of some underlying system that is different from the system that underlies fluid intelligence). Second, with respect to causality, crystallized intelligence must be considered to be a formative variable, whereas verbal comprehension can reasonably be assumed to be a reflective variable (differences in verbal comprehension give rise to differences in its indicators). Third, investment theory aims to give a developmental account of cognitive abilities, and aims to explain differences in age trends of tests that depend mainly on reasoning and tests that depend mainly on knowledge (Cattell, 1963). In itself, the interpretation of Gc as verbal comprehension does not provide such explanation (Hunt, 2000).

In this light, Beauducell, Brocke, & Liepmann (2001), Beauducell & Kersting (2002), and Amthauer, Brocke, Liepmann, & Beauducell (2001) are relevant. In these studies, participants were tested using a battery comprising verbal, numerical, and figural *reasoning* tasks, as well as verbal, numerical, and figural *knowledge* tests. In factor analyses all tests loaded on factors denoted Gf and Gc. However, if tasks required mainly reasoning they loaded primarily on Gf, and if they required mainly prior knowledge they loaded primarily on Gc. These results are consistent with Ackerman's distinction between intelligence as a process and intelligence as knowledge (Ackerman, 1996, 2000), and with Cattell's (1987, p.114) remark that "verbal (synonym, analogies) tests can be made to load fluid intelligence very substantially (and crystallized relatively little) if the words are chosen to be easily within the vocabulary of the group tested" (for examples, see Horn, 1970, p. 59, and Jensen, 1998, p. 123). Thus, at first sight, the distinction between a domain transcending (fluid) reasoning factor and a domain transcending knowledge factor seems better supported empirically than a distinction between a nonverbal and verbal factor. However, a close inspection of item content suggests a finer distinction is necessary.

In the Beauducell et al. studies, the figural and numerical knowledge items are stated in such a way that they appeal to verbally stored information (e.g. "what's the name of this figure?"). Furthermore, loadings on Gf and Gc generally depend on the proportion of verbal and nonverbal content in the battery (Carroll, 1993). Finally, since transmission of knowledge commonly takes place verbally (Hunt, 2000), differences in verbal comprehension may have resulted in differences in acquired knowledge. We suggest that in investment theory the acquisition of knowledge requires investment of fluid intelligence and exposure to education, and mediation by verbal comprehension. This would give an account, *inter alia*, of why numerical and figural knowledge tests often have loadings on a verbal factor. We conclude that verbal comprehension is a plausible alternative interpretation of factor Gc. However, one final phenomenon, namely the fact that under certain circumstances factors can merge together, is important, because here the interpretations of Gc as exposure to education and Gc as verbal comprehension intersect.

Table 5.1 Interpretations of Gc (see main text)

Interpretation of Gc	Appropriate Measurement model	Broadness of domain	Source	Kind	Plausibility
Crystallized intelligence (knowledge)	Formative	Transcending	-	-	-
Crystallized intelligence (Potential to perform)	Reflective	Transcending	Psychological	Cognitive	Implausible
Motivation	Reflective	Transcending	Psychological	Noncognitive	Implausible
Persistence	Reflective	Transcending	Psychological	Noncognitive	Implausible
Broad interest	Reflective	Transcending	Psychological	Noncognitive	Implausible
Exposure to education	Reflective	Transcending	Environmental	Noncognitive	Plausible
Verbal cognitive ability	Reflective	Specific	Psychological	Cognitive	Plausible

5.5.3 Interpretations intersect

As mentioned above, in culturally homogeneous samples, Gf and *g* may correlate perfectly and thus appear as one, while in heterogeneous samples, Gf and *g* correlate less than one, and thus may appear distinct (Kvist & Gustafsson, 2008). The fact that Gf and *g* can merge is thus important in understanding the mixed results of studies concerning the distinctiveness of *g* and Gf (see above). In addition, such merging is also important in the discussion and understanding of Gc as representing an underlying variable. We contend that in culturally and educationally homogeneous samples Gc and verbal comprehension merge into one factor (note that this would explain why Carroll (1993) did not always find a separate crystallized intelligence factor). In other words: Given investment theory, once differences in language, cultural influences, and education have been taken into account, individual differences in fluid intelligence (the same ability as general intelligence) and verbal comprehension can account for individual differences in crystallized intelligence.

Obviously, Gc cannot represent both exposure to information via education and verbal comprehension. To clinch the meaning of Gc, we have to scrutinize the common factors in the factor structure of a given test battery. We note that researchers do not always label CHC factors consistently. For example, consider the Woodcock-Johnson batteries, in which Gc can refer to (1) a variable that summarizes covariance among tests of verbal knowledge in the cognitive battery (McGrew & Woodcock, 2001); (2) a variable that summarizes covariance among general and academic knowledge in the achievement battery (McGrew & Woodcock, 2001); (3) a variable that summarizes covariance among all tests in the achievement battery (Horn & Blankson, 2005). In the last case, Gc is a general achievement factor, and, as a general factor, resembles the *g* factor in the cognitive battery. If this general achievement factor (Gc) is not found to be identical to the general factor in the cognitive battery (*g*), a third level *g*-Gc model seems appropriate, which can be regarded as a reintroduction of Cattell's original bipartition of *g*. If the verbal ability factor in the cognitive battery and the general achievement factor (Gc) are indistinguishable, Gc can be interpreted as verbal comprehension.

5.5.4 Conclusion

The realist interpretation of the CHC factor that has been designated Gc depends on the exact constitution of the battery. Theoretically, Gc can either represent verbal comprehension or exposure to information through education. We suspect that in practice in most cognitive batteries, the CHC factor Gc will represent verbal comprehension (Carroll, 1993).

5.6 Reanalysis of the HCA ALLI00 Dataset

To corroborate our conclusions we reanalyzed a (HCA) data set (Carroll, 1993, McGrew, 2009) based on a sample that was homogeneous with respect to age and cultural background. We chose the HCA ALLI00 data set, because educational attainment was measured, which opened the possibility to

control for educational differences. The data set was originally analyzed by Allison (1960), and reanalyzed by Snow, Kyllonen and Marshalek (1984), and by Carroll (1993). According to Carroll (1993, p. 281), the study was well designed and well conducted. A point of concern might be that the sample consists entirely of male subjects, which, in principle, may limit our conclusions to males. However, there is no clear theoretical reason to assume that investment theory only applies to men and not to women. Moreover, since measurement invariance with respect to sex may not be tenable (e.g. factor loadings may differ in male and female samples), mixed samples may introduce additional variance not ascribable to cognitive factors. We do not consider the sex limited sample as a real limitation with respect to testing our hypotheses. The sample was representative for the population otherwise.

Because the battery was designed to measure cognitive abilities (rather than educational achievement), we hypothesized that in the HCA ALLI00 data set, once educational differences have been taken into account (1) Gc and verbal comprehension are indistinct as common factors, and 2) Gf and *g* are indistinct as common factors. If so, Gc and one of the factors *g* and Gf can be removed from the model without a significant decrease in model fit.

We stress that in this chapter we limit ourselves to the test of the hypotheses that Gc and Gf (or *g*) are redundant as explanatory variables. We do not test, for example, whether the CHC model is the preferred statistical model (cf, Johnson et al, 2007), or via what pathways education influences intelligence exactly (cf, Dolan, et. al, 2001). These questions are important, but beyond the scope of the current chapter. Our present analyses are meant to serve as an example of how to test Gc is not a substantive latent variable, and as a starting point for future research into investment theory. We content that to investigate investment theory properly, one should use a longitudinal design, using same-aged, same-sex, cultural and educational homogeneous subject samples.

Table 5.2 Indicator variables (subtests) and the factors on which they are taken to load across analyses

Subtest	Carroll (1993) Factor	Allison (1960) Factor	Current analysis
Vocabulary	Verbal	Verbal Knowledge	Verbal Comprehension
Sentence Completion	Verbal	Verbal Knowledge	Verbal Comprehension
General Classification Test	Verbal	Verbal Knowledge	Verbal Comprehension
Arithmetic	Verbal	General Reasoning	Quantitative Reasoning
Division	Verbal	Number Facility	Quantitative Reasoning
Math Aptitude	Verbal	General Reasoning	Quantitative Reasoning
Number Series	Verbal	Induction	Reasoning
Letter Sets	Verbal	Induction	Reasoning
Ship Destination	Verbal	General Reasoning	Reasoning
Reasoning	Verbal	Deduction	Reasoning
False Premises	Verbal	Deduction	Verbal Comprehension
Words Association	Perceptual Speed	Speed of Association	(Verbal) Fluency
Word Checking	Perceptual Speed	Speed of Association	(Verbal) Fluency
Recognition	Perceptual Speed	Associative Memory	(Verbal) Fluency
Mechanical	Mechanical/Space	Mechanical Knowledge	Spatial ability
Mechanical Knowledge	Mechanical/Space	Mechanical Knowledge	Spatial ability
Paper Form Board	Mechanical/Space	Visualization	Spatial ability
Cards	Mechanical/Space	Spatial Relations & Orientation	Spatial ability
Paper Folding	Mechanical/Space	Visualization	Spatial ability
Cubes	Mechanical/Space	Spatial Relations & Orientation	Reasoning
Picture Discrimination	Mechanical/Space	Perceptual Speed	Spatial ability
Clerical Aptitude	Writing	Perceptual Speed	Speed
Picture Number	Association Memory	Associative Memory	Speed
Word Number	Association Memory	Associative Memory	Memory
First Names	Association Memory	Associative Memory	Memory
Addition	Numerical	Number Facility	Quantitative Reasoning

5.6.1 Method

Participants The sample consisted of 483 17-to-22-year-old (male) U.S. Naval recruits. Educational attainment ranged from 0 years in high school to 2 years in college. Mean IQ was 'slightly below average' (Allison, 1960, p. 48). The data analysis is based on the correlation matrix of cognitive ability test scores of 315 participants. The correlation matrix is available at www.iapsych.com/wmfhcaarchive/ALLI00.html.

Measures In his (re)analysis of data set ALLI00, Carroll (1993) interpreted the extracted first order factors as verbal reasoning, perceptual speed, mechanical/spatial ability, writing speed, associative memory, and numerical facilitation. In addition, he interpreted these variables as being dependent on fluid intelligence, crystallized intelligence, and broad memory. Because the CHC factor model is based largely on Carroll's analyses, we followed his analyses as closely as possible in our confirmatory factor analysis. However, we omitted the Armed Forces Qualification Test, the Otis Self-Administering Achievement Test, and the Oral Directions Test, because these are full scale IQ tests (Allison, 1960), hence they measure aggregated effects. Measures of writing speed were also omitted because we consider these to represent psychomotor abilities, which are not relevant to the present discussion. Table 5.2 includes the indicator variables.

5.6.2 Statistical analysis.

Because we omitted several subtests, we could not implement Carroll's model directly into a confirmatory factor model. Since, in general, the factor structure depends to a certain extent on the tests in the test battery, we considered it possible that the factor structure differed somewhat from Carroll's results (especially at the first order level). We therefore carried the statistical analyses out in the following five steps.

Step 1: Following Carroll, we carried out a Schmid-Leiman hierarchical factor analysis (Schmid & Leiman, 1957).

Step 2: Next, to implement the higher order factor model, we used maximum likelihood confirmatory factor analysis (CFA). Initially the factor structure was thus based on the Schmid-Leiman results. Loadings that were found to be small (< 0.15) in step 1 were omitted, but otherwise we included cross-loadings (as in Carroll, 1993).

Step 3: To evaluate the fit (Schermelleh-Engel, Moosbrugger, & Müller, 2003), we considered the ratio of the χ^2 to the degrees of freedom (χ^2/df), the Root Mean Square Error of Approximation (RMSEA), and the Non-normed Fit Index (NNFI). We considered a value of the RMSEA of 0.05 or lower and a value of χ^2/df of less than 2 as indicative of good fit. We considered an NNFI between .90 and .95 as acceptable, and above .95, as good. To obtain an acceptable model, we allowed for correlated errors in the model obtained in step 2 (but only if these correlations were open to reasonable theoretical interpretable).

Step 4: Since we hypothesized that once education differences are taken into account, Gc and verbal ability, on the one hand, and Gf and *g,* on the other, would be indistinct, we partialled out the educational variable. We did so by regressing all indicator variables on this variable in the model obtained in step 3. To evaluate our hypotheses, we examined correlations among the factors and the amount of explained variance in given factors (squared multiple correlations). If the correlation between two factors did not differ from 1, we took this to mean that we could remove one factor as being redundant in accounting for the observed covariance structure. If the variance of a given factor was fully explained by other factors (i.e., not significantly different from 100% explained variance), we interpreted the first as being nonexistent.

Step 5: Finally, we implemented the model leaving out any redundant factors.

5.6.3 Results

Step 1: On the basis of the results of the Schmid-Leiman procedure and theoretical considerations, we arrived at a model including 7 first order factors and, following Carroll's interpretation of the factors at the second stratum, 3 second-order factors. We interpreted the first order factors as inductive reasoning (Re), verbal comprehension (Ve), spatial reasoning (SA), quantitative reasoning (QR), (verbal) fluency (Fl), associative memory (AM), and speed (Sp). We denoted the second order factors fluid intelligence (Gf), a Gc factor (its interpretation yet to be evaluated), and BM, which Carroll interpreted as broad memory. Following Carroll (1993), speed (Sp) and associative memory (AM) were taken to load on the broad memory factor in the CFA. Following investment theory and Cattell's original observation that in his bipartite model fluid and crystallized tests loaded on both general factors, the other first order factors were taken to load both on Gf and on Gc. In the CFA, this leads to identification problems, which we solved by fixing the factor loadings of Re on Gc, and of Ve on Gf to zero. We consider this acceptable as these loadings are expected to be low: The tests that loaded on Reasoning can be assumed to require a minimum of prior knowledge, and the tests that loaded on Verbal comprehension were clearly knowledge tests (hence crystallized tests), so they can be assumed to require only a minimum of relation perceiving.

Step 2: We fitted the model and removed small loadings on Gc and Gf. Ultimately, Ve and Sp loaded on Gc, Re and Qu on Gf, and Fl on both Gc and Gf.

Step 3: We allowed for correlated residuals of the tests *vocabulary* and *sentence completion* (because *sentence completion* can be considered to rely on vocabulary), and correlated residuals of the tests *words association* and *word checking* (because *words association* can be assumed to involve mentally checking of words in order to prevent giving answers that have already been given).

For this model, we obtained a χ^2 of 492.05 (df = 275 , p <0.001), hence, $\chi^2/df < 2$, with RMSEA = 0.049 and NNFI = 0.94. That is, the fit was acceptable. Factor loadings are given in Table 5.3. Squared multiple correlations for this model are given in Table 5.4.

Step 4: Partialling out the educational variable resulted in a model (see Figure 5.4) in which the correlation between Gf and *g*, and the correlation between verbal comprehension and Gc were found to equal one. So statistically, Gf and *g* were indistinguishable. In addition, the correlation between the first order reasoning factor and Gf was also one. Hence, statistically, all variance in reasoning was explained by Gf (or *g*). Finally, statistically, Gc explained all variance in verbal comprehension.

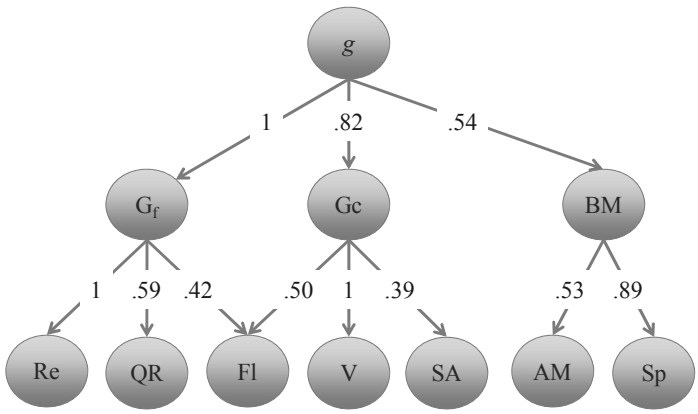


Figure 5.4 Hierarchical factor model derived from the Schmid-Leiman procedure and Confirmatory Factor Analysis (without Education regressed out).

Step 5: The redundant factors were omitted. This model (see Figure 5.5) yielded the exact same fit statistics as the hierarchical factor model.

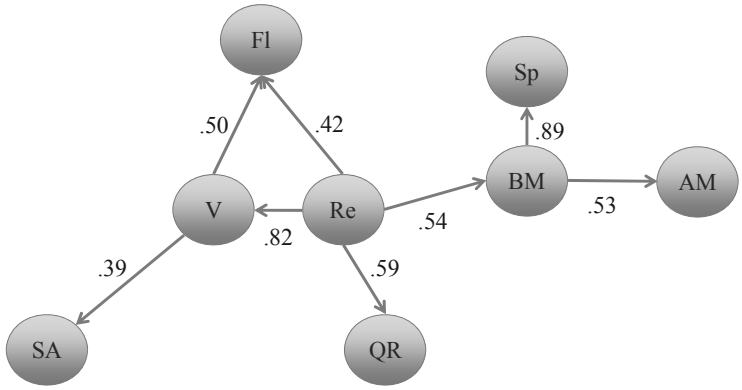


Figure 5.5 Causal model (interpretation of the factor model in Figure 5.4)

5.6.4 Conclusions

We hypothesized that in the ALLI00 data set, once educational differences were taken into account, Gc and verbal comprehension on the one hand, and Gf and *g*, on the other, would be indistinct. These hypotheses were confirmed. Because Gc was found to be redundant, we conclude that crystallized intelligence is not a unique capacity. Furthermore, Gf accounted for all the variance in the first order factor Re (reasoning). So, if the three factors *g,* Gf and Reasoning are interpreted as realistic underlying variables (e.g. as representing individual differences in psychological or biological capacities), there is no need to assume that they constitute different variables (hence capacities). We note that in the model of Step 5, verbal comprehension mediates not only (directly) the relationship between reasoning and verbal knowledge, but also (indirectly) the relationship between reasoning and spatial (mechanical) knowledge.

5.7 Discussion

The CHC factor Gc has been interpreted in different ways. On the one hand, Gc has been interpreted as a realistic, underlying cause of individual differences in intelligence subtest scores. On the other hand, Gc has been interpreted as crystallized intelligence, consistent with Cattell's investment theory of fluid and crystallized intelligence. We demonstrated that these interpretations are mutually incompatible, because in investment theory crystallized intelligence is not an underlying cause, but a statistical entity that (under certain circumstances) may emerge in factor analysis. We argued that, if one subscribes to investment theory, and to the interpretation of Gc as underlying cause of individual differences in intelligence subtest scores, two interpretations are tenable. We maintained that, depending on the exact constitution of the test battery, Gc can be interpreted as (1) exposure to information through education, or (2) verbal comprehension. We corroborated our main conclusions by presenting a reanalysis of a HCA (Human Cognitive Abilities project) data set. Our results supported the interpretation of Gc as verbal comprehension. We conclude that crystallized intelligence does not represent an underlying variable, i.e., a real, causal cognitive capacity, or any other psychological or biological capacity. Since crystallized intelligence is a statistical entity, it is better conceptualized as a formative variable (see Diamantopoulos, 2008, for example).

The objective of this chapter was to contribute to the theoretical framework of the CHC model, and to elucidate the relations among *g*, Gc and Gf in the light of investment theory. Therefore,

we took the CHC model and investment theory to be correct. However, we do not mean to say that we consider the CHC taxonomy to be the best taxonomy or that investment theory is correct in reality. We do not rule out that other taxonomies (e.g. based on Johnson's VPR model) may provide more satisfactory results. In addition, in view of the mixed support for investment theory (Ferrer & McArdle, 2004; Kvist & Gustafsson, 2008; Schweizer & Koch, 2002), we also content that investment theory is itself is open to debate.

However, despite the mixed results concerning investment theory, we note that Gf is still found to be a predictor of Gc (Ferrer & McArdle, 2004). We argue that the essence of investment theory - the hypothesis that cognitive processing (including reasoning) is involved in the acquisition of knowledge, skills, and solving strategies - remains plausible. We also note that in Cattell's investment theory, which is broader than the investment hypothesis of fluid intelligence, Gf is not the only predictor of crystallized intelligence; exposure to information through education is another. From our results, we conclude that the acquisition of knowledge as a result of reasoning may be mediated by verbal comprehension.

We believe that Cattell's investment theory still provides a viable framework to test hypotheses in intelligence research.[12] For example, Kvist & Gustafsson (2008) hypothesized that the relation between fluid intelligence and the general factor is a function of cultural background, which they consider to be evidence for investment theory. They predicted a perfect correlation between Gf and *g* in samples homogeneous with respect to culture, age, and education. In our re-analysis of the HCA ALLI00, we confirmed this prediction. To study investment theory properly, it is important to be aware of the role of sample heterogeneity. Ideally, it should be investigated in a longitudinal design using same-aged, same-sex, culturally and educationally homogeneous samples.

We contend that the explication of one's scientific/philosophical position regarding the CHC factors will facilitate the integration of empirical results, and the investigation of old (e.g. the hypotheses in investment theory) and new hypotheses. Like Gc, Grw (Reading and writing), Gq (Quantitative knowledge), and Gkn (General specialized knowledge) (McGrew, 2009) are regarded alternately as achievement clusters (McGrew & Woodcock, 2001), the results of investment of cognitive abilities (McGrew, 2009), subsets of crystallized intelligence (Snow, 1996; Woodcock, 1994), and as non-unitary variables (McGrew & Evans, 2004). We conjecture that the appropriate measurement models for these achievement variables are formative, rather than reflective.

In the spirit of Cattell (1987, pp. 365-433), we argue for the usage of different notations for domain specific, domain transcending, and domain general variables, and for different notations for constructivist variables and realist variables. The adoption of such explicit notation will benefit the CHC theoretical framework. We expect that this will further increase our understanding of empirical findings, help to generate novel hypotheses, and, ultimately, reinvigorate the study of human intelligence.

[12] Although our hypotheses, derived from investment theory, were supported by our analyses, we leave open the possibility that the theory may be incorrect. The development of intelligence may be more dynamic and complex (e.g. reciprocally interactive) than suggested by investment theory. As dynamic modeling has shown that factors may arise from such interactions (van der Maas et al., 2006), each variable in the CHC model deserves critical examination.

Table 5.3 Subtests' loadings on the first order factors in the current Confirmatory Factor Analysis.

		First order factor loadings						
	Subtest	Ve	SA	Qu	Fl	AM	Sp	Re
1	Vocabulary	**.82**						
2	Sentence Completion	**.74**						
3	General Classification Test	**.99**						
4	Arithmetic			.45				**.54**
5	Division			**.85**			.24	
6	Math Aptitude			.26				**.53**
7	Number Series							**.75**
8	Letter Sets	.29					.18	**.31**
9	Ship Destination		.21					**.76**
10	Reasoning							**.68**
11	False Premises	**.44**						
12	Words Association				**.69**			
13	Word Checking				**.65**			
14	Recognition				**.56**			
15	Mechanical	.16	**.83**					
16	Mechanical Knowledge		**.77**					
17	Paper Form Board		**.54**					.23
18	Cards		**.50**					.22
19	Paper Folding		**.39**					.39
20	Cubes							.16
21	Picture Discrimination		.34				**.53**	.19
22	Clerical Aptitude						**.77**	
23	Picture Number					**.78**		
24	Word Number					**.55**		
25	First Names				**.43**	.41		
26	Addition			**.56**			.44	

Note: Ve = Verbal comprehension, SA = Spatial ability, Qu = Quantitative Reasoning, Fl= (Verbal) Fluency, Me = Memory, Sp = Speed, Re = Reasoning. Bold faced loadings denote most salient loadings

Table 5.4 Squared multiple correlations of the higher order factors in the Confirmatory Factor Analyses.

Order	Squared multiple correlations without education partialled out						
Third	*g*						
	100.0%						
Second	Gc	Gf	BM				
	21.9%	0.0%	12.0%				
First	Verbal	Spatial	Quantitative	Fluency	Memory	Speed	Reasoning
	0.0%	50.3%	12.9%	10.9%	43.4%	0.7%	0.0%
Order	Squared multiple correlations with education partialled out						
Third	*g*						
	100,0%						
Second	Gc	Gf	BM				
	20.0%	0.0%	13.9%				
First	Verbal	Spatial	Quantitative	Fluency	Memory	Speed	Reasoning
	0.0%	46.9%	13.3%	11.1%	43.3%	0.8%	0.0%

Note: *g* = general intelligence, Gf = fluid intelligence, Gc = to be interpreted

CHAPTER 6

THE RELATIONS AMONG *g* LOADING, HERITABILITY, AND CULTURAL LOADING: DO CURRENT THEORIES OF INTELLIGENCE EXPLAIN THEM?

Abstract

A positive correlation between intelligence subtests' g *loadings and their heritability coefficients (*h^2*) has been interpreted in support of a real* g*, i.e., a unitary, largely genetically influenced (biological) variable that mediates between genes and all cognitive abilities measured by these subtests. This correlation may exist, but does not provide sufficient evidence for such substantive interpretation, because it is accompanied by an association that* g *theories do not account for: the highest heritabilities and* g *loadings are for the most culturally loaded tests (see Chapter 3). Sampling theories and reciprocal interaction theories do not explain this joint effect either, but combinations of these theories are promising.*

6.1 The Jensen Effect for Heritability

A correlation between intelligence subtests' factor loadings on the general factor of intelligence and their heritability coefficients has figured prominently in theoretical discussions concerning the scientific status of this factor. These discussions include related questions as whether the general factor (*g*, for short) is unitary or not (e.g, Petrill, 1997), whether it represents a realistic variable or is merely a statistical entity (e.g. Gray & Thompson, 2004; Rushton & Jensen 2010a), and whether specific cognitive abilities 'feed into' a general cognitive ability ('bottom up') or vice versa ('top down') (Plomin et al., 2008, p.183). According to Gray and Thompson (2004, p. 476), for instance, a correlation between *g* factor loadings and their heritability coefficients (h^2) "favours a biological over a purely statistical explanation of *g*".

Rushton and Jensen (2010a) made a similar point, and interpreted this correlation, which they dubbed a 'Jensen effect for heritability' as follows:

> "A Jensen Effect for heritability provides biological evidence for a true genetic *g*, as opposed to the mere statistical reality of *g*. It makes problematic theories of intelligence that do not include a general factor as an underlying biological variable […] such as the model proposed by Dickens and Flynn (2001), and the mutualism model by van der Maas, Dolan, Grasman, Wicherts, Huizenga, and Raijmakers (2006)." (Rushton & Jensen, 2010a, p. 213)

We take this to mean that a positive correlation between *g* loadings and heritability coefficients supports theories of intelligence that include *g* as a (single) variable that mediates between genes and all cognitive abilities measured by IQ tests. We denote these theories *g* theories. By implication, the correlation poses a problem for theories in which the general factor of intelligence is viewed as purely a statistical entity (Dickens and Flynn, 2001; van der Maas et al., 2006; Thomson, 1951; Bartholomew, Deary, & Lawn, 2009). We denote these theories alternative theories.

Indeed, it is important theoretically (but also empirically) to distinguish between, on the one hand, the general factor of intelligence as a statistical entity and, on the other hand, its interpretation as a realistic (e.g. psychological or biological) common cause of individual differences. First, because different processes can produce very similar factorial structures of intelligence tests (Anderson, 2001; Bartholomew, Deary, & Lawn, 2009; Dickens, 2008; Thomson, 1951; van der Maas et al. 2006), factor analysis in itself cannot determine whether the statistical general factor actually represents a realistic variable. Second, both *g* theories and alternative theories can explain the facts that the statistical general factor is highly heritable and has biological correlates, hence these facts do not differentiate either. In this light, the Jensen effect for heritability is regarded as a key

argument that *g* represents a real, rather than a statistical variable. Putting aside the issue whether there is empirical evidence for this effect, the question that arises immediately is *why* this effect should provide such evidence.

6.2 Does a Jensen Effect for Heritability Provide Evidence for a Real *g*?

Consider the (fully standardized) regression model displayed in Figure 6.1 (left panel). Here, intelligence subtest scores (IQ scores) are regressed on the latent (unobserved, unmeasured[13]) variable *g* (with regression weights λ_i, for $i = 1... n$); *g* in turn is regressed on latent (unobserved, unmeasured) genetic (A) and environmental (E) influences (with regression weights *h*, respectively *e*). First suppose that the more specific variables that influence IQ subtests are purely environmental (so that each h_i^2 is zero). In this situation, if the factor *g* represents a real, genetically influenced, common cause of individual differences in IQ, the subtests' regression weights on *g* (λ_i, '*g* loadings' in factor analytic terms) and the regression weights on the genetic influences (i.e., the *g* loadings multiplied by *h*: $h\lambda_i$) are collinear. In this case, the subtests' heritability coefficients will take the value of $h^2\lambda_i^2$. The rank correlation between the subtests' *g* loadings (λ_i) and their heritability coefficients ($h^2\lambda_i^2$) will be perfect (+1). Next, suppose that the more specific factors of intelligence are heritable as well, due to specific genetic variance. The correlation between *g* loadings (λ_i) and heritability coefficients (which now take the value of $h^2\lambda_i^2 + [1- \lambda_i^2]h_i^2$) will be attenuated. However, we may still expect a positive correlation if the common genetic variance, h^2, is relatively large compared to the subtests' specific genetic variances, h_i^2.

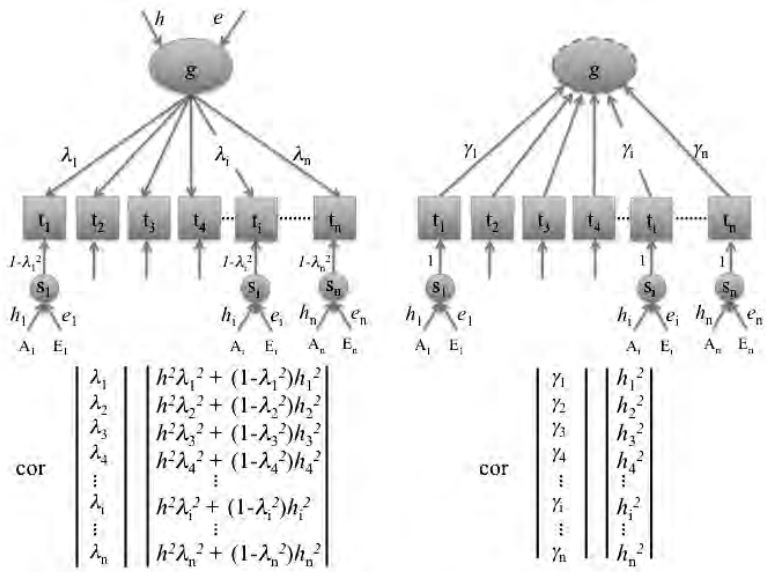


Figure 6.1 If *g* is a realistic, genetically influenced variable (left panel) and the only genetic influence on the subtests, the ranks of the subtests' factor loadings and heritability coefficients are collinear, because both vectors only depend on *l*. The rank correlation is then +1. If specific are also heritable this correlation will be attenuated. If *g* is merely a statistical summary of heritable indicators (right panel), a relation between the subtests' regression weights and heritability coefficients is not expected, unless additional assumptions are made.

[13] *g* is a higher order factor. IQ tests *estimate g* to a certain extent, because the cognitive abilities measured by IQ tests (on the zero or first order level) are influenced by *g*. The tests do not *measure g* directly.

Now suppose that the common factor is merely a summary of diverse heritable cognitive abilities and does not exist in reality. For instance, suppose that *g* is a latent variable of the formative kind (see Figure 6.1, right panel; Bollen & Lennox, 1991; Borsboom; Mellenberg, & van Heerden, 2003; Diamantopoulos, Riefler, & Roth, 2008), which means that individual differences in cognitive abilities (as measured by IQ tests) determine individual differences in the statistical variable general intelligence rather than vice versa. In this case, unless we make additional assumptions, there is no reason to expect an association between heritability coefficients (h_i^2) and summation weights ('*g* loadings', g_i). The claim that the Jensen effect for heritability provides evidence for a realistic *g* is thus derived from the argument that "what is observed is what would have been expected if an underlying *g* did in fact exist" (Rushton & Jensen, 2010a. p. 214, see Plomin et al., 2008, p.185, for a similar argument, and Bartholomew, 2004, p. 73 for the logic of inferences concerning *g*).

To test whether a positive correlation between *g* loading and heritability coefficient is significant does not provide the means to statistically test whether the relationships between genes and IQ scores are actually mediated by *g* (Dolan & Hamaker, 2001, see also Appendix B). First, it might be the case that general genetic influences and general environmental influences follow independent pathways (e.g. Plomin, 2003). Second, the possibility that the Jensen effect for heritability is due to some other process than a general mechanism (see, for example, van der Maas et al., 2006) cannot be ruled out. Putting aside this issue of testing, we agree that if a correlation between heritabilities and *g* loadings exists, this correlation requires explanation. Will any theory of general intelligence be substantive, it must be able to account for it. In addition, such theory must not conflict with other empirical data.

Our research has shown that - next to the Jensen-effect - there is a relationship between heritabilities and cultural load, and one between cultural load and *g* loading (see Chapter 3). In this chapter we evaluate whether, how, and to what extend the joint effects, which we proposed to call the Cattell-Jensen effect, poses problems for theories of intelligence. First, we evaluate *g* theories. We split these up in complexity theory and investment theories. Next, we consider alternative theories, which we further categorize into sampling theories and reciprocal interaction theories.

6.3 The Relations Among *g* Loading, Heritability, and Cultural Influences in *g* Theories

Concerning the explanatory status of common factors and *g* loadings, Jensen rightly stated:

> "Factors, including *g*, are not themselves explanatory constructs. They are constructs which themselves require explanation." (Jensen, 1987, p. 95)

> "The salient characteristics of the most highly *g*-loaded tests are not essential or definitional, but are empirical phenomena in need of theoretical explanation in their own right" (Jensen, 1998, p.92).

Nonetheless, *g* theorists are rather unspecific about the exact nature of *g* (for a critique, see Demetriou, 2002; Ackerman & Lohman, 2003), and how *g* affects test scores. For example, although Rushton & Jensen (2010a) maintain that *g* is an underlying biological variable, they do not indicate what this biological variable *g* represents. Yet, *g* theorists do have offered accounts for the finding that some tests have high *g* loadings (or, as it is sometimes expressed, are 'highly *g* loaded'). We can categorize these accounts as follows. Either they are unspecific about the role of *g* in the acquisition of cultural knowledge (e.g. Jensen, 1973; 1987), or they involve an account in terms of, or similar to, Cattell's investment theory (Carroll, 1993, p. 658; Cattell, 1987; Jensen, 1998, 2001). We denote the latter investment theories, and the former complexity theory.

6.3.1 Complexity theory

Complexity theory offers the following account for the relation between *g* and *g* loadings:

> "Probably the most undisputed fact about *g* is that the *g* loadings of cognitive tasks are an increasing monotonic function of the perceived complexity of the task." (Jensen, 1987, p. 111).

So, one may infer a relation between complexity and heritabilities of subtests (see also Beaujean, 2005; Plomin et al. 2008, p.183): on the one hand *g* loadings are indicative of complexity; on the other hand they are predictive of heritability (see Figure 6.2). Hence, we expect most complex tasks (e.g. Raven's Progressive Matrices, the so-called the prime 'marker of *g*'; see Jensen, 1998, p. 38) to be most *g* loaded and most heritable, and the culturally loaded knowledge tests (such as information and vocabulary, see Chapter 3) to be less *g* loaded (and given the Jensen-Effect for heritability, less heritable) because *g* theorists do not regard them as complex:

> "Information tests consisting of questions like 'Who was the first President of the United States?', 'Whose picture is on a penny?', and so on, make poor test items mainly for two reasons: (a) they do not get at complex mental process, and (b) they cannot be steeply graded in difficulty level without introducing items of information to which there is a relatively low probability of exposure, in which case social status and educational differences become practically impossible to avoid. The same holds true for vocabulary tests […] The difficulty levels differ only because of frequency of exposure. Such items based on information and vocabulary are rightly regarded as more culturally loaded than items which vary in difficulty because of the complexity of the mental processes involved." (Jensen, 1973, p. 184)

Indeed, in his older work Jensen predicted that culture loaded tests show low heritabilities and culture reduced tests high heritabilities (Jensen, 1973, p.194-195; see also Rushton & Jensen, 2005; Rushton, 1995). As is clear from Chapter 3, the empirical evidence does not support these predictions: the largest heritabilities (and *g* loadings) are for the most culturally loaded, noncomplex, knowledge tests. We conclude that complexity theory does not explain the observed relations among *g* loading, cultural load, and heritability.

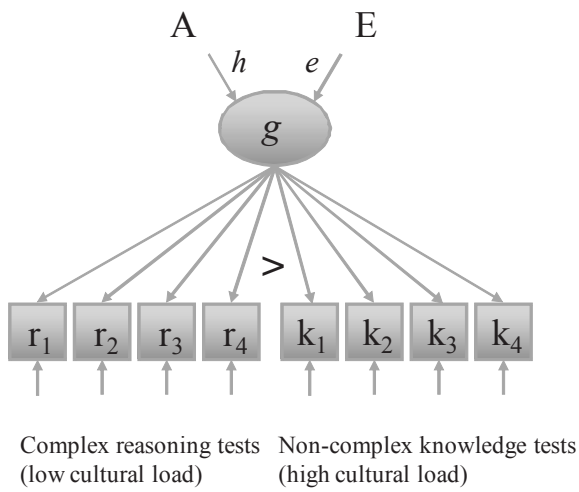


Figure 6.2 In complexity theory, the most culturally loaded tests (i.e. knowledge tests) have low *g* loadings and low heritabilities, because they are not complex.

6.3.2 Investment theories

In investment theories, *g* and *g* loadings are still related to complexity, but additionally, assumptions are made about the role of *g* in the acquisition of knowledge. Solving the items on a knowledge tests is not complex, but the acquisition of the required knowledge is.

Investment theories are derived from or inspired by Cattell's theory of fluid and crystallized intelligence (e.g. Cattell, 1987). Key in this theory is that knowledge is acquired through exposure to information, and that compared to people with low levels of *g*, people with high levels require less exposure to acquire the same amount of knowledge. Stated otherwise, compared to less intelligent people, more intelligent people are able to acquire more knowledge in the same amount of time.

In Cattell's original theory, fluid intelligence is a realistic variable, or, in his words, a 'source trait'. It is a single reasoning capacity and described as follows:

> "Fluid intelligence [..] is an expression of the level of complexity of relationships which an individual can perceive and act upon when he does not have recourse to answers to such complex issues already stored in memory" (Cattell, 1987, p. 96).

Fluid intelligence (abbreviated gf or Gf) is considered to be related to the maturation of the brain and to be highly genetically influenced. Crystallized intelligence (abbreviated gc or Gc), in contrast, is a not a realistic variable (Cattell, 1987; Kan, Kievit, Wicherts, Dolan & van der Maas, 2011; see Chapter 5), but rather a summary statistic, or, in Cattell's words, a 'surface trait', best conceptualized as a formative variable (Bollen & Lennox, 1991; Diamantopoulos et al., 2008). Crystallized intelligence is the result of 'investment' of fluid intelligence during one's learning experiences throughout the lifespan.

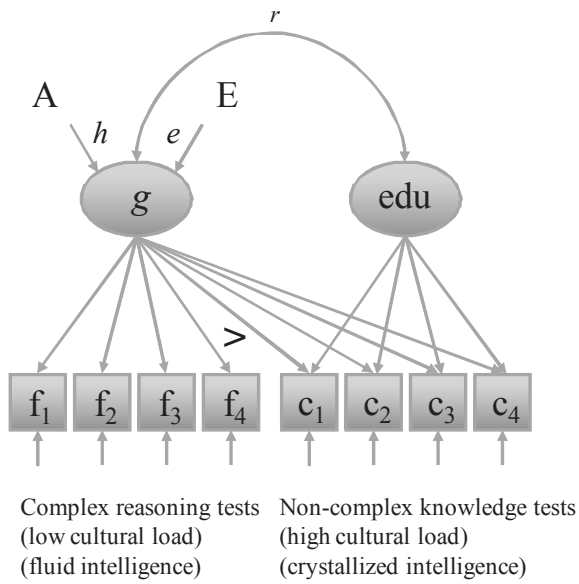


Figure 6.3 In investment theories, the culturally loaded knowledge tests (crystallized tests) may have large *g* loadings, but should show smaller heritability estimates than fluid tests, because of the intrusion of environmental influences.

In his later work, Jensen seems to subscribe to investment theory in stating that "the psychological distinction that Cattell makes between [fluid and crystallized intelligence] are valid" (Jensen, 1998, p. 124). However, because Jensen equates gf with *g* (Jensen, 1998, p.124), he considers crystallized intelligence to be the result of investment of *g* (see Carroll, 1993, p. 658, for a similar point of view). In other words, one may read investment theory as a theory that includes a general factor as a realistic, genetically influenced common cause of variance, that is, as a *g* theory (Rushton & Jensen, 2010a).

From investment theories it follows that people who possess higher degrees of fluid intelligence (or alternatively *g*) tend to acquire more knowledge and skills in all kinds of domains, which are the crystallized abilities. Crystallized intelligence is the total of the crystallized abilities. According to Cattell (1987), one might expect lower loadings for crystallized ability tests (crystallized tests, for short) than fluid intelligence tests (fluid tests, for short), because of "the intrusion of intermediate influences in the specific learnings" (p. 140). However, as he continued, a second common influence, assumed to be exposure to information via education, may render the loadings of crystallized tests on the *g* factor as large as the loadings of fluid tests (see Figure 6.3).[14]. In investment theories crystallized tests can thus have the largest *g* loadings. What about their heritabilities?

Cattell believed that

> "the population heritability is noticeably lower for crystallized intelligence than fluid intelligence, which is precisely what we should expect from the investment theory [...]". (Cattell, 1987, p. 324)

He expected this "on the hypothesis that gf is physiologically determined whereas gc is a product of environmentally varying, experientially determined investments of gf." (Cattell, 1963, pp. 3-4, see Carroll, 1993, p.374, for a similar point of view).

Exposure to information through education itself is clearly an environmental factor (but may be influenced by genetic factors as a result of gene-environment interplay, e.g. see Scarr & McCartney, 1985; Chapter 7). So, crystallized abilities (e.g. vocabulary and general information) may have large *g* loadings in investment theories, but, compared to the markers of *g* or gf (e.g. Raven's progressive matrices), they should have lower heritabilities, because of the intrusion of environmental influences (see Figure 6.3). As shown in Chapter 3, Cattell's expectation is not supported by empirical evidence.

Although researchers, including Jensen (1998), have mentioned the high heritability of crystallized intelligence, it is unclear whether this observation should be taken as evidence against investment theories, or as an erroneous expectation of Cattell (Mackintosh, 1998). In any case, no investment theory (Carroll, 1993; Cattell, 1987; Jensen; 1998) predicts that the most cultural influenced tests should be subject to the highest genetic influences. We conclude that invest theories do not explain the observed relation between *g* loading, cultural load, and heritability.

6.3.3 Conclusion

Both complexity theory and investment theories predict that heritabilities of the culturally loaded knowledge tests (crystallized tests) are lower (or at least not higher) than the heritabilities of the most complex processing tests (complex fluid tests). As is clear from Chapter 3, these predictions are not supported by empirical evidence. This is not to say that future *g* theories cannot account for the relations among *g* loading, cultural influences, and heritabilities at all or in principle, but that the existing *g* theories do not fully account for them. These theories are likely to require additional assumptions.

As mentioned above, a positive correlation between *g* loadings and heritabilities has been considered to be problematic for alternative theories of general intelligence (Rushton & Jensen,

[14] We note that the literature gives another reason for crystallized tests having larger loadings on the first principal factor than fluid tests, namely the composition of the test battery. Crystallized tests tend to have large loadings on this factor if the battery contains relatively many crystallized tests (Ashton & Lee, 2005).

2010a). Indeed, if *g* is a formative variable (Bollen & Lennox, 1991), and unless additional assumptions are made, there is no reason to expect this correlation (see above). At first glance *g* theories seemed thus more parsimonious, but, in view of the above, in order to explain the relations among *g* loading, heritability, and cultural influences jointly, *g* theories are likely to require additional assumptions as well. One question is thus how plausible the assumptions in theories of general intelligence are. In the next sections we consider the assumptions in alternative theories of general intelligence.

6.4 The Relations Among *g* Loading, Heritability, and Cultural Influences in Other Theories

We divide alternative theories in sampling theories and reciprocal interaction theories. In sampling theories *g* is essentially due to a measurement problem. If we could measure the underlying variables of intelligence independently, the positive manifold would disappear, and a *g* factor would not be present. In reciprocal interaction theories the *g* factor would remain.

6.4.1 Sampling theories

As early as 1903, hence before the publication of the original *g* theory (Spearman, 1904), Thorndike had observed positive, but imperfect correlations among cognitive tasks. He explained this observation as follows: Whenever people aim to solve a problem, they 'tap' from a pool of learned stimulus-response associations, which are hardwired as neural connections. Thorndike called these connections 'bonds'. Tests tap (sample) from some of the same bonds, so across tests there will be an overlap in the tapping. This overlap causes the positive intercorrelations among test scores.

Based on this principle, Thomson (1919; 1927; 1951) developed an alternative to *g* theory, which is now called sampling theory (Jensen, 1998) or the bonds model of intelligence (Bartholomew et al, 2009). Like Thorndike, Thomson (1919) accounted for the positive manifold (hence the statistical general factor) by assuming a simultaneous sampling from (statistically) independent variables. In Thomson's theorizing, and elaborations of his bonds model (e.g. Bartholomew, Deary, & Lawn, 2009; Thomson, 1951) (henceforth sampling theories), the brain consists of a (large) number of bonds. These bonds are called upon when completing an IQ test or item. The nature of the bonds usually remains implicit, but their underpinnings are usually sought in neural terms (Bartholomew et al., 2009; Jensen, 1998).

We note a certain shifts in the bonds model, which is not without consequences for presenting sampling theories as a true alternative for *g* theories. At one point in Thomson's theorizing, the sampled variables represented (statistically) independent group and specific factors of intelligence (Thomson, 1919). The subtests determine which variables (i.e., factors of intelligence) are called upon. We will refer to this kind of sampling as 'test sampling'. In other models (Thomson, 1951; Bartholomew et al. 2009), the sampled variables are not interpreted as the factors of intelligence, but as their underlying 'elements'. Here it is not assumed that *tests* sample from the factors of intelligence, rather it is assumed that *people* sample from the underlying elements when they attempt to solve a test item. We will refer to this type of sampling as modern sampling theory.

Recently, alternative theories of general intelligence have been developed that can be regarded as sampling theories in which the underlying elements represent genes or genetic mutations (Anderson, 2001; Penke et al., 2007). In this case, tests do not sample from cognitive abilities or capacities (factors of intelligence), nor do people sample from the underlying elements when attempting to solve an item, rather *cognitive capacities* sample from the genes. We refer to these theories as genetic sampling theories.

Next we examine how sampling is related to the general factor of intelligence and *g* loading. Furthermore, we ask ourselves whether sampling theories explain the relations among *g* loading, heritability, and cultural influences.

Test sampling In test sampling theory, the sampled variables represent independent factors of intelligence. Here, 'independent' means statistically independent across members of the population. This implies that the number of variables is essentially the same for each member of the population in question. The variables may represent elementary cognitive processes (Jensen, 1998), for example. Each cognitive tests measures a number (sample) of these variables (e.g. elementary processes). Factor loadings represent the (mean) fraction of overlap among group factors. In test sampling theory the positive manifold is thus essentially a measurement problem, because tests are multi-dimensional. If we would and could construct unidimensional tests (measuring each factor, e.g. elementary process, apart) the positive manifold would disappear, hence the *g* factor as well.

Test sampling theory does not give an explanation for a relation between factor loadings and heritability. Nevertheless, under certain specific assumptions the effect can be present. It will be present when it is assumed that the most heritable factors are called upon the most frequently, for example. We conclude that the theory does not give an account for the Jensen effect, but also that it does not conflict with it. Test sampling theory is silent about knowledge acquisition and the role of culture.

Modern sampling In modern sampling theory the independent variables (bonds) are interpreted as the underlying elements. People may differ in the *values* of these elements as well as in the *number* of elements (Bartholomew et al., 2009). As mentioned, in modern sampling tests do not sample from group factors, rather people sample from the elements when attempting a test item. Some items call upon a small number of the elements, whereas other items call upon many of them.

In some versions of modern sampling theory, one may encounter difficulties in interpreting them as true alternatives of *g*. These difficulties concern the level of analysis (see also Burt 1940). First, in interpreting bonds at the level of neurons or synapses (e.g. Thomson, 1951) it is hard to conceive of *particular* neurons showing *inter-individual* differences. It implies that neurons can be labeled (e.g. Neuron A, Neuron B, etc.) and identified in each member of the population (or subpopulation); across people there is a distribution of values of a certain property of Neuron A, a distribution of values of a certain property of Neuron B, etc., and these values are statistically independent. Second, if the properties are all of the same kind and their values are additive (e.g. when they represent mass, or amount of myelin), we actually have a single, quantitative inter-individual variable (the total or mean value of this property, for example neuronal mass). It is thus important to state what the elements or bonds represent, whether they are identifiable across individuals, and if their values can be truly added, or are merely regressed on them.

In modern sampling theory, factor loadings represent the (mean) fraction of elements used in the test (Bartholomew et al. 2009, p. 576). In our interpretation of certain bonds models, the factors loadings of subtest i (λ_i) on the estimated variable k (hence λ_{ik}) just need a subscript for the individual, i, (hence λ_{ijk}), and perhaps even for the item, l, (λ_{ijkl}). Intelligence tests scores are eventually manifestations of variable k, for instance a capacity (e.g. neuronal mass) that constrains working memory.

Like test sampling theory, modern sampling theories do not give an account for the Jensen-effect. Neither do they make explicit the role of environment or the impact of cultural influences.

Genetic sampling Multiple genes can affect one phenotype. This is called polygeneity. In addition, a single gene can influence multiple phenotypes. This is called pleiotropy. Given polygenetic and pleiotropic influences on diverse cognitive capacities are present, these capacities can share genetic influences. It is thus possible that genetic intercorrelations between cognitive processes are present. This can be due to genes that have a general effect (Kovas & Plomin, 2006), for example. The intercorrelations on the genetic level will give rise to intercorrelations on the phenotypic level. When general genetic effects are present, that is, when certain genes influence all cognitive processes, one can conceive of the general factor as representing a genetic factor: the total of general genetic effects.

The presence of a statistical general genetic factor does not necessarily mean that general effects are present in reality. A genetic factor can be the result of what we call genetic sampling, which we define as follows: *genetic sampling means that any two cognitive processes always share*

some of their genetic determinants (genes or genetic mutations), but there are no genes that influence all cognitive processes. In genetic sampling theories (Anderson, 2001; Penke et al., 2007) the sampled elements thus represent genes. Here, it is easier to conceive of the elements showing individual differences than in certain interpretations of modern sampling theories. Genes can be labeled (e.g. gene A, gene B, etc.), and in principle they are identifiable in each member of the population (or subpopulation). Across people, there is a distribution of values of gene A, a distribution of values of gene B, etc.; in genetic sampling theories these values are statistically independent. In quantitative genetics, genetic values are additive, so also in genetic sampling theory we (can) have a single, quantitative interindividual variable again (e.g. the total genetic value of 'genes that influence working memory capacity'). Whether genetic sampling is a true alternative theory is thus debatable.

The presence of genetic sampling implies that if cognitive abilities are (solely) intercorrelated due to genetic effects, the rank ordering of people on Full scale IQ will be an estimation of the rank ordering on the total genetic value of the genes that influence these cognitive abilities. Across individuals, this total genetic value predicts the levels of the cognitive abilities to a certain degree, some better than others. Factor loadings will represent the mean fraction of genes that influence a cognitive ability. Given that mutations are generally harmful, and lower genetic values, one way to conceive *g* is 'mutation load'. The more genetic mutations one carries, the lower one's genetic values, which probably means that cognitive development is disrupted to higher degrees.

Unless the pathways of genetic and environmental influenced are modeled explicitly, the relation between heritability coefficients and *g* loadings (loadings on the phenotypic psychometric *g*) is *a priori* unknown (see also Chapter 4). Genetic sampling theories are silent about the role of culture.

Conclusion In test sampling theories and modern sampling theories, factor loadings represent the (mean) fraction of bonds used in the test (Bartholomew et al., 2009). Differences in these loadings are usually explained from the tests or items' levels of complexity (Jensen, 1998): Complex items will call upon many of the bonds, hence show larger overlaps (hence larger intercorrelations, hence larger *g* loadings) than noncomplex items. If the bonds' values are heritable, sampling theories resemble Jensen's complexity theory, which, as we have shown, cannot explain the mutual relationships among heritability, *g* loading, and cultural load. There is only one important difference. If *g* loadings represent fractions of statistically independent, heritable bonds, there will be no relation between subtests *g* loadings and heritabilities. This is because a sum of two statistically independent variables will have a heritability that falls somewhere in between the heritability of the two variables. A Jensen effect might be present however, namely if the most heritable bonds are called upon the most frequently. We conclude that sampling theories can incorporate a correlation between heritabilities and *g* loadings, by making additional assumptions, but that they do not give an account for it.

Sampling may occur at every level, so genetic sampling, test sampling and modern sampling can be combined. Sampling does not account for a correlation between *g* loading and cultural load. This is because sampling theories do not make explicit the role of cultural and environmental influences that affect intelligence during its development. However, we do believe that sampling may play a role in explaining the Jensen effect (see below, and Chapter 7).

6.4.2 Reciprocal interaction theories

In reciprocal interaction theories (Dickens & Flynn, 2001, Dickens, 2008; van der Maas et. al, 2006), the statistical *g* factor is the result of mutual beneficial interactions between cognitive processes or abilities. These theories do not include *g* as a realistic latent variable. The general factor is not a measurement problem, as in sampling theories.

The mutualism theory In the mutualism theory of van der Maas et al. (2006), general intelligence is assumed to be the result of mutual, largely beneficial, interactions among basic cognitive processes (such as perceptual, memory, and reasoning processes), which occur throughout development (see Figure 6.4). One assumption is that the growth of each cognitive process depends on limited resources, which are conceptualized in terms of biological constraints, such as neuronal speed and the

size of neural systems associated with each of the cognitive processes. Another assumption is that the development of cognitive processes is largely an autonomous, self-regulating process. By incorporating these assumptions in a multivariate dynamical systems model (henceforth, the mutualism model), it was shown that mutual, largely beneficial, interactions among variables can yield a statistical common factor.

The mutualism model is formulated mathematically as follows:

$$\frac{dx_i}{dt} = a_i x_i (1 - x_i / K_i) + a_i \sum_{\substack{j=1, \\ j \neq i}}^{W} M_{ij} x_j x_i / K_i \qquad \text{for } i, j = 1 \ldots W.$$

The K's represent W limited resources of the growth processes of an individual system. Parameters a are growth parameters, determining the steepness of the (logistic) growth function associated with each variable x. Weights M_{ij} determines the influence of x_i on x_j. In the population, the parameters a and K are considered to be parameters that differ over individuals, whereas the weights $M_{ij}$are considered to be equal for all individuals (i.e., they are population parameters, like factor loadings in factor models of intelligence).

When all individual systems are equilibrium, the covariance matrix of the x's equals:

$$\Sigma = [\mathbf{I} - \mathbf{M}]^{-1} \Psi [\mathbf{I} - \mathbf{M}]^{-\mathrm{T}}$$

where $\mathbf{\Psi}$ is a diagonal matrix containing the variances of the K's, and $\mathbf{I}$ is the identity matrix. Superscript -1 denotes matrix inversion, and superscript -T denotes inversion and transposition.

In simulating the development of general intelligence, van der Maas et al. (2006) modeled all underlying variables (a's, K's, and starting values) as initially uncorrelated. During the development of individual systems the values of the observed variables (x's) increased while becoming intercorrelated due to their mutual beneficial interactions. Once the systems were all in equilibrium, a factor analysis on these variables yielded a 'general factor'. This factor is a purely statistical entity, however, as it does not represent any common, underlying psychological or biological variable; it is simply a summary index of the positive correlations among the observed variables. When genetic and environmental influences were introduced in the model (via the limited resources, K), the general factor was found to be heritable, thus even when all genetic and environmental influences on the limited resources were assumed to be uncorrelated. Notably, the common factor appeared *more* genetically influenced (i.e., had a higher heritability) than the limited resources themselves.

These simulations did not produce the Jensen effect, but following the introduction of very small positive correlations among the genetic influences on the limited resources K, the Jensen effect was observed (van der Maas et al., 2006). Using the equation for the covariance matrix, van der Maas et al. simulated data for 16 K's with normally distributed M_{ij} (mean = .05). Heritabilities of the K's were normally distributed with mean .5 and SD .02). Their genetic intercorrelations were assumed to be equal (SD = 0), whereas their environmental intercorrelations were fixed at zero. Generally, the correlations between factor loadings on the first principal factor and observed heritabilities were positive (see Figure 9 in van der Maas et al.). The variance in **M** was introduced to investigate the robustness of the Jensen effect, and is not a necessary or essential aspect in this setup. The essential assumption was the presence of nonzero genetic intercorrelations. If the interaction weights in **M** are uniform and the K's are heritable to different degrees, the introduction of any nonzero genetic intercorrelation, no matter how small, will yield a vector correlation between heritabilities and g loadings of 1.

In the simulations of van der Maas et al. (2006) only very weak genetic intercorrelations were thus required. Here, we interpret these weak intercorrelations in terms of genetic sampling (Anderson, 2001; Penke et al., 2007), which we took to mean that any two cognitive processes always share some of their genetic determinants, but that there are no genes that influence all cognitive processes (genetic sampling may be interpreted as a weak form of pleiotropy). We conclude that a correlation between g loadings and heritabilities can be reasonably explained without including a

general, largely genetically influenced, mediating variable or the assumption of general genetic effects. We stress we do not argue for or against the presence of general genetic effects in reality. Rather, we note that a correlation between heritabilities and *g* loadings does not necessarily imply that general effects are present.

Mutualism provides a plausible explanation of the positive manifold. As shown in van der Maas et al. (2006), mutualism can also account for the hierarchical factor structure of intelligence (e.g. Carroll, 1993) and developmental effects, such as integration and differentiation effects (Deary et al., 1996). In addition, it can account for the (high) heritability of *g* (e.g. Plomin et al. 2008), and the increase of heritability of *g* during the lifespan (Bartels et al., 2002; Haworth et al., 2009; Plomin et al., 2008). It is also consistent with the finding that more specific factors of intelligence (spatial ability, verbal ability, memory, and processing speed) are genetically independent initially, but become increasingly genetically interrelated as development unfolds (Hoekstra, Bartels & Boomsma, 2007). Such developmental effects are not readily explained by *g* theories or sampling theories. We conclude that mutualism explains more empirical findings (including behavior genetic results) within intelligence research than *g* theories and sampling theories.

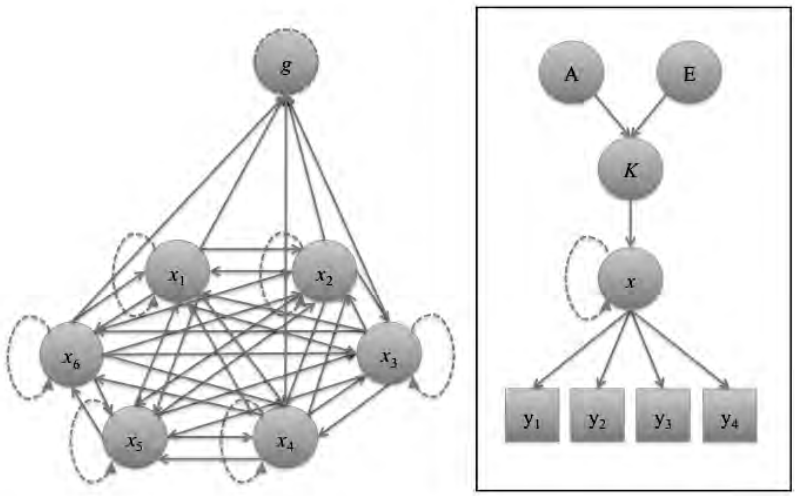


Figure 6.4 The mutualism model. The *g* factor is the result of mutual beneficial interactions among cognitive processes (*x*'s) during development. The growth of these processes is constrained by limited resources (*K*), which are influenced by genetic (A) and environmental variables (E). Variables *y* are indicators of individual differences in *x*.

Mathematically, the mutualism model bears strong similarities with Dickens' model (Dickens, 2008; unpublished). In fact, the latter can be regarded as a mutualism model in which the role of the environment is made explicit.

Dickens' model Though not published in a peer-reviewed journal, we acknowledge Dickens' model (Dickens, 2008; unpublished) as an important alternative theory of general intelligence, and discuss it here, because Dickens provides another account for the Jensen effect without assuming a realistic *g*. In addition, we believe that Dickens (unpublished, p. 7) gives an important perspective on the relations among *g* loading, heritability, and cultural influences by suggesting that "those skills that are emphasized in practice will have a tendency to be the ones that will have the highest heritability." (see also Chapter 7).

Dickens' model extends the theory of Dickens & Flynn (2001), but the basic assumption is the same, namely that people who possess higher levels of any cognitive ability are more likely to end up in environments conducive to the development of their cognitive abilities. In Dickens' model it is assumed that people differ in cognitive abilities as a result of both environmental and genetic

differences. More specifically, individual differences in cognitive abilities are assumed to be heritable, but not due to a common dependence on a single underlying biological, heritable variable. In essence, Dickens' model is a model of (evocative, passive and active) gene-environment correlation (Scarr & McCartney, 1983). Current levels of ability are the result of a feedback process in which environmental influences are influenced by genetic influences (via the abilities), whereas environmental demands cause the heritable cognitive abilities to become correlated (see Figure 6.5).

Dickens was able to account for salient findings of behavior genetics and intelligence research, including the Jensen effect for heritability, by making a few additional assumptions (Dickens, unpublished, p. 9). The most relevant assumptions are that (1) abilities that are valued the most important within the cognitive environment will be practiced most often in that environment, and (2) people who excel in the most valued abilities will have the greatest chance to end up in cognitive demanding environments. By incorporating these assumptions into his model, Dickens demonstrated that the most valued abilities became the most highly *g* loaded.

Dickens' (2008; unpublished) account of the Jensen effect differs from that of van der Maas et al. (2006). A Jensen effect for heritability appeared in Dickens simulations by assuming that cognitive abilities that are the least important are more ubiquitous in non-cognitive activities, and, as a result, are more open to environmental influences outside the cognitive environmental system. We believe that this assumption is plausible, but future research is necessary to establish whether this is indeed the case.

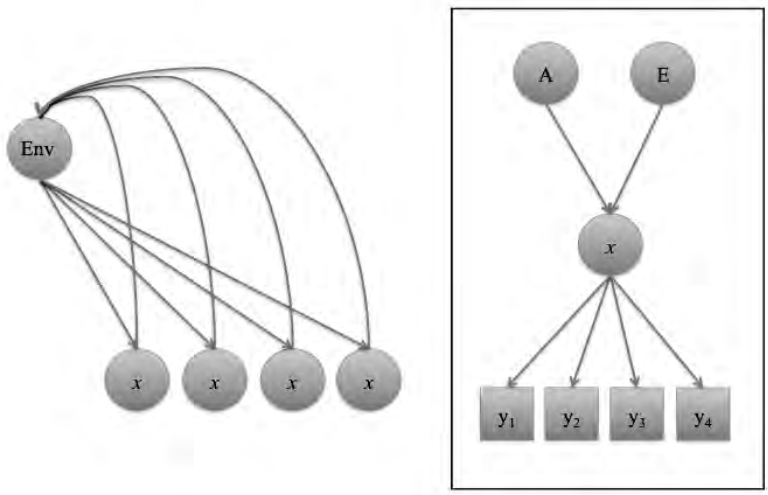


Figure 6.5 Dickens' model (based on Dickens & Flynn, 2001). Environment is influenced by heritable cognitive abilities, whereas environmental demands cause the abilities to become correlated.

Conclusion As shown in Chapter 3, the most culturally loaded tests (generally verbal knowledge tests) have the highest *g* loadings. It is conceivable that in society in general, and in educational settings in particular, verbal skills and knowledge are most highly valued, and that these are subject to more training and exercise than, say, perceptual speed. Following Dickens' model, the most highly valued skills (verbal skills and knowledge) will show the highest *g* loadings, and, given the Jensen effect, also the highest heritabilities.

As mentioned, Dickens' model can be accommodated within the mutualism model by making the role of the environment explicit, i.e., by introducing mediation by the environment of the mutual interactions among the cognitive processes. Therefore, we believe that mutualism, in combination with genetic sampling, is promising in accounting for the relations among *g* loading, heritability, and cultural load when incorporating assumptions from the theory of Dickens & Flynn (2001) and Dickens' (2008; unpublished) model.

6.5 Discussion

In this chapter, we reviewed the issue whether the relation between *g* loadings and heritabilities supports the interpretation of *g* as a realistic (biological), largely genetic, mediating variable (e.g. Gray & Thompson, 2004; Rushton & Jensen, 2010a). We subscribe to Rushton and Jensen's statement:

> "[T]here is no absolute claim that *g* effects have been proven; only that what is observed is what would have been expected if an underlying *g* did in fact exist." (Rushton & Jensen, 2010a, p.214)

Indeed, if a real, largely genetic influenced, mediating variable is present, we would expect a Jensen effect for heritability. This effect may exist, but in view of the following, the Jensen effect for heritability effect is not convincing evidence for a realistic, biological *g* as Rushton and Jensen put forward.

First, in *g* theories theory large genetic influences and large cultural influences are juxtaposed, but cultural load and heritability are not mutually exclusive. Indeed, empirical evidence has shown that that large genetic and large cultural influences on intelligence can go together: Chapter 3 revealed that the highest heritability coefficients and *g* loadings are of the most culturally loaded (knowledge) tests. Current *g* theories do not account for this joint effect. Sampling theories (Anderson, 2001; Bartholomew et al., 2009; Penke et al. 2007; Thomson, 1951) do not account for it either, because they do not make explicit the role of environment in the development of intelligence. Furthermore, some interpretations of certain sampling theories (Bartholomew et al., 2009) must be regarded as *g* theories, in which factor loadings are person specific. Reciprocal interaction theories (Dickens, 2008; van der Maas et al., 2006) in isolation do not account for the relations among *g* loading, heritability, and cultural load either (but see below and Chapter 7).

We believe that integration of (elements from) alternative theories provide a promising perspective. First, the mutualism theory (van der Maas et al., 2006) accounts for developmental effects that are not accounted for by other theories of general intelligence. Second, assuming genetic sampling (Anderson, 2001; Penke et al. 2007), a Jensen effect for heritability is present in the mutualism model. Third, Dickens's (2008; unpublished) model, can be incorporated in the mutualism model, which can explain that the most practiced abilities are the most *g* loaded. Fourth, elements of the investment theory of fluid and crystallized intelligence can be incorporated. We believe that mutualism can account for the joint effect if it is assumed that the cognitive environment puts demands primarily on crystallized abilities.

We also contend that the general factor of intelligence does not represent a realistic (biological) common cause of individual differences, but is a merely a statistical entity, best conceptualized as an index measure, or formative variable (Bollen & Lennox, 1991; Borsboom et al., 2003), i.e., an informative summary descriptive. However, as such, general intelligence certainly has utility. Specifically, we consider it to be similar to 'general health'. Like general health is an informative summary descriptive of physical functioning, '*g*' is an informative summary descriptive of cognitive functioning.

In the introduction of this chapter we mentioned that factor analysis does not discriminate between *g* theories and alternative theories, and that both *g* theories and alternative theories can explain the facts that the general intelligence is (highly) heritable and has (strong) biological correlates. As shown, the Jensen effect for heritability does not help discriminating among theories either. In order to explain the relations among *g* loading, heritability, and cultural load, we need better theory. In order to discriminate among these theories we need true differential predictions.

CHAPTER 7

DISCUSSION

Abstract

Knowledge tests display higher g *loadings and heritability coefficients than cognitive processing tests. In isolation, mainstream theories of intelligence fail to explain this effect. To account for it, we present an integrated model of intelligence. We propose the following. The development of cognitive processes is initially largely autonomous and self-regulating, but benefits from knowledge acquisition (learning), whereas knowledge (including skills and solving strategies) is the result of cognitive processing. Growth in cognitive processing and knowledge (cognitive abilities) is constrained by limited resources, which show genetically and environmentally influenced individual differences. As a result, some individuals develop as being more intelligent than others. Society selects on intelligence. Its demands influence the interactions among cognitive abilities, hence their intercorrelations, and so their factor loadings. Computer simulations showed that if society promotes the acquisition of knowledge rather than training of cognitive processes, knowledge tests demonstrate higher* g *loadings and heritability coefficients than cognitive processing tests.*

7.1 Better Theory of Intelligence is Warranted

As pointed out in the introduction, to attach a meaning to intelligence other than its interpretation as a rank order on psychometric tests, both the inter-individual differences perspective and the developmental (intra-individual differences) perspective are required. The former is required to account for the covariance structure among the cognitive abilities measured by intelligence tests; the latter to account for cognitive growth. As most models of intelligence (factor models) only concern inter-individual differences in cognitive abilities and do not incorporate cognitive growth, the development of an integrated theory is warranted. The first aim of this thesis was to reinvigorate the development of such theory. Of course, like any theory of intelligence, such integrated theory should be able to account for the most robust and salient empirical findings in intelligence research. These include, for example, the positive correlations among IQ test scores (hence the general factor of intelligence), the heritability of intelligence, the increasing heritability of intelligence throughout development, and the fact that subtests' estimated heritability coefficients correlate with their loadings on the general factor. While working on a model that can account for all of these effects, we noticed that certain empirical findings in intelligence research are not fully addressed in the literature. The second aim of this thesis was to review these findings in more detail.

From the previous chapters we can conclude that in general the dynamical character of intelligence is not fully appreciated in intelligence research. This concerns the dynamic interplay between genes and environment in particular. In this last chapter we focus on this issue, but the message is broader: The mutualism model of intelligence (van der Maas et al. 2006) provides an excellent framework to study effects in intelligence research (in isolation or combined), and to check old and new hypotheses. In Table 7.1 we summarized the most well-known of these findings. The table also includes findings discussed in the previous chapters. We first collate these again below. Next, we provide an example of how the mutualism model can help researchers to study combined effects. We aim to explain one of the most intriguing empirical effects: The more culturally loaded an intelligence subtest is, the higher its heritability estimate is. That is, the better culturally dependent, specific knowledge differentiates between people, the better individual differences in test scores reflect genetic differences. Together with this effect, we aim to explain the fact that the test with the highest cultural load and heritability coefficients are also the most *g* loaded. We implement our hypotheses in the mutualism model and run a number of series of simulations with this model. The results will be discussed.

Table 7.1 Knowns and Unknowns in intelligence research

Factor analytical & Measurement
Positive manifold
Factor indeterminacy
Group factors: e.g. Gf, Gc, Verbal, Perceptual, Image Rotation, etc.
g factor convergence across batteries
High correlation between Gf and *g*
High *g* loading of Raven's
High *g* loadings of knowledge (crystallized intelligence) tests
Lack of unidimensial measures
No direct measure of *g* (*g* is a higher order factor)
Cultural specifity of measurement
Behavior genetics
Heritability of IQ and the factors of intelligence
Gene-environment correlations
Gene-environment interactions
No genes/QTL/SNPs identified
Jensen effect for heritability
Relation heritability, *g* loading and cultural loading
Inbreeding effects
Developmental
Cognitive development
Increasing stability of IQ
Increase of heritability with age (decrease of shared environmentality)
Domain-specific aging trends
Lack of transfer of training
Poor health & pre/perinatal effects on IQ
Specific effects of brain damage
Stage transitions
Correlates
Physical/physiological measures, e.g. brain size
Imaging results
Jensen effects for physical/physiological measures, e.g. brain size
Social economic status
Educational attainment & school performance
Job performance and income
Social outcomes & health
Group differences
Racial and ethnic differences
Jensen effect for Black-White differences
Black-white differences are most pronounced on culture loaded tests and least so on culture reduced tests
Jensen effect for inbreeding effects
Flynn Effect (+ lack of invariance across subtests)

7.2 The Role of Gene Environment Interplay in the Dynamic Development of Intelligence

The dynamical interplay between genes and environment is largely ignored in the statistical (cross sectional) modeling of heritable intelligence. This makes that the further interpretation of the statistical results requires caution. This is so for at least two reasons. First, whereas the hypothesized underlying variables in statistical models are generally modeled as having linear effects, dynamical systems - such as humans - often display behavior that is nonlinear. If nonlinearity is present, a linear statistical model does thus not represent the true data generating mechanism. Second, dynamical gene-environment interplay can lead to gene-environment correlation, which is usually not modeled in statistical models of intelligence. Both the presence of nonlinearity and gene-environment correlation affect heritability estimates. How or to what extent heritability estimates of intelligence are affected is unknown and requires research.

Chapter 2 dealt with the problem of nonlinearity. We noted that the genetic and environmental variables in behavior genetic models are generally not measured, but merely inferred, using genetically informative research designs (e.g. twin and adoption studies), and that these genetic and environmental variables are modeled as having linear relationships with the observed variables, e.g. cognitive abilities measured by IQ tests. However, in reality cognitive development is typically nonlinear and is characterized by stage transitions (see Table 7.1). By means of computer simulations, we investigated the effects on estimated heritabilities and environmentalities in statistical linear behavior genetic modeling in case the underlying mechanisms were nonlinear. In doing so, we were unable to use the mutualism model of van der Maas et al. (2006), because this model does not contain nonlinear terms. We could have developed a nonlinear version of the mutualism model and use this to investigate the effects of nonlinear effects, but this would have required first a full systematic analysis of its dynamical behavior, which was beyond the scope of the thesis. Instead of the mutualism model we used the two-cell model of van Oss & van Ooyen (1996). The main advantage of using the two cell-model was that its dynamics were rich and already known. Although originally the two cells or units represent an inhibitory and an excitatory neuron, they can be interpreted as representing inhibitory and excitatory networks constituting a neural system (van Oss & van Ooyen, 1996), for example a system that underlies working memory. We implemented heritable individual differences in the parameters of the model, ran a series of simulations, and investigated whether the estimated genetic and environmental structure resembled the implemented structure.

In the results of our simulations we observed effects that are in line with empirical findings in intelligence research. First, when nonlinearity was present, estimated heritabilities appeared to increase over time (see Table 7.1), so that the estimated relative contribution of environmental influences decreased. The heritabilities were overestimated, sometimes by as much as 25%. Second, more particular, the estimated relative effects of shared environmental influences decreased (often to values of zero) rather than those of nonshared influences (see Table 7.1). Third, in the behavior genetic modeling the ultimate underlying causes were hard to detect, which fits the fact that to date the search for specific environmental and genetic influences on intelligence has been largely unsuccessful (see Table 7.1). We concluded that within the linear statistical framework, nonlinear development can be considered as constituting a source of phenotypic variance on its own, next to the genetic and environmental sources (see also Molenaar et al., 1993). As in reality cognitive growth is characterized by stage transitions, we hope new mutualism models will be developed that can account for such transitions. We also hope that such model will be used to investigate the role of nonlinear development as a third source of variance in intelligence. Since the presence of stage transitions can lead to an increasing overestimation of heritability coefficients over time, this provides an additional or alternative explanation to the mainstream explanation of the increase of heritability. This increase is usually explained in terms of increasing gene-environment correlation.

Increasing gene-environment correlations also constituted our explanation of the results discussed in Chapter 3, in which we summarized data from 23 independent behavior genetic studies into intelligence. These data showed the contra-intuitive result that the most cultural dependent intelligence tests (i.e. knowledge tests) demonstrated higher heritabilities than the least cultural dependent tests (i.e. cognitive processing tests). These cultural dependent knowledge tests also showed the relatively highest *g* loadings (see Table 7.1). In isolation, mainstream theories of intelligence fail to explain this combined effect (see Chapter 6). Some of them even predict the opposite. We hypothesized that if society puts demands on acquired knowledge rather than cognitive processing, gene-environment correlations would become higher for individual differences in knowledge than individual differences in cognitive processing. In standard behavior genetic modeling gene-environment correlations will be subsumed under the genetic variance, leading to higher heritability estimates.

On the basis of the results from Chapter 2, we can also formulate an alternative or additional hypothesis. As was shown, stage transitions in cognitive development can result in overestimation of the true heritability. The development of certain cognitive abilities may be characterized by more transitions than the development of other cognitive abilities, so that the increase in heritability due to nonlinear effects will differ across abilities. For example, verbal comprehension, usually measured by knowledge tests, may develop stage wise, whereas nonverbal

cognitive abilities, usually measured by cognitive processing tests, develop gradually (whether this is the case should be investigated empirically). Hence, the heritability coefficients of the verbal tests can be overestimated, while those of nonverbal tests are correct or less affected. Whether this alternative explanation of increasing heritability can account for the high *g* loadings of verbal knowledge tests (see Table 7.1) is unknown, but can be investigated with a nonlinear model of general intelligence. Simulations with a mutualism model that includes nonlinear terms may help to discriminate between the two explanations.

Both explanations of the increase in heritability are also relevant in the discussion of racial and ethnic group differences. Black-White differences in IQ exist, for example (see Table 7.1). Chapter 4 showed that they are the most pronounced on verbal knowledge tests (see Table 7.1). If verbal comprehension develops nonlinearly, small differences in initial conditions, e.g. small language deficits in early development or differences in language use (vocabulary), can have large consequences for the further acquisition of knowledge, hence for development of intelligence and scholastic and academic achievement, especially when selection to educational systems is based largely on knowledge and when transmission of knowledge (teaching) takes place verbally. In principle, Black-White mean group differences in later IQ can thus be the result of small initial differences in verbal comprehension or language use, for example. To date, the actual causes are unknown, however, although some researchers (Rushton & Jensen, 2010a) have argued that certain empirical results imply Black-White differences are genetic.

As mentioned Mean Black-White differences in IQ are the most pronounced on verbal knowledge tests. These tests also happen to demonstrate the highest heritabilities and *g* loadings (see above). This effect has been interpreted as a genetic effect (Rushton & Jensen, 2010a). In Chapter 4 we demonstrated that the reasoning behind this interpretation was invalid. We showed analytically that group differences can be the most pronounced on the most heritable and the most *g* loaded tests even when the origin is purely environmental. Moreover, as Chapter 3 showed, the most heritable, most *g* loaded tests are also the most cultural loaded. The finding that Black-White differences are the most pronounced on the most culturally loaded subtests are problematic for Rushton and Jensen's (2010a) biological *g* theory, because this theory predicts the opposite: the mean group differences are most pronounced on the least culturally loaded tests.

In line with Cattell's investment theory of fluid and crystallized intelligence, cultural loaded knowledge tests are often referred to as crystallized tests. Crystallized tests often load highly on a factor that is separate from nonknowledge tests (cognitive processing tests). For this reason, such factor (abbreviated Gc) is often interpreted as 'crystallized intelligence'. In addition, crystallized intelligence is interpreted as the result of the 'investment' of fluid intelligence (reasoning). From our review of the fluid-crystallized theory (Chapter 5), we argued that 1) crystallized intelligence is purely a statistical summary, and that 2) the investment hypothesis of fluid intelligence does not explain why crystallized tests should load on a factor other than the factor fluid intelligence (Gf); 3) a second influence is necessary to dislocate the factor crystallized intelligence from the factor fluid intelligence. Also from the review, we concluded that this second causal influence is most likely verbal comprehension or education. Supported by results from factor analyses, we concluded that once cultural, educational and age differences (hence developmental differences) are taken into account, factor Gc is equivalent to verbal comprehension. We maintained that if researchers find a separate Gc factor, this is due to sample heterogeneity with respect to developmental differences. Likely, such sample heterogeneity also leads to the separation of *g* and Gf. In homogeneous samples they will be equivalent (see Table 7.1), as we found.

Chapter 6 was a theoretical chapter. We addressed in detail whether and to what extent one of the most intriguing empirical findings poses a problem for current theories of general intelligence That is, we asked ourselves whether *g* theories, investment theories, sampling theories, and reciprocal interaction theories explain the finding that the most cultural dependent cognitive abilities (crystallized abilities) are the most *g* loaded and most heritable. We concluded that (in isolation) they do not. By implication, the reviewed theories do not explain how Black-White differences become the most pronounced on the most culturally dependent, most heritable, most *g* loaded subtests. To explain this effect better theory is required. An adequate theory of general intelligence certainly needs to incorporate the developmental perspective.

In the next section we aim to reinvigorate the development of such theory. We hope to so by giving an account for the empirical relations among intelligence subtests' *g* loadings, their heritability coefficients, and their cultural load**.** To this end, we integrate various hypotheses from current theories of general intelligence and implement them in one (mutualism) model.

7.3 An Integrated Model of Intelligence

In developing a new theory, it helps to make a distinction between the factors as statistical entities and their interpretation as realistic common causes of individual differences in cognitive abilities. When doing so, two kinds of theories of general intelligence can be distinguished (see Chapter 6). The one kind, which we denote *g* theories, regards the general factor as representing a realistic (e.g. biological) variable. The positive manifold is regarded as the result of the influence of a general, but unobserved, influence, *g*. Examples of *g* theories are Spearman's (1904) two-factor theory, Jensen's (1968) Level I and level II theory, Cattell's (1963) original investment theory of fluid and crystallized intelligence, and Rushton and Jensen's (2010a) biological *g* theory. The other kind, which we denote alternative theories, do not deny the general factor as a robust statistical phenomenon, but they do not include a variable *g* as an explanatory construct. That is, in these theories the general factor does not explain the positive manifold, rather the positive manifold explains the general factor. Examples of alternative theories are Thomson's (1919; 1951) sampling theory and its elaborations (e.g. Bartholomew et al, 2009b), van der Maas et al.'s (2006) mutualism theory, Dickens' (2008) multivariate version of Dickens and Flynn's (2001) social multiplier theory, and the genetic theories of Anderson (2001) and Penke, Denissen, and Miller (2007).

Like we can distinguish among multiple explanations of the positive manifold, we can also distinguish among multiple explanations of *g* loadings (i.e., regression weights on the statistical general factor). Usually, *g* loadings are explained in terms of complexity. This is the case in Spearman's (1904) and Jensen's (1968) *g* theory, but also in certain alternative theories, for example in Thomson's sampling theory (see Jensen, 1998, for a discussion). We denote these theories complexity theories. Complexity theories predict that the more complex a item or test is the larger its loading on the general factor. Alternatively, *g* loadings have been explained in terms of environmental demands (Dickens & Flynn; Dickens, 2001). Here, the prediction is that the larger the demands of society are, the larger the *g* loading will be. Finally, mutualism theory explains *g* loadings in terms of developmental interactions (van der Maas et al., 2006). The highest *g* loadings are for the cognitive abilities that influence other cognitive abilities the most, and for the cognitive abilities that receive the most influences from other cognitive abilities.

Explanations of *g* loadings are not necessarily mutually exclusive. It is possible that society puts more demands on complex problem solving than noncomplex rote memory or perceptual speed (as in the social multiplier theory), and that the former is practiced more during the course of development than the latter, and so has larger influences on other cognitive abilities during development (as in the mutualism theory). As a result the complex tests display higher *g* loadings (as in complexity theory).

With respect to cognitive development, we can make yet another distinction between theories of intelligence. Sampling theories and *g* theories do not incorporate the developmental perspective, whereas mutualism theory and investment theory do. However, mutualism is not specific enough concerning knowledge acquisition, while investment theory often predicted effects opposite to the empirical effects reviewed in this thesis. Nevertheless, we believe that the essence of investment theory remains tenable. This is the assumption that knowledge acquisition is not only influenced by learning experiences, but also by genetic influences, via cognitive processing. This assumption can be easily implemented in the mutualism model. The assumption that the learning environment is an important mediator of these genetic effects can also be implemented. These assumptions may sound trivial, but since current theories do not fully account for the effect we want to explain, we contend it is important to be explicit about which hypotheses we retain and which we do not.

Our integrated model thus borrows from the investment theory of fluid and crystallized intelligence, but its basis is the mutualism model of van der Maas et al. (2006). The model also

incorporates the idea of a social multiplier effect (Dickens & Flynn, 2001). That is, more intelligent people end up in more cognitive stimulating environments. In the next section, we present the model in more detail and aim to explain how knowledge, the result of heritable cognitive processing and environmental influenced learning experiences, displays higher heritability than these causes themselves. We contend that the answer lies in a dynamical interplay between heritable abilities and the demands of society.

7.3.1 Hypotheses

We make the usual distinctions between crystallized abilities (knowledge, including skills and solving strategies) and fluid cognitive abilities (variables that represent differences in cognitive processing). In line with the mutualism theory, we do not consider fluid intelligence to constitute a single, realistic latent variable, but rather as the result of mutual beneficial interactions among multiple basic cognitive processes. In line with the investment hypothesis, we assume relatively large influences of fluid abilities on crystallized abilities. Since transfer of training is generally absent (e.g. Cattell, 1987; see Table 7.1), we assume that crystallized abilities do not influence each other directly. In line with mutualism, we assume that due to the process of learning, cognitive processing benefits from the development of crystallized abilities. The development of working memory, for example, benefits from actively learning the knowledge and skills that are demanded by the cognitive environment. In line with Dickens and Flynn (2001), we assume that more intelligent people end up in more cognitive stimulating environments which provide opportunities to further train cognitive skills and to acquire knowledge. In line with mutualism, the growth of such cognitive abilities is constrained by limiting capacities. Although we believe genetic sampling is present in reality (any two limiting capacities can share some genetic determinants), we also believe that this assumption is not necessary to account for the effects under investigation. The same holds for sampling on test level. With test sampling we mean that, for example, solving an item of an intelligence tests always requires previous acquired knowledge, as well as cognitive processing, but that items call upon acquired knowledge and cognitive processing to different degrees. With test sampling, we also mean that certain processes are called upon across different tests, for instance that working memory is involved in both verbal and spatial tasks. Figure 7.1 shows a graphical representation of the fully integrated model.

7.3.2 Explorations of the integrated model

Incorporating the assumptions above in the mutualism model (van der Maas, 2006), preliminary results have shown that multiple variants of this setup can lead to the observation that *g* loading and heritabilities correlate positively, crystallized abilities being the most heritable and having the highest loadings on the general factor. The variants differed with respect to the presence or absence of details of the setup (see Appendix B). The variants fall in two categories. In the first category, the systems have reached their equilibrium points, hence approximate their heritable limited capacities. In the second category they have not.

Variants of the first category only differed from the simulations of van der Maas et al. (2006) in that the mutualism matrix is made more specific (see Appendix B). We specified this matrix in accordance with the present model. For explanation of the simulation procedure and programming codes, we refer to van der Maas et al. (2006) and the R mutualism package. Both are available at http://hvandermaas.socsci.uva.nl/Homepage_Han_van_der_Maas/Home.html. Here one can also find all programming codes of the simulations that are discussed in this thesis. Here, we merely mention our main conclusions.

For the effect to appear that crystallized abilities display the largest heritabilities, the presence of genetic sampling (weak genetic intercorrelations) was required. In addition, it had to be assumed that the cognitive environment selects on general intelligence, i.e., both crystallized and fluid abilities, rather than either crystallized or fluid intelligence. Last, large genetic effects on the limiting capacities as well as large environmental effects on cognitive abilities were required.

Because in simulations of the first category all systems were in equilibrium, this may be interpreted to mean that the associated abilities were fully developed for every person in the sample.

We consider this as not very plausible. We therefore also studied the situation in which the systems were not in equilibrium. This setup is consistent with the situation in which the subjects in a given sample are heterogeneous with respect to the amount of practice. This implies the presence of individual differences in the abilities that stem from differences in development. We contend that further increase in intelligence is possible, and that crystallized abilities are usually more practiced than fluid abilities, and thus reflect better the genetically influenced limited capacities. In this setup, we could relax the assumption that environment selects on general intelligence. That is, for the Jensen effect for heritability to appear, we could assume that cognitive environment selects on knowledge and skills (hence, on crystallized intelligence) rather than general intelligence, which we consider more plausible. This assumption was not necessary, however. Neither necessary was the assumption of genetic sampling. Crystallized abilities had the largest *g* loadings and heritabilities, provided environmental influences had a large effect.

Because the simulations of the second category differ substantially from the simulations in van der Maas et al. (2006) in a number of aspects, we discuss below one of the variants in more detail (any other variants are described briefly in Appendix B). As the simulations could not be based on the equilibrium covariance matrix, we simulated data using the mutualism model itself (basic simulations with the mutualism R package are discussed in van der Maas et al).

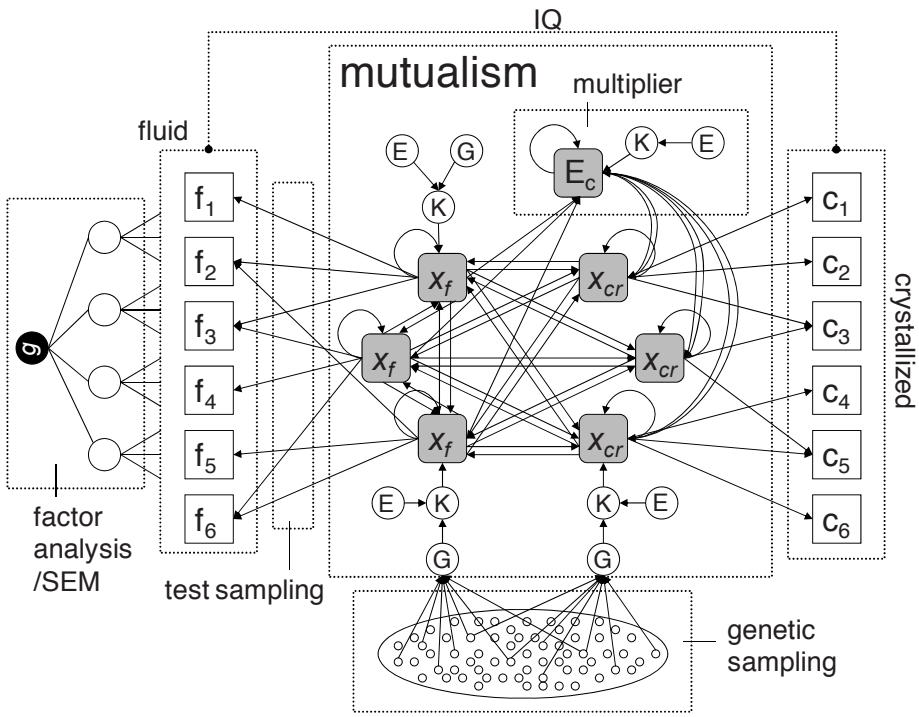


Figure 7.1 The fully integrated model. Correlations among fluid abilities are the result of mutual beneficial interaction between basic cognitive processes during development. Crystallized abilities (knowledge and skills) are the result of cognitive processing. The development of crystallized abilities has a beneficial effect on the development of fluid abilities. Some individuals turn out to be more intelligent, because they posses higher levels of cognitive abilities. More intelligent individuals will be more likely to end up in cognitive environments conducive to the further development of crystallized abilities (which is beneficial to the further development of fluid abilities). The growth of the cognitive abilities will be constrained by genetically and environmentally influenced limited capacities. These capacities are possibly weakly intercorrelated due to genetic sampling. Sampling on test level may also be present. General intelligence, as the outcome of factor analysis, is an index measure of cognitive functioning.

A simulation example One way to accomplish that crystallized abilities are practiced more than fluid abilities in the mutualism model is to assume different time scales for the development of crystallized and fluid abilities. We can accomplish this by assigning differences in growth rates (i.e., the parameters *a*, in van der Maas et al. 2006; see also Appendix B); crystallized abilities having higher growth rates than fluid abilities.

We simulated data for 17 observed variables (8 fluid abilities, 8 crystallized abilities, and 1 environmental variable) for N = 500 'monozygotic (MZ) twin pairs raised apart'. The mutualism matrix is taken as in Figure 7.2, with an additional, but arbitrary, amount of noise (mean 0; SD .02). Following van der Maas et al. (2006), starting values were sampled from a (17 dimensional) uncorrelated multivariate normal distribution with means equal to .05 and SD's of .01. Growth parameters were sampled from a 17 dimensional uncorrelated multivariate normal distribution with arbitrary means of 6 (SD = .5) for fluid abilities and 9 (SD = .5) for crystallized abilities and the environmental variable.

	Fluid abilities (x_f)								Crystallized abilities (x_c)								Cognitive environment (x_e)	
	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.02	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	x_f
	.02	.02	.02	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.02	.02	.02	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.02	.00	
	.02	.02	.02	.02	.02	.02	.02	.00	.02	.02	.02	.02	.02	.02	.02	.02	.00	
M	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	x_c
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	.04	.04	.04	.04	.04	.04	.04	.04	.00	.00	.00	.00	.00	.00	.00	.00	d	
	d	d	d	d	d	d	d	d	d	d	d	d	d	d	d	d	00	x_e

Figure 7.2 Matrix of mutualistic weights used in the example simulation study (see text). It contains the interactions (weights) among fluid and crystallized abilities and the cognitive environment.

The parameter values of the limited capacities (the parameters K in van der Maas et al., 2006; see also Appendix B) were determined as follows. N values were sampled from a 16 dimensional multivariate normal distribution with means 0 and SD's of 0.5 (these SD's equal the arbitrary SD's in van der Maas et al., 2006), with intercorrelations r_g. These parameter values represent the genetic values of the limited capacities associated with the cognitive abilities. They were taken identical for each twin (so we obtained 2*N values for each capacity associated with the cognitive abilities). Across simulations r_g varied (from 0 to .5), but within each simulation r_g was fixed (i.e., did not vary over cases). The genetic value for the limited capacity associated with the cognitive environmental variable was fixed at zero for every individual. In addition, 2*N values were sampled from a 17 dimensional uncorrelated multivariate normal distribution with means 0 and standard deviations .5. These represent the environmental values of the limited capacities.

We added the genetic and environmental values together with weights $\sqrt{(h^2)}$ and $\sqrt{(1- h^2)}$ respectively, where h^2 was set at an relatively high, but arbitrary value of 0.7. This implies that the capacities associated with the cognitive variables had an expected heritability of 0.7, and that the capacity associated with the environmental variable had an expected heritability of 0. Next, we

rescaled the values to accomplish they had a mean of 3. In this way, the mean and SD of the parameters equaled the arbitrary settings of van der Maas et al. (2006).

To obtain data of N twins, we ran the mutualism model $2*N$ times with the twin parameter values. Data were obtained for one point in time (set arbitrarily at $t = 1$), after the beginning of development ($t = 0$), but before the systems reached their end states (which was around $t = 2.5$). Each simulation resulted in a data set of measurements of the 17 variables of the $2*N$ subjects. The observed heritabilities were estimated by calculating the MZ twin correlation for each of the 16 cognitive abilities. The values on the cognitive abilities of the first members of the twin pairs were submitted to factor analysis. That is, we followed the common practice of inspecting the intercorrelations of the observed variables, plotting the correlation matrix' eigenvalues, and fitting the common factor model.

Figure 7.3 shows the results of a typical simulation run in which d (the interaction weights between the cognitive environment and cognitive abilities; see Figure 7.2) is 0.2, and in which the limited capacities were all completely independent (i.e., $r_g = 0$). Thus, genetic or environmental intercorrelations among the underlying limited capacities were absent. Correlations among the observed cognitive abilities were all positive, the first eigenvalue was dominant and a common factor model fitted the data. Heritabilities and g loadings showed a positive correlation. Only when d was positive and relatively large did the crystallized abilities display the highest g loadings and heritabilities.

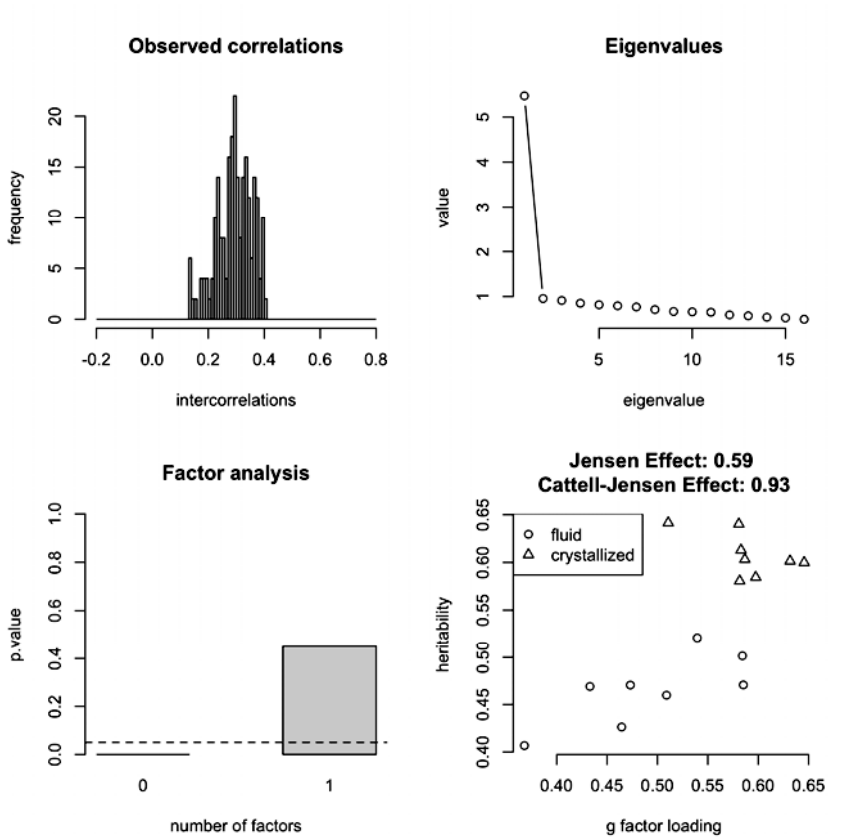


Figure 7.3 The results of a typical simulation run with the mutualism model, using the design matrix from Figure 7.2, and assuming completely independent limiting capacities. A dominant first eigenvalue is present and a common factor model fits the data. Crystallized tests have the largest loadings on this factor. They are also the most heritable.

7.4 Overall Conclusions

Current theories of intelligence in isolation do not account for the empirical relations among intelligence subtests' loadings on the general factor, heritability coefficients, and their cultural load. One of the purposes of this chapter was to account for these relations. To this end, we integrated hypotheses form different theories of intelligence. Another purpose of this chapter was to show that the mutualism model of intelligence can be used to study (combined) effects and new and old hypothesis. As an example we implemented the hypotheses above in the mutualism model. From our simulation studies, we concluded that the relations among *g* loading, heritability, and cultural load will be positive when the society fosters crystallized abilities rather than fluid abilities.

The simulations also showed that the general factor of intelligence does not have to represent a realistic (e.g. biological) common cause of individual differences. In our integrated model it is a merely a statistical entity, i.e., an informative summary descriptive. However, as such, general intelligence certainly has utility. Specifically, we consider it to be similar to 'general health'. Like general health is an informative summary descriptive of physical functioning, '*g*' is an informative summary descriptive of cognitive functioning. A distinction between *g* as a statistically entity, and *g* as interpreted as a realistic, common cause of individual differences, is not only important theoretically, but also empirically. Consider genetic association and linkage studies of intelligence, for example. So far, the search for genes for general intelligence has met with relatively little success (Deary, Johnson, & Houlihan, 2009; Plomin and Spinath, 2004; Chabris et al., *in press*). The alternative theories of Dickens and Flynn (2001; Dickens, 2008) and of van der Maas et al. (2006), in which the general factor of intelligence is a statistical entity originating in reciprocal beneficial interactions among cognitive processes or abilities, are able to provide a plausible explanation of this lack of success. In these theories, there are no direct genetic influences general to all abilities. If general intelligence is indeed the outcome of such interactions, the search for specific 'genetic influences on *g*' is a questionable undertaking (Dolan, Kan, van der Maas, 2008; van der Sluis, Kan & Dolan, 2010), in particular when (1) the genetic determinants of these processes are under mutation selection balance, which means that the process of natural selection reduces genetic variance but cannot deplete all of it because new variation - due to mutations - is continually reintroduced (Penke et al., 2007), and (2) the interactions are nonlinear (e.g. Molenaar, Boomsma, & Dolan, 1993; Kan, Ploeger, Raijmakers, Dolan & van der Maas, 2011).

Cross-sectional factor analysis cannot discriminate between *g* theories and alternative theories of general intelligence. In addition, both *g* theories and alternative theories can explain the facts that the general intelligence is (highly) heritable and has (strong) biological correlates. So these facts will not discriminate between different theories either. Developmental analysis on the other hand, may have the potential to discriminate between theories of intelligence. For example, group factors of intelligence are more or less genetically independent initially, but become increasingly genetically interrelated as development unfolds (Hoekstra et al., 2007). This observation is consistent with mutualism, but is not readily explained by *g* theories or sampling theories. Developmental effects are thus important in the understanding of individual differences. We hope future (behavior genetic studies) studies into intelligence will use longitudinal modeling. Together with formal, dynamical system modeling, this will enrich intelligence research.

APPENDIX A

MATHEMATICAL DESCRIPTION OF THE TWO-CELL MODEL

Van Oss and van Ooyen (1997) model neuronal activity as:

$$\frac{dx}{dt} = -x + (1-x)wf(x) - (h+x)wpf(y) \quad (1)$$

and

$$\frac{dy}{dt} = -y + (1-y)wpf(x) \quad (2)$$

where

$$f(u) = \frac{1}{1+e^{(\theta-u)/\alpha}} \quad (3)$$

The variable x is interpreted as the average membrane potential of a population of excitatory neurons, and y that of inhibitory ones. The cells outputs are taken their mean firing rates, which are a sigmoid function f of their membrane potentials. The function f has two parameters: α, the steepness of the function, and θ, the firing threshold. In equation 1 and 2, w represents the connection strength between cells. Furthermore, $w = w_{xx}$, $w_{xy} = w_{yx}$, $w_{yy} = 0$ (See Figure 2.1). Parameter h represents the relative inhibitory saturation potential (the finite maximum membrane potential) compared to the excitatory saturation potential, and p represents the relative strength of the inhibitory connection ($p = w_{xy} / w_{xx}$).

The dynamics of w are governed by the equation:

$$\frac{dw}{dt} = q(\varepsilon - bw^2 - x) \quad (4)$$

where q determines the outgrowth rate of neurons. Finally, b stands for the degree of saturation, while ε stands for the membrane potential at which the cell neither retracts nor extends its neurites (i.e., at which neurite outgrowth is 0).

In this model the dynamics of the connection strength are considerably slower compared to the dynamics of neuronal activity. Hence, variables x and y can be displayed in a *(w, x)* or *(x, y)* -plane as a function of w (as depicted in Figure A.1). In such a diagram one can observe a characteristic S-shaped curve, which van Oss and van Ooyen (1997) call the slow manifold. The set of points defined by $dw/dt = 0$, is called the null-cline. Intersections of the slow manifold with the w-nullcline are the equilibrium points of the system. Variables x and y are able to show transient behavior jumps from the one to another trajectory. In addition, as w changes x and y can either end up in a point attractor or in a stable or unstable limit cycle, depending on the w-nullcline.

By means of bifurcation analysis van Oss and van Ooyen (1997) distinguished different regions in the parameter plane *(ε, p)*. These regions are characterized by the number and stability of equilibrium points. Furthermore, three types of limit cycles appeared to exist: 1) 'relaxation oscillations' (very slow oscillations in x, y, and w with high amplitude), 2) 'bursting oscillations' (a single fast oscillation in x and y, and w almost non-oscillating), 3) 'fast oscillations' (a relatively high amplitude in x and y, and low amplitude in w). Subregions are defined as having equal number and stability of equilibrium points, but having different types of limit cycle attractors. Figure A.2 displays

the parameter regions, subregions, and bifurcation lines within the parameter plane *(ε, p)*. Multistability can occur in (sub)regions 1b, 2b, 4d, 5b, 6, 7, 9, 10, 11, and 12.

In most cases of multistability, there exist a point attractor at low values of connectivity (*w*) and either a second point attractor or an additional limit cycle of type 2 or 3 for high values of connectivity. At intermediate values there exists an unstable equilibrium (unstable node or limit cycle). However, for values of $p > 0.43$ - in some cases - the second attractor can be reached even for low values of *w*. Switching from the first attractor to the second is possible when *w* is decreased, *x* is chronically decreased, or *y* is chronically increased. Reversely, if *w* is increased, x chronically increased, or *y* chronically decreased, a switch from the second to the first attractor can be accomplished.

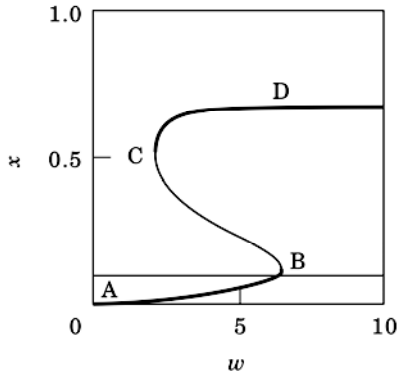


Figure A.1 Copy of Fig. 2b in van Oss and van Ooyen (1997): An S-shaped curve and the w-nullcline (horizontal line). Bold lines indicate stable equilibrium points with respect to x, when w is regarded as a parameter.

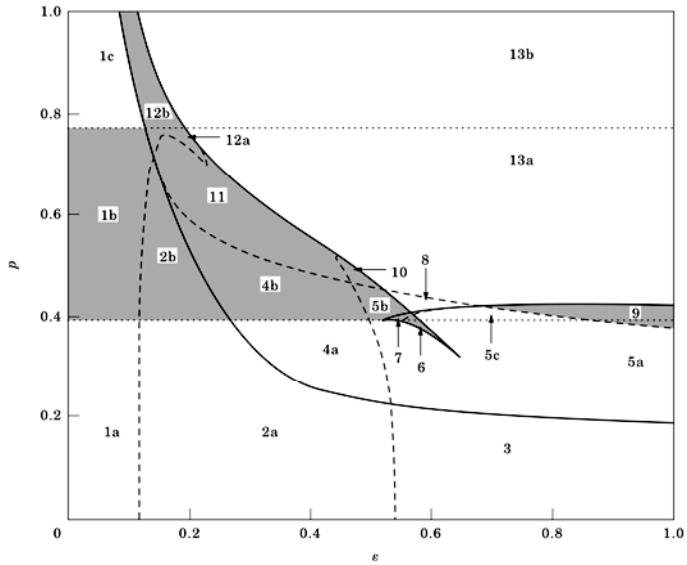


Figure A.2 Copy of Fig. 1 in van Oss & van Ooyen (1997): Regions and subregions in the (ε, p)-parameter plane. In the grey regions multistability can occur; between the dotted lines oscillations can occur. Continuous lines represent fold lines; dashed lines represent Hopf lines.

APPENDIX B

SETUP OF THE SIMULATIONS IN CHAPTER 7

The mutualism package, codes of basic simulations with the mutualism model, and the codes pertaining to the simulations in Chapter 7 can be found at http://hvandermaas.socsci.uva.nl/Homepage_Han_van_der_Maas/Home.html. The mutualism model used in these simulations is formulated mathematically as follows:

$$\frac{dx_i}{dt} = a_i x_i (1 - x_i / K_i) + a_i \sum_{\substack{j=1, \\ j \neq i}}^{W} M_{ij} x_j x_i / K_i \qquad \text{for } i, j = 1 \ldots W.$$

The covariance matrix of the observed x variables in equilibrium is:

$$\Sigma = [\mathbf{I} - \mathbf{M}]^{-1} \Psi [\mathbf{I} - \mathbf{M}]^{-\mathrm{T}}$$

where $\mathbf{\Psi}$ is a diagonal matrix containing the variances of the K's, and $\mathbf{I}$ is the identity matrix. Superscript -1 denotes matrix inversion, and superscript -T denotes inversion and transposition.

In the simulations in Chapter 7, we generated data for $W = 17$ variables (8 fluid abilities, 8 crystallized abilities, and 1 environmental variable). The K's are normally distributed with means of 3 and SD's 0.5. Parameters a are normally distributed with means of 6 (unless stated otherwise) and SD's 0.5. **M** is taken as in Figure B.1. Values *ec* and *ce* are always set equal. Values *cc* are always set 0.

Highest *g* loadings and heritabilities for crystallized abilities could be obtained in multiple ways (see and below, and Table B.1). Two categories of simulations can be made. In the first category, simulations were based on the covariance matrix, i.e., when systems are in equilibrium (see the mutualism R help file, topic Jensen effect). In the second category simulations are based on the mutualism model (see van der Maas et al, 2006, and the mutualism R help file).

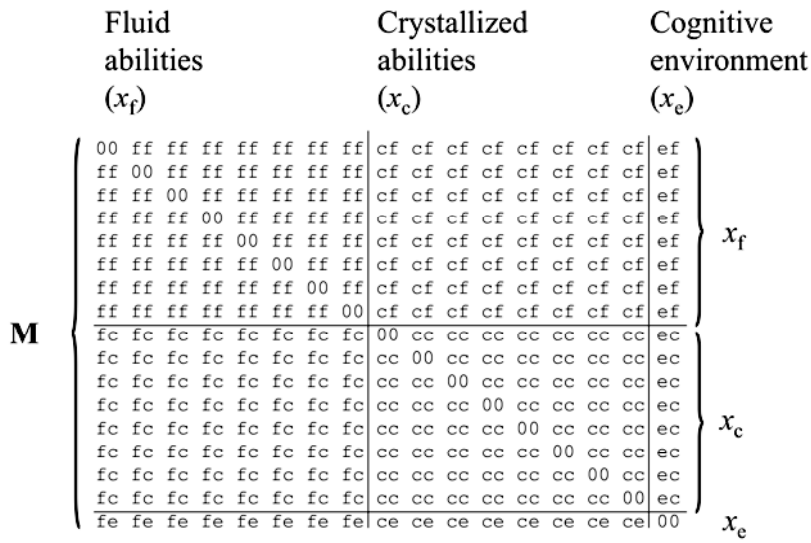


Figure B.1 Generalized design Matrix M used in the simulations in Chapter 7.

Category 1

Variant 1: *fc* is positive, other values in **M** are zero, and intercorrelations between the *Kg*'s are high. This variant can be regarded as a revision of investment theory. Here, *g* represents the total genetic value of the genes that influence the limited capacities which constrain the growth of fluid and crystallized abilities.

Variant 2: *ff* and *cf* are positive and relatively low, *fc* is positive and relatively high, *fe*, *ce* and *ec* are large (and equal), while other weights in **M** are 0, and the *Kg*'s are intercorrelated. Only weak intercorrelations are required. This is a scenario in which mutualism, a multiplier effect, and genetic sampling are present. The environment selects on general intelligence (both fluid and crystallized abilities), but gives practice primarily in crystallized abilities.

Category 2

Variant 3: All *ff* and *cf* are positive and relatively low, *fc* are positive and relatively high, *fe*, *ce* and *ec* are large (and equal), while all other weights in **M** are 0, and the subjects sample is assumed to be heterogeneous. This is a scenario in which mutualism, and a multiplier effect are present. The cognitive environment selects on general intelligence, but gives practice primarily in crystallized abilities.

Variant 4: All *ff* and *cf* are positive and relatively low, *fc* are positive and relatively high, *ce* and *ec* are large (and equal), while all other weights in **M** are 0 and the subjects sample is assumed to be heterogeneous. This is a scenario in which mutualism and a multiplier effect are present. The cognitive environment primarily selects on, and is conducive to practice in crystallized abilities. Heterogeneity can be introduced by specifying larger *a*'s for crystallized abilities than fluid abilities, while *t* is equal over subjects (in our simulations *t* is set at 1), or by assuming normally distributed *a*'s with equal means and heritable *t* (in our simulations *t* is normally distributed, with mean 1, and SD = .02).

Table B.1 Assumptions under which crystallized abilities have higher heritabilities than fluid abilities. The effect can be obtained in multiple ways.

Variant	Mutualism is present	Systems in equilibrium	Multiplier is present	Environment selects on	Environment gives practice in	Genetic sampling is present
1	Yes	Yes	No	-	-	Yes, high overlap (some general genes)
2	Yes	Yes	Yes	General intelligence	Crystallized intelligence	Yes, low overlap (no general genes)
3	Yes	No	Yes	General intelligence	Crystallized intelligence	No, not necessarily
4	Yes	No	Yes	Crystallized intelligence	Crystallized intelligence	No, not necessarily

REFERENCES

Ackerman P. L. (1996). A theory of adult intellectual development: process, personality, interests, and knowledge. *Intelligence 22*, 227-57.

Ackerman, P. L. 2000. Domain-specific knowledge as the "dark matter" of adult intelligence: gf/gc, personality and interest correlates. *Journal of Gerontology: Psychological Sciences*, *55B(2)*, 69-84.

Ackerman, P. (2003). Cognitive ability and non-ability trait determinants of expertise. *Educational Researcher, 32*, 15-20.

Ackerman, P. L., & Lohman, D. F. (2003). Education and *g*. In H. Nyborg (Ed.), *The scientific study of general intelligence. Tribute to Arthur. R. Jensen.* (pp. 275-292). Amsterdam: Pergamon Press.

Alarcon, M., Knopik, V. S., & DeFries, J. C. (2000). Covariation of mathematics achievement and general cognitive ability in twins. *Journal of School Psychology, 38*, 63-77.

Alfonso, V.C., Flanagan, D.P., & Radwan, S. (2005). The impact of Cattell-Horn-Carroll theory on test development and the interpretation of cognitive abilities. In D.P. Flanagan & P.L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests and issues* (2nd ed., pp. 185-202). New York: Guilford Press.

Allison, R. B. (1960). *Learning parameters and human abilities*. Princeton, N. J.: Educational Testing Service, Office of Naval Research Technical Report.

Amthauer, R., Brocke, B., Liepmann, D., & Beauducell, A. (2001). Intelligenz-Struktur-Test 2000 R. Göttingen: Hogrefe.

Anderson, B. (2001). *g* as a consequence of shared genes. *Intelligence, 29*, 367-371.

Archer, G. S., Dindot, S., Friend, T.H., Walker, S., Zaunbrecher, G., Lawhorn, B., & Piedrahita, J.A. (2003). Hierarchical phenotypic and epigenetic variation in cloned swine. *Biology of Reproduction, 69*, 430-436.

Archer G. S., Friend T. H., Piedrahita J., Nevill C. H. & Walker S. (2003). Behavioral variation among cloned pigs. *Applied Animal Behavior Science, 82*, 151-161.

Arnold, V. I., Afrajmovich, V. S., Il'yashenko, Y. S., & Shil'nikov, L. P. (1994). *Bifurcation theory and catastrophe theory*. Berlin: Springer-Verlag.

Ashton, M.C., & Lee, K.M. (2005). Problems with the method of correlated vectors. *Intelligence, 33*, 431- 444.

Baltes, P.B., Schaie, K.W. 1974. Aging and IQ: The Myth of the Twilight Years. *Psychology Today*, *7*, 35-40.

Bartels, M., Rietveld, M. J. H., Baal, G. C. M. van, & Boomsma, D. I. (2002). Heritability of educational achievement in 12-year olds and the overlap with cognitive ability. *Twin Research, 5*, 544-553.

Bartholomew, D. J. (2004). *Measuring intelligence: Facts and fallacies*. Cambridge, UK: Cambridge University Press.

Bartholomew, D. J., Deary, I. J. & Lawn, M. (2009a). A new lease of life for Thomson's bonds model of intelligence. *Psychological Review, 116*, 567-579.

Bartholomew, D. J, Deary, I. J. & Lawn. M. (2009b). The origin of factor scores: Spearman, Thomson and Bartlett. *British Journal of Mathematical and Statistical Psychology 62*, 569 - 582.

Beauducell, A., Brocke, B., & Liepmann, D. (2001). Perspectives on fluid and crystallized intelligence: Facets for verbal, numerical, and figural intelligence. *Personality and Individual Differences, 30*, 977-994.

Beauducell, A., & Kersting, M. (2002). Fluid and crystallized intelligence and the Berlin Model of Intelligence Structure (BIS). *European Journal of Psychological Assessment, 18*, 97-112.

Beaujean, A. A., (2005). Heritability of cognitive abilities as measured by mental chronometric tasks: A meta-analysis. *Intelligence, 33*, 187-201.

Beloussov, L. V. (2006). A Morphomechanical Aspect of Epigenesis. *Russian Journal of Genetics, 42 (9)*, 966-969.

Benno, R. H. (1990). Development of the nervous system: Genetics, epigenetics and phylogenetics. In Hahn, M. E., Hewitt, J. K., Henderson, N. D., and Benno, R. H. (eds.), *Developmental behavior genetics: Neural, biometrical, and* e*volutionary approaches* (pp. 113-143). New York: Oxford University Press.

Betjemann, R. S., Johnson, E. P., Barnard, H., Boada, R., Filley, C. M., Filipek, P. A., Willcutt, E. G., DeFries, J. C., & Pennington, B. F. (2010). Genetic covariation between brain volumes and IQ, reading performance, and processing speed. *Behavior Genetics, 40*, 135-145.

Bickley, P. G., Keith, T. Z., & Wolfle, L. M. (1995). The three-stratum theory of cognitive abilities: Test of the structure of intelligence across the life span. *Intelligence, 20*, 309-328.

Blecker, E. G., (1983). Cognitive defense style and WISC-R P > V sign in juvenile recidivists. *Journal of Clinical Psychology, 39*, 1030-1032.

Bollen, K., & Lennox, R. (1991). Conventional wisdom on measurement: A structural equation perspective. *Psychological Bulletin, 110*, 305-314.

Block, J. B. (1968). Hereditary components in the performance of twins on the WAIS. In S. G. Vandenberg (Ed.), *Progress in human behavior genetics* (pp. 221-228). Baltimore: The Johns Hopkins University Press.

Borsboom, D., & Dolan, C. V. (2006). Why *g* is not an adaptation: A comment on Kanazawa. *Psychological Review*, *113*, 433-437.

Borsboom, D., & Mellenbergh, G. J. (2002). True scores, latent variables, and constructs: A comment on Schmidt and Hunter. *Intelligence*, *30*, 505-514.

Borsboom, D., Mellenbergh, G. J., & Van Heerden, J. (2003). The theoretical status of latent variables. *Psychological Review, 110*, 203-219.

Bouchard, T. J. (2004). Genetic influence on human psychological traits. *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.

Brody, N. (1992). *Intelligence*. San Diego: Academic Press.

Burt, C. (1940). *The factors of the mind*. London: University of London Press.

Camazine, S. (2001). *Self-organization in biological systems*. Princeton, NJ: Princeton University Press.

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge, United Kingdom: Cambridge University Press.

Carroll, J. B. (1996). A three-stratum theory of intelligence: Spearman's contribution. In I.Dennis, I. & P. Tapsfield (Eds.), *Human abilities: Their nature and measurement* (pp. 1-18). Mahwah, NJ: Erlbaum.

Carroll, J. B. (1998). Human cognitive abilities: A critique. In J. McArdle, & R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 5-24). Mahwaw, NJ: Lawrence Erlbaum Associates.

Carroll, J. B. (2003). The higher stratum structure of cognitive abilities: Current evidence supports *g* and about ten broad factors. In H. Nyborg (Ed.), *The scientific study of general intelligence: Tribute to Arthur R. Jensen*. New York: Pergamon.

Cattell, R. B. (1941). Some theoretical issues in adult intelligence testing. [Abstract] *Psychological Bulletin, 38*, 592.

Cattell, R. B. (1943). The measurement of adult intelligence. *Psychological Bulletin 40*, 153-193.

Cattell, R. B. (1957). *Personality and motivation structure and measurement*. New York: World.

Cattell, R. B. (1963). Theory for fluid and crystallized intelligence: A critical experiment. *Journal of Educational Psychology, 54*, 1-22.

Cattell, R. B. (1971). *Abilities: Their structure, growth, and action*. Boston: Houghton-Mifflin.

Cattell, R. B. (1987*). Intelligence: Its structure, growth and action*. Amsterdam: North-Holland.

Ceci, S. J. (1996). *On intelligence. . . more or less: A bio-ecological treatise on intellectual development.* Englewood Cliffs, NJ: Prentice Hall.

Chabris, C. F., Hebert, B. M., Benjamin, D. J., Beauchamp, J., Cesarini, D., van der Loos, M. J. H. M., Johannesson, M., Magnusson, P. K. E., Lichtenstein, P., Atwood, C. S. et al. (*in press*). Most Reported Genetic Associations with General Intelligence Are Probably False Positives. *Psychological Science*.

Daucé, E. (2007). Learning and control with large dynamic neural networks. *European Physical Journal Special Topics, 142*, 123-161.

Deary, I. J., Egan, V., Gibson, G. J., Austin, E. J., Brand, C. R., & Kellaghan, T. (1996). Intelligence and the differentiation hypothesis. *Intelligence, 23*, 105-132.

Deary, I. J., Johnson W., Houlihan L., M. (2009) Genetic foundations of human intelligence. *Human Genetics, 126*, 215-232.

Demetriou, A. 2002. Tracing psychology's invisible giant and its visible guards. In R. J. Sternberg & E. L. Grigorenko (Eds.). *The general factor of intelligence: How general is it?* (pp. 3-18). Mahwah, NJ: Erlbaum.

Diamantopoulos, A., P. Riefler, K. P. Roth. 2008. Advancing Formative Measurement Models. *Journal of Business Research, 61 (12),* 1203-1218.

Dick, D.M., & Rose, R.J. (2002). Behavior genetics: What's new, what's next? *Current Directions in Psychological Science, 11*, 70-74.

Dickens, W. T. (2008). *What is g*? Paper presented at the 38th annual meeting of the Behavior Genetics Association. University of Louisville. June 28 2008.

Dickens, W. T. (unpublished manuscript). *What is g?* Retrieved on 01/01/10 from http://www.brookings.edu/papers/2007/0503education_dickens.aspx

Dickens, W. T., & Flynn, J. R. (2001). Heritability estimates versus large environmental effects: The IQ paradox resolved. *Psychological Review, 108*, 346-369.

Dolan, C. V. (2000). Investigating Spearman's hypothesis by means of multi-group confirmatory factor analysis. *Multivariate Behavioral Research, 35*, 21-50.

Dolan, C. V., & Hamaker, E. L. (2001). Investigating Black-White differences in psychometric IQ: Multi-group confirmatory factor analyses of the WISC-R and K-ABC and a critique of the method of correlated vectors. In Frank Columbus (Ed.), *Advances in Psychology Research, vol. 6* (pp. 31-59). Huntington, NY: Nova Science Publishers.

Dolan, C.V., & Molenaar, P.C.M. (1995). A note on the scope of developmental behaviour genetics. *International Journal of Behavioral Development, 18*, 749-760.

Dolan, C. V., Kan, K. J., & van der Maas, H. L. J. (2008). *The consequences of the mutualistic theory of "g" for QTL searches*. Paper presented at the 38th annual meeting of the Behavior Genetics Association. University of Louisville. June 28 2008.

Dolan, C. V., Colom, R., Abad, F. J., Wicherts, J. M., Hessen, D. J., & van der Sluis, S. (2006). Multi-group covariance and mean structure modelling of the relationship between the WAIS-III common factors and sex and educational attainment in Spain. *Intelligence, 34*, 193-21.

Eaves, L. J., & Erkanli, A. (2003). Markov Chain Monte Carlo approaches to analysis of genetic and environmental components of human developmental change and G x E interaction. *Behavior Genetics, 33*, 279-99.

Eaves, L. J., Kirk, K. M., Martin, N. G., & Russell, R. J. (1999). Some implications of chaos theory for the genetic analysis of human development and variation. *Twin Research, 2(1)*, 43-48.

Eckert, M. A., Leonard, C. M., Molloy, E. A., Blumenthal, J., Zijdenbos, A., &, Giedd, J. N. (2002). The epigenesis of planum temporale asymmetry in twins. *Cerebral Cortex, 12*, 749-755.

Edmonds, C. J., Isaacs, E. B., Visscher, P. M., Rogers, M., Lanigan, J., Singhal, A. Lucas, A. Gringras, P., Denton, J., Deary, I. J. (2008). Inspection time and cognitive abilities in twins aged 7 to 17 years: age-related changes, heritability, and genetic covariance. *Intelligence, 36*, 210-225.

Edwards, J. R. (2011). The Fallacy of Formative Measurement. *Organizational Research Methods, 14(2),* 370-388.

Edwards, J. R., & Bagozzi, R. P. (2000). On the nature and direction of relationships between constructs and measures. *Psychological Methods, 5*, 155-174.

Ekström, R. B., French, J. W., Harman. H., & Dermen, D. (1976). *Kit of factor-referenced cognitive tests* (Rev. ed.). Princeton, NJ: Educational Testing Service.

Evans, J., Floyd, R., McGrew, K. S., & LeForgee, M. (2002). The relations between measures of Cattell-Horn-Carroll (CHC) cognitive abilities and reading achievement during childhood and adolescence. *School Psychology Review, 31*(2), 246.

Falconer, D. S., & Mackay, T. F. C. (1996). *Introduction to quantitative genetics (4th ed.)*. Harlow, UK: Addison Wesley Longman.

Ferrer, E., & McArdle, J. J. (2004). An experimental analysis of dynamic hypotheses about cognitive abilities and achievement from childhood to early adulthood. *Developmental Psychology*, *40*, 935-952.

Finch, C. E., Kirkwood, T. B. L. (2000). *Chance, development, and aging*. New York: Oxford University Press.

Fink, B., Manning, J. T, Neave, N., & Grammer, K. (2004). Second to fourth digit ratio and facial asymmetry. *Evolution and Human Behavior, 25(2)*, 125-132.

Flanagan, D. P. & Ortiz, S. O. & Alfonso, V.C. (2007). *Essentials of Cross-Battery Assessment* (*2nd. Ed)*. New York : Wiley Press

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive function are almost entirely genetic in origin. *Journal of Experimental Psychology, 137*, 201-225.

Gärtner, K. (1990). A third component causing random variability beside environment and genotype: A reason for the limited success of a 30 year long effort to standardize laboratory animals? *Laboratory animals, 24 (1)*, 71-77.

Georgas, J., van de Vijver, F.J.R., Weiss, L.G., & Saklofske, D.H. (2003). A cross-cultural analysis of the WISC-III. In J. Georgas, L. G. Weiss, F.J.R. van de Vijver, & D. H. Saklofske (Eds.), *Culture and children's intelligence. Cross-cultural analysis of the WISC-III,* pp. 277-313. Amsterdam: Academic Press.

Gignac, G. E. (2006). Evaluating subtest 'g' saturation levels via the Single Trait-Correlated Uniqueness (STCU) SEM approach: Evidence in favour of crystallized subtests as the best indicators of 'g'. *Intelligence*, *34*, 29-46.

Gillespie, N. A., & Neale, M. C. (2006). A finite mixture model for genotype and environment interactions: Detecting latent population heterogeneity. *Twin Research and Human Genetics, 9,* 412-423.

Gilmore, R. (1981). *Catastrophe theory for scientists and engineers.* New York: Dover.

Ghosh, J., Chang, H.-J. & Liano, K. (1998). A Macroscopic Model of Oscillation in ensembles of inhibitory and excitatory neurons. In Omidvar, O. and Dayhoff, J. (eds.), *Neural Networks and Pattern Recognition*, pp. 143-169. San Diego: Academic Press.

Gosso, M. E. (2007). *Common genetic variants underlying cognitive ability*. Unpublished doctoral dissertation. Vrije Universiteit Amsterdam, the Netherlands

Gray, J. R., Thompson, P. M. (2004). Neurobiology of intelligence: science and ethics. *Nature Reviews Neuroscience, 5 (6)*, 471-482.

Guckenheimer, J., & Holmes, P. (2002). *Nonlinear Oscillations, Dynamical Systems and Bifurcations of Vector Fields*. New York: Springer.

Guilford, J. P. (1964). Zero correlations among tests of intellectual abilities. *Psychological Bulletin, 61*, 401-404.

Guilford, J. P. (1980). Fluid and Crystallized Intelligences: Two fanciful concepts. *Psychological Bulletin, 88 (2)*, 406-412.

Gustafsson, J. E. (1984). A unifying model for the structure of intellectual abilities. *Intelligence, 8*, 179-203.

Gustafsson, J. E. (1988). Hierarchical models of individual differences in cognitive abilities. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 35-71). Hillsdale, NJ7 Erlbaum.

Hakstian, A. R., & Cattell, R. B. (1975). *The Comprehensive Ability Battery*. Champaign, IL: Institute for Personality and Ability Testing.

Harris, J. R. (1995). Where is the child's environment? A group socialization theory of development. *Psychological Review, 102*, 458-489.

Haworth, C. M. A., Wright, M. J., Luciano, M., Martin, N. G., de Geus, E. J. C., van Beijsterveldt, C. E. M., Bartels, M., Posthuma, D., Boomsma, D. I., Davis, O. S. P., Kovas, Y., Corley, R. P., DeFries, J. C., Hewitt, J. K., Olson, R. K., Rhea, S. A., Wadsworth, S. J., Iacono, W. G., McGue, M., Thompson, L. A., Hart, A., Petrill, S. A., Lubinski, D., & Plomin, R. (2009). The heritability of general cognitive ability increases linearly from childhood to young adulthood. *Molecular Psychiatry, 15*, 1112-1120.

Helms-Lorenz, M., van der Vijver, F. J. R., & Poortinga, Y. H. (2003). Cross-cultural differences in cognitive performance and Spearman's hypothesis: *g* or *c*? *Intelligence, 31*, 9-29.

Hoekstra, R. A., Bartels, M., & Boomsma, D. I. (2007). Longitudinal genetic study of verbal and nonverbal IQ from early childhood to young adulthood. *Learning Individual Differences, 17*, 97-114.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review, 75*, 242-259.

Horn, J. L. (1970). Organization of data on life-span development of human abilities. In L.R. Goulet & P. B. Baltes (Eds.), Life-span developmental psychology: Research and theory (pp. 423-466). New York: Academic Press.

Horn, J. L. (1980) Concepts of intellect in relation to learning and adult development. Intelligence, 4, 285-317.

Horn, J. L. (1991). Measurement of intellectual capabilities: A review of theory. In: K.S. McGrew, J.K. Werder and R.W. Woodcock, Editors, WJ-R technical manual, DLM, Allen, TX.

Horn, J. L., & Blankson, N. (2005). Foundation for better understanding cognitive abilities. In D. P. Flanagan & P. Harrison (Eds.), *Contemporary intellectual assessment* (pp. 41-68), 2nd Ed. New York: Guilford Press.

Horn, J. L., & Cattell, R. B. (1967). Age differences in fluid and crystallized intelligence. *Acta Psychologica, 26*, 107-129.

Horn, J. L., & Noll, J. (1997). Human cognitive capabilities: Gf-Gc theory. In D. P. Flanagan, J. L. Genshaft, & P. L. Harrison (Eds.), *Contemporary intellectual assessment: Theories, tests, and issues* (pp. 53-91). New York: Guilford Press.

Horn, J. L., & Stankov, L. (1982). Auditory and visual factors of intelligence. Intelligence, 6, 165-185.

Hubel, D. H., &Wiesel, T. N. (1962). Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *Journal of Physiology, 160*, 106-154.

Hunt, E. B. (2000). Let's hear it for crystallized intelligence. Learning and Individual Differences, 12 (1), 123-130.

Hunt, E. & Sternberg, R. J. (2006). Sorry, wrong numbers: An analysis of a study of a correlation between skin color and IQ, *Intelligence, 34*, pp. 131-139.

Jablonka, E. & Lamb, M. J. (2002). The changing concept of epigenetics. *Annals of the New York Academy of Sciences, 981*, 82-96.

Jacobs, N., Van Gestel, S., Derom, C., Thiery, E., Vernon, P., Derom, R., et al. (2001). Heritability estimates of intelligence in twins: Effect of chorion type. *Behavior Genetics, 31*, 209-217.

Jensen, A. R. (1968). Patterns of mental ability and socioeconomic status. *Proceedings of the National Academy of Sciences, 60*, 1330-1337.

Jensen, A. R. (1973). *Educational Differences*. Methuen. London.

Jensen, A. R. (1976). Test bias and construct validity. *Phi Delta Kappa, 58*, 340-346.

Jensen, A. R. (1980). *Bias in mental testing*. New York: Free Press.

Jensen, A. R. (1985). The nature of the black-white difference on various psychometric tests: Spearman's hypothesis. *Behavioral and Brain Sciences, 8*, 193-263.

Jensen, A. R. (1987a). Further evidence for Spearman's hypothesis concerning the black-white differences on psychometric tests. *Behavioral and Brain Sciences, 10*, 512-519.

Jensen, A. R. (1987b). The *g* beyond factor analysis. In R. R. Ronning, J. A. Glover, J. C. Conoley, & J. C. Witt (Eds.), *The influence of cognitive psychology on testing* (pp. 87-142). Hillsdale, NJ: Erlbaum.

Jensen, A. R. (1998). *The g Factor: The Science of Mental Ability*. London: Praeger.

Jensen, A. R. (2001). Vocabulary and general intelligence. *Behavioral and Brain Sciences, 24*, 1109-1110.

Johnson, W., & Bouchard, T. J. (2005). The structure of human intelligence: it is verbal, perceptual, and image rotation (VPR), not fluid and crystallized. *Intelligence, 33*, 393-416.

Johnson, W., Bouchard, T. J., Jr., Krueger, R. J., McGue, M., & Gottesman, I. J. (2004). Just one *g*: Consistent results from three test batteries. *Intelligence, 32*, 95-107.

Johnson, W., & Gottesman, I. I. (2006). Clarifying process versus structure in human intelligence: Stop talking about fluid and crystallised. Behavioral and Brain Sciences, 29, 136-137.

Johnson, W., te Nijenhuis, J., & Bouchard, T. J. (2008). Still just 1 *g*: consistent results from five test batteries. *Intelligence*, *36*, 81-95.

Johnson, W., Bouchard, T. J., McGue, M., Segal, N. L., Tellegen, A., Keyes, M., & Gottesman, I. I. (2007). Genetic and environmental influences on the verbal-perceptual-image rotation (VPR) model of the structure of mental abilities in the Minnesota study of twins reared apart. *Intelligence, 35*, 542-562.

Kan, K.-J., Ploeger, A., Raijmakers, M. E. J., Dolan, C. V., & van der Maas, H. L. J. (2011). Nonlinear Epigenetic Variance: Review and Simulations. *Developmental Science, 13(1)*, 11-27

Kan, K.-J., Kievit. R. A., Dolan, C. V., & van der Maas, H. L. J. (2011). On the interpretation of the CHC factor Gc. *Intelligence, 39*, 292-302.

Kandel, E. R., Schwartz, J. H. & Jessel, T. M., (2000), *Principles of Neural Science (4th Ed.)*. New York: McGraw-Hill.

Keith, T. Z. (2005). Using confirmatory factor analysis to aid in understanding the constructs measured by intelligence tests. In D. P. Flanagan & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests and issues (pp. 581-614). (2nd edition). New York, NY: Guilford Press.

Keith, T. Z., & Reynolds, M. R. (2010). Cattell-Horn-Carroll abilities and cognitive tests: What we've learned from 20 years of research. *Psychology in the Schools, 47(7)*, 635-650.

Kendler, K. S., Heath, A. C., Martin, N. G., & Eaves, L. J. (1987). Symptoms of Anxiety and Symptoms of Depression. Same Genes, Different Environments? *Archives of General Psychiatry, 44*, 451-457.

Kievit, R. A., Romeijn, J. W., Waldorp, L. J., Wicherts, J. M., Scholte, H. S. & Borsboom, D. (2011). Mind the gap: A psychometric approach to the reduction problem. *Psychological Inquiry, 22(2)*, 67-87.

Kimble, G.A. (1993). Evolution of the nature-nurture issue in the history of psychology. In R. Plomin & G. E. McClearn (Eds.), *Nature, nurture and psychology, pp. 3-26.* Washington D.C.: American Psychological Association.

Kelso, J. A. S. (1995). *Dynamic Patterns: The Self-Organization of Brain and Behavior.* Cambridge, Mass.: MIT Press.

Kirillov, A. B., Myre, C. D., & Woodward, D. J. (1991). Bistable neurons and memory patterns in the inhibitory-feedback model inspired by neostriatum. *Society for Neuroscience Abstracts, 17*, 51.

Kirillov, A. B., Myre, C. D., & Woodward, D. J. (1993). Bistability, switches and working memory in a two-neuron inhibitory-feedback model. *Biological Cybernetics, 68*, 441-444.

Kowner R. (2001). Psychological perspective on human developmental stability and fluctuating asymmetry: Applications, sources and implications. *British Journal of Psychology, 92*, 447-469.

Kovas, Y. & Plomin, R. (2006). Generalist genes: implications for the cognitive sciences. *Trends in Cognitive Science, 10*, 198-203.

Kranzler, J. H., & Jensen, A. R. (1991). The nature of psychometric g: Unitary process or a number of independent processes? *Intelligence, 15*, 397-422.

Kvist, A. V., & Gustafsson, J. E. (2008). The relation between fluid intelligence and the general factor as a function of cultural background: A test of Cattell's investment theory, Intelligence, 36, 422-436.

LaBuda, M.-C., DeFries, J. C., & Fulker, D.-W. (1987). Genetic and environmental covariance structures among WISC-R subtests: A twin study. *Intelligence, 11*, 233-244.

Law, J. G. Jr. & Faison, L., (1998). WISC-III and KAIT results in adolescent delinquent males. Journal of clinical psychology, 52 (6), 699 -703.

Lee, K. H., Choi, Y. Y., & Gray, J. R. (2007).What about the neural basis of crystallized intelligence? Behavioral and Brain Sciences, 30, 159-161.

Lewontin, R. C. (1970). Race and intelligence. *Bulletin of the Atomic Scientists, 26*, 2-8.

Lewontin, R. C. (1974). The analysis of variance and the analysis of causes. *American Journal of Human Genetics, 26*, 400-411.

Luo, D., Petrill, S.-A., & Thompson, L.-A. (1994). An exploration of genetic *g*: Hierarchical factor analysis of cognitive data from the Western Reserve Twin Project. *Intelligence, 18*, 335-347.

Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford: Oxford University Press.

Malykh, S. B., Iskoldsky, N. V., & Gindina, E. D. (2005). Genetic analysis of IQ in young adulthood: A Russian twin study. *Personality and Individual Differences, 38*, 1475-1485.

Marks, D. F. (2010) IQ Variations across time, race, and nationality: An artifact of differences in literacy skills. *Psychological Reports, 106*, 643-664.

McArdle, J. J., & Goldsmith, H. H. (1990). Alternative Common Factor Models for Multivariate Biometric Analyses. *Behavior Genetics, 20(5)*, 569-608.

McArdle, J. J., Ferrer-Caja, E., & Hamagami, F. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. Developmental Psychology, 38, 115-142.

McGrew, K.S. (1997). Analysis of the major intelligence batteries according to a proposed comprehensive Gf-Gc framework. In D.P. Flanagan, J.L. Genshaft, & P.L. Harrison (Eds.), Contemporary intellectual assessment: Theories, tests, and issues (pp. 131-150). New York: Guilford.

McGrew, K. S. (2009). CHC theory and the human cognitive abilities project: Standing on the shoulders of the giants of psychometric intelligence research. *Intelligence*, *37*, 1-10.

McGrew, K. S., & Evans, J. J. (2004). Internal and external factorial extensions to the Cattell-Horn-Carroll (CHC) theory of cognitive abilities: A review of factor analytic research since Carroll's seminal 1993 treatise (Carroll Human Cognitive Abilities Project Research Report No. 2). Retrieved February 1, 2010, from the Institute for Applied Psychometrics Website: http://www.iapsych.com/ carrollproject.htm

McGrew, K. & Flanagan, D. (1998). *The Intelligence Test Desk Reference: Gf-Gc cross-battery assessment*. Boston: Allyn & Bacon.

McGrew, K. S., & Woodcock, R. W. (2001). Technical manual. Woodcock-Johnson III. Itasca, IL: Riverside Publishing.

McGue, M. & Bouchard, T. J. (1998). Genetic and environmental influences on human behavioral differences. *Annual Review of Neuroscience, 21,* 1-24.

Meinhardt, H. (1982). *Models of biological pattern formation*. London: Academic Press.

Mingroni, M. A. (2007). Resolving the IQ paradox: Heterosis as a cause of the Flynn effect and other trends". *Psychological Review, 114*, 806-829.

Moffitt, T. E., Caspi, A., & Rutter, M. (2005). Interaction between measured genes and measured environments: A research strategy. *Archives of General Psychiatry, 62*, 473-481.

Molenaar, P. C. M., & Raijmakers, M. E. J. (1999). Additional aspects of third source variation for the genetic analysis of human development and behaviour: A commentary on Eaves et al. *Twin Research, 2(1),* 49-52.

Molenaar, P. C. M., & Raijmakers, M. E. J. (2000). A causal interpretation of Piaget's theory of cognitive development: Reflections on the relationship between epigenesis and nonlinear dynamics. *New Ideas in Psychology, 18 (1),* 41-55.

Molenaar, P. C. M., Boomsma, D. I., & Dolan, C. V. (1993). A third source of developmental differences. *Behavior Genetics, 23,* 519-524.

Mountcastle, V. (1957). Modality and topographic properties of single neurons of cat's somatic sensory cortex. *Journal of Neurophysiology, 20*, 408-434.

Muthén, B., Asparouhov, T., & Rebollo, I. (2006). Advances in behavioral genetics modeling using Mplus: Applications of factor mixture modeling to twin data. *Twin Research and Human Genetics, 9*, 313-328.

Nijenhuis, J. te, Resing, W. C. M., Tolboom, E., & Bleichrodt, N. (2004). Short-Term Memory as an Additional Predictor of School Achievement for Immigrant Children? *Intelligence, 32*, 203-213.

Owen, D. R., & Sines, J. O. (1970). Heritability of personality in children. *Behavior Genetics, 1*, 235-248.

Oyama, S. (1985). *The Ontogeny of Information, Developmental Systems and Evolution.* New York: Cambridge University Press.

Pedersen, N. L., Plomin, R., Nesselroade, J. R., & McClearn, G. E. (1992). A quantitative genetic analysis of cognitive abilities during the second half of the life span. *Psychological Science, 3*, 346-353.

Penke, L., Denissen, J. J. A., & Miller, G. F. (2007). The evolutionary genetics of personality. *European Journal of personality, 21*, 549-587.

Petrill, S. A. (1997). Molarity versus modularity of cognitive functioning? A behavioral genetic perspective. *Current Directions in Psychological Science, 6*, 96-99.

Petzoldt, T. (2003). R as a simulation platform in ecological modeling. *R News, 3*, 8-16.

Ploeger, A., van der Maas, H. L. J., & Raijmakers, M. E. J. (2008). Is evolutionary psychology a metatheory for psychology? A discussion of four major issues in psychology from an evolutionary developmental perspective. *Psychological Inquiry, 19*, 1-18.
Plomin, R. (2003). Molecular genetics and *g*. In H. Nyborg, *The Scientific Study of General Intelligence: Tribute to Arthur R. Jensen*. Oxford: Pergamon Press.

Plomin, R., Spinath, F. M. (2002). Genetics and general cognitive ability (*g*). *Trends in Cognitive Science, 6*, 169-176.

Plomin, R., Asbury, K., & Dunn, J. (2001). Why are children in the same family so different? Nonshared environment a decade later. *Canadian Journal of Psychiatry, 46*, 225-233.

Plomin, R., DeFries, J. C., McClearn, G. E., & McGuffin, P. (2008). *Behavioral genetics* (5th ed.). New York: Worth.

Plomin, R., DeFries, J. C., & Loehlin, J. C. (1977). Genotype-environment interaction and correlation in the analysis of human behavior. *Psychological Bulletin, 84 (2)*, 309-322.

Pollack, J. B. (1991). The induction of dynamical recognizers. *Machine Learning, 7,* 227-252.

Prigogine, I. (1980). *From being to becoming: Time and complexity in the physical sciences*. San Francisco: Freeman.

Purcell, S. (2002). Variance components models for gene-environment interaction in twin analysis. *Twin Research, 5*, 554-571.

Raijmakers, M.E.J., & Molenaar, P.C.M. (2004). Modeling developmental transitions in adaptive resonance theory. *Developmental Science, 7 (2)*, 149-157.

Reynolds, M. R., & Keith, T. Z. (2007). Spearman's law of diminishing returns in hierarchical models of intelligence for children and adolescents. Intelligence, 35(1), 267-281.

Rijsdijk, F. V., Vernon, P. A., & Boomsma, D. I. (1998). The genetic basis of the relation between speed-of-information processing and IQ. *Behavioural Brain Research, 95*, 77-84.

Rijsdijk, F. V., Vernon, P. A., & Boomsma, D. I. (2002). Application of hierarchical genetic models to Raven and WAIS subtests: A Dutch twin study. *Behavior Genetics, 32*, 199-210.

Roberts, R. D., Goff, G. N., Anjoul, F., Kyllonen, P. C., Pallier, G., & Stankov, L. (2000). The Armed Services Vocational Aptitude Battery: Little more than acculturated learning (Gc)!? *Learning and Individual Difference*, *12*, 81-103.

Rodriguez, P., Wiles, J., & Elman, J. L. (1999). A recurrent neural network that learns to count. *Connection Science, 11 (1*), pp. 5-40.

Rushton, J. P. (1995). *Race, evolution, and behavior: A life history perspective.* New Brunswick, NJ: Transaction Publishers.

Rushton, J.P. (1998) The 'Jensen Effect' and the 'Spearman-Jensen Hypothesis' of Black-White IQ differences. *Intelligence, 26*, 217-225.

Rushton, J. P., & Jensen, A. R. (2005). Thirty years of research on race differences in cognitive ability. *Psychology, Public Policy, and Law, 11*, 235-294.

Rushton, J. P. & Jensen, A. R. (2010a). The rise and fall of the Flynn Effect as a reason to expect a narrowing of the Black-White IQ gap. *Intelligence, 38*, 213-219.

Rushton, J. P., & Jensen, A. R. (2010b). Race and IQ: A theory-based review of the research in Richard Nisbett's Intelligence and How to Get It. *The Open Psychology Journal, 3*, 9-35.

Rutter, M., Dunn, J., Plomin, R., Simonoff, E., Pickles, A., Maughan, B., Ormel, J., Meyer, J., & Eaves, L. (1997). Integrating nature and nurture: Implications of person-environment correlations and interactions for developmental psychopathology. *Development and Psychopathology, 9*, 335-364.

Sameroff, A. J., Seifer, R., Baldwin, A., & Baldwin, C. (1993). Stability of intelligence from preschool to adolescence: The influence of social and family risk factors. *Child Development, 64*, 80-97.

Scarr, S., & McCartney, K. (1983). How people make their own environments: A theory of genotype → environment effects. *Child Development, 54*, 424-435.

Schaie, K. W., & Strother, C. R. (1998). A cross-sequential study of age changes in cognitive behavior. *Psychological Bulletin, 70*, 671-680.

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research, 8(2)*, 23-74.

Schmid, J., & Leiman, J. N. (1957). The development of hierarchical factor solutions. *Psychometrika, 22,* 53-61.

Schweizer, K., & Koch, W. (2002). A revision of Cattell's investment theory: Cognitive properties influencing learning. *Learning and Individual Differences, 13*, 57-82.

Segal, N.-L. (1985). Monozygotic and dizygotic twins: A comparative analysis of mental ability profiles. *Child Development, 56*, 1051-1058.

Smith, D. D. (1993). Brain, environment, heredity, and personality. *Psychological Reports*, 72(1), 3-13.

Snow, R. E.(1996). Aptitude development and education. *Psychology, Public Policy, and Law*, *3* (4), 536-60.

Snow, R. E., Kyllonen, C. P., & Marshalek, B. (1984). The topography of ability and learning correlations. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence* (pp. 47-103). Hillsdale, NJ: Erlbaum.

Spearman, C. (1904). General intelligence objectively determined and measured. *American Journal of Psychology, 15*, 201-293.

Spearman, C. (1927). *The abilities of man: Their nature and measurement.* New York: Macmillan.

Stankov, L. (2000). Structural extensions of a hierarchical view on human cognitive abilities. Learning and Individual Differences, 12(1), 35-51.

Steinmetz, H. (1996). Structure, function and cerebral asymmetry - in vivo morphometry of the planum temporale. *Neuroscience and Biobehavioral Reviews, 20*, 587-591.

Stent, G. S. (1978). *Paradoxes of Progress*. San Francisco: Freeman.

Sternberg, R. J., Grigorenko, E. L., & Bundy, D. A. (2001). The predictive value of IQ. *Merrill-Palmer Quarterly, 47*.(1), 1-41.

Stige, L. C., Slagsvold, T., Vøllestad, L. A. (2005). Individual fluctuating asymmetry in pied flycatchers (Ficedula hypoleuca) persists across moults, but is not heritable and not related to fitness. *Evolutionary Ecology Research, 7*, 381-406.

Tambs, K., Sundet, J.-M., & Magnus, P. (1984). Heritability analysis of the WAIS subtests: A study of twins. *Intelligence, 8*, 283-293.

Thomson, G. H. (1919). On the cause of hierarchical order among correlation coefficients. *Proceedings of the Royal Society*, *95*, 400-408.

Thomson, G. H. (1951). *The factorial analysis of human ability* (5th ed.). London: University of London Press.

Thorndike, E. L. (1903). *Educational psychology*. New York: Lemcke & Buechner.

Thorndike, R. L. (1927). The measurement of intelligence. New York: Teachers College.

Thoma, R. J., Yeo, R. A., Gangestad, S. W., Lewine. J., & Davis, J. (2002).Fluctuating asymmetry and the human brain. *Laterality, 7*, 45-58.

Thurstone, L.L. (1938). *Primary mental abilities*. Chicago: University of Chicago Press.

Townsend, G. C., Richards, L., Hughes, T., Pinkerton, S., & Schwerdt, W. (2005). Epigenetic influences may explain dental differences in monozygotic twin pairs. *Australian Dental Journal, 50(2),* 95-100.

Turkheimer, E. (2000). Three laws of behavior genetics and what they mean. *Current Directions in Psychological Science, 9(5),* 160-164.

Turkheimer, E. (2004). Spinach and ice cream: Why social science is so difficult. In L.F. DiLalla (Ed.), *Behavior genetics principles: Perspectives in development, personality and psychopathology*. Washington D.C.: American Psychological Association.

Turkheimer, E. & Gottesman, I. I. (1996). Simulating the dynamics of genes and environment in development. *Development and Psychopathology, 8,* 667-677.

Turkheimer, E. & Waldron, M. (2000). Nonshared environment: A theoretical, methodological, and quantitative review. *Psychological Bulletin, 126(1)*, 78-108.

Turkheimer, E., D'Onofrio, B.M., Maes, H.H., Eaves, L.J. (2005). Analysis and interpretation of twin studies including measures of the shared environment. *Child Development, 76 (6),* 1217-1233.

Vandenberg, S, G. (1968). The nature and nurture of intelligence. In Glass, D. C, (ed.), *Biology and Behavior: Genetics* (pp. 3-58). New York: Rockefeller University Press.

van der Maas, H. L. J., & Molenaar, P. C. M. (1992). Stage-wise. Cognitive development: an application of catastrophe theory. *Psychological Review, 99 (3)*, 395-417.

van der Maas, H. L. J., Verschure, P. F. M. J., & Molenaar, P. C. M. (1990). A note on chaotic behavior in simple neural networks. *Neural Networks, 3*, 119-122.

van der Maas, H. L. J., Dolan, C. V., Grasman, R. P. P. P., Wicherts, J. M., Huizenga, H. M., & Raijmakers, M. E. J. (2006). A dynamical model of general intelligence: The positive manifold of intelligence by mutualism. *Psychological Review, 113*, 842-861.

van der Sluis, S., Kan, K. J., & Dolan C. V. (2010). Consequences of a network view for genetic association studies. *Behavioral and Brain Sciences, 33 (2),* 173-174.

van der Sluis S., Willemsen, G., de Geus, E. J., Boomsma, D. I., & Posthuma, D. (2008). Gene-environment interaction in adults' IQ scores: measures of past and present environment. *Behavior Genetics, 38*, 348-360.

Vanderwood, M. L., McGrew, K. S., Flanagan, D. P., & Keith, T. Z. (2001). The contribution of general and specific cognitive abilities to reading achievement. Learning & Individual Differences, 13, 159-188.

van Fraassen, B. C. (1980). The scientific image. Oxford, England: Clarendon

van Ooyen, A. (1994). Activity-dependent neural network development. *Network, 5,* 401-423.

van Ooyen, A., & van Pelt, J. (1994). Activity-dependent outgrowth of neurons and overshoot phenomena in developing neural networks. J*ournal of Theoretical Biology, 167*, 27-43.

van Ooyen, A., & van Pelt, J. & Corner, M. (1995). Implications of activity-dependent neurite outgrowth for neuronal morphology and network development. J*ournal of Theoretical Biology, 172*, 63-82.

van Oss, C., van Ooyen, A. (1997). Effects of inhibition on neural network development through activity-dependent neurite outgrowth. *Journal of Theoretical Biology, 185*, 263-280.

Waddington, C. H. (1957). *The Strategy of the Genes*. London : Allen & Unwin.

Williams, F. (1975). Family resemblance in abilities: The Wechsler Scales. *Behavior Genetics, 5*, 405-409.

Wechsler D. (1949). *Wechsler Intelligence Scale for Children (WISC)*. New York: Psychological Corporation.

Wechsler, D. (1955). *Manual for the Wechsler Adult Intelligence Scale*. New York: Psychological Corporation.

Wechsler, D. (1974). *Manual for the Wechsler Intelligence Scale for Children-Revised*. New York: Psychological Corporation.

Wechsler, D. (1981). *Manual for the Wechsler Adult Intelligence Scale- Revised*. New York: Psychological Corporation.

Wechsler, D. (1991a). *Manual for the Wechsler Intelligence Scale for Children-Third Edition*. San Antonio, TX. The Psychological Corporation.

Wechsler, D. (1991b). *WISC-III UK. Manual*. London: Psychological Corporation Limited.

Wechsler, D. (1997). *Wechsler Adult Intelligence Scale-Third Edition: administration and scoring manual*. San Antonio, TX: Psychological Corporation.

Wechsler, D. (2002*). WISC-III NL. Handleiding*. London: Psychological Corporation Limited.

Wechsler, D. (2005). *WAIS-III Nederlandstalige bewerking. Technische handleiding*. Amsterdam: Hartcourt Assessment B.V.

Wilson, H. R., & Cowan, J. D. (1972). Excitatory and inhibitory interactions in localized populations of model neurons. *Biophysical Journal, 12*, 1-24.

Wilson, J. R., De Fries, J. C., McClearn, G. E., Vandenberg, S. G., Johnson, R. C., & Rashad, M. N. (1975). Cognitive abilities: Use of family data as a control to assess sex and age differences in two ethnic groups. *International Journal of Aging and Human Development, 6*, 261-276.

Woodcock, R.W. (1994). Extending Gf-Gc theory into practice. In J. C. McArdle& R. W. Woodcock (Eds.), *Human cognitive abilities in theory and practice* (pp. 137-156). Mahwaw, NJ: Lawrence Erlbaum Associates.

Wright, I. C., Sham, P., Murray, R. M., Weinberger, D. R., & Bullmore, E. T. (2002). Genetic contributions to regional variability in human brain structure: Methods and preliminary results. *Neurolmage, 17*, 256-271.

SUMMARY IN ENGLISH / SAMENVATTING IN HET ENGELS

Better Theory

This thesis stressed the importance of the development of better theory in intelligence research. In addition, it illustrated that taking explicit scientific philosophical standpoints (e.g. realist or nonrealist) with respect to intelligence, and the variables that relate to it, helps to understand empirical results. Without adequate theory and without researchers' philosophical standpoints, it is extremely difficult to make sense of the intelligence literature, which can hinder scientific progress in intelligence research.

In the introduction we explained why it is so difficult to make sense of the intelligence literature. First, we pointed out that intelligence relates to how well systems process information and that - when we restrict ourselves to human intelligence - the level of intelligence is usually determined using psychometric tests via ordinal scaling. Next, it was made clear that without further commitment to any specific theory that relates intelligence to a quantitative property (or multiple quantitative properties), intelligence is nothing more than a rank order on intelligence test scores (Bartholomew, 2004), whether these comprise raw scores or scores on variable extracted from those raw scores. The problem that arises next, is that different rank orders can be made, so that based on the same data one researcher can (legitimately) conclude that intelligence grows, hence changes throughout development, while another researcher can (also legitimately) conclude that intelligence is stable over time and does not change at all (see Figure 1.1). What conclusion is drawn in practice largely depends on the scientific perspective of the researcher. With scientific perspective we mean an intra-individual perspective, which is common in developmental psychology, or an interindividual perspective, which is common in differential psychology. We stressed that the causes of inter-individual differences can be entirely different from the causes of intra-individual differences. In order to understand, explain, and model intelligence well, we need both perspectives. The interindividual perspective is needed to explain the covariance structure of intelligence test scores, the intraindividual perspective to explain cognitive growth. We noted that the majority of studies into intelligence lacks the developmental perspective and that discussions and theories are mainly concerned with (statistical) descriptions of individual differences according to factor models.

Aims

The aims of the thesis were twofold. The first was to reinvigorate the development of an adequate theory of intelligence by providing a model that accounts for both cognitive growth and (heritable) interindividual differences in intelligence. This model had to be able to explain salient findings in intelligence research, such as a correlation between intelligence subtests' heritability coefficients and their loadings on the (statistical) general factor of intelligence. While developing such model, we encountered theoretical issues that were not fully addressed in the literature. They concern mostly the interpretation of the heritability of intelligence. The second aim was to address these issues in more detail.

Working Hypothesis

Throughout the thesis we provisionally accepted the Cattell-Horn-Carroll (CHC) model as a working hypothesis. Taking a realist scientific philosophical position concerning the variables of intelligence, we hypothesized that the second order factors in this model (including Gf and Gc) represent individual differences in unique cognitive systems, which are constrained by genetically and environmentally influenced capacities. We did not posit a substantive underlying general factor, because we believed positive intercorrelations among the cognitive systems are due to reciprocal interactions among those systems that occur throughout cognitive development, as in the mutualism model (see Figure 1.8). We also assumed that in principle individual differences in these systems can

be measured by intelligence tests, although we maintained that in practice intelligence tests are likely not uni-dimensional (see Figure 7.1).

Summary per Chapter

In the introduction we pointed out that the dynamical interplay between genes and environment is largely ignored in the statistical (cross sectional) modeling of heritable intelligence. This makes that the further interpretation of the statistical results requires caution. This is so for at least two reasons. First, whereas the hypothesized underlying variables in statistical models are generally modeled as having linear effects, dynamical systems - such as humans - often display behavior that is nonlinear. If non-linearity is present, a linear statistical model does thus not represent the true data generating mechanism. Second, dynamical gene-environment interplay can lead to gene-environment correlation, which is usually not modeled in statistical models of intelligence. Both the presence of non-linearity and gene-environment correlation affect heritability estimates. How or to what extent heritability estimates of intelligence are affected by these effects is unknown and requires research.

Chapter 2 dealt with the problem of nonlinearity. We noted that the genetic and environmental variables in behavior genetic models are generally not measured, but merely inferred, using genetically informative research designs (e.g. twin and adoption studies), and that these genetic and environmental variables are modeled as having linear relationships with the observed variables, e.g. cognitive abilities measured by IQ tests. However, in reality cognitive development is typically nonlinear and is characterized by stage transitions (see Table 1). By means of computer simulations, we investigated the effects on estimated heritabilities and environmentalities in statistical linear behavior genetic modeling in case the underlying mechanisms were nonlinear. In doing so, we used the two-cell model of van Oss & van Ooyen (1996). Although originally the two cells or units represent an inhibitory and an excitatory neuron, they were interpreted as representing inhibitory and excitatory networks constituting a neural system (van Oss & van Ooyen, 1996), for example a system that underlies working memory. We implemented heritable individual differences in the parameters of the model, ran a series of simulations, and investigated whether the estimated genetic and environmental structure resembled the implemented structure.

In the results of our simulations, we observed effects that are in line with empirical findings in intelligence research (see Table 7.1). First, when nonlinearity was present, estimated heritabilities appeared to increase over time, so that the estimated relative contribution of environmental influences decreased. The heritabilities were overestimated, sometimes by as much as 25%. Second, and more particular, the estimated relative effects of shared environmental influences decreased (often to values of zero) rather than those of nonshared influences. Third, in the behavior genetic modeling the ultimate underlying causes were hard to detect, which fits the fact that to date the search for specific environmental and genetic influences on intelligence has been largely unsuccessful. We concluded that within the linear statistical framework, nonlinear development can be considered as constituting a source of phenotypic variance on its own, next to the genetic and environmental sources (see also Molenaar et al., 1993). As in reality cognitive growth is characterized by stage transitions, we hope new mutualism models will be developed that can account for such transitions. We also hope that such models will be used to investigate the role of nonlinear development as a third source of variance in intelligence. Since the presence of stage transitions can lead to an increasing overestimation of heritability coefficients over time, this provides an additional or alternative explanation to the mainstream explanation of the increase of heritability. Usually this increase is explained in terms of increasing gene-environment correlation.

Gene-environment correlations also constituted our explanation of the results discussed in Chapter 3, in which we summarized data from 23 independent behavior genetic studies into intelligence. These data showed the contra-intuitive result that the most cultural dependent intelligence tests (i.e. knowledge tests) demonstrated higher heritabilities than the least cultural dependent tests (i.e. cognitive processing tests). Notably, these cultural dependent knowledge tests also showed the relatively highest *g* loadings (see Table 1). In isolation, mainstream theories of intelligence fail to explain this combined effect (see Chapter 6). Some of them even predict the opposite. We hypothesized that if society puts demands on acquired knowledge rather than cognitive

processing, gene-environment correlations would become higher for individual differences in knowledge than individual differences in cognitive processing. In standard behavior genetic modeling gene-environment correlations will be subsumed under the genetic variance, leading to higher heritability estimates.

On the basis of the results from Chapter 2, we can also formulate an alternative or additional hypothesis. As was shown, stage transitions in cognitive development can result in overestimation of the true heritability. The development of certain cognitive abilities may be characterized by more transitions than the development of other cognitive abilities, so that the increase in heritability due to nonlinear effects will differ across abilities. For example, verbal comprehension, usually measured by knowledge tests, may develop stage wise, whereas nonverbal cognitive abilities, usually measured by cognitive processing tests, develop gradually (whether this is the case should be investigated empirically). Hence, the heritability coefficients of the verbal tests can be overestimated, while those of nonverbal tests are correct or less affected. Whether this alternative explanation of increasing heritability can account for the high *g* loadings of verbal knowledge tests (see Table 7.1) is unknown, but can be investigated with a nonlinear model of general intelligence. Future simulations with a mutualism model that includes nonlinear terms may help to discriminate between the two explanations.

Both explanations of the increase in heritability are also relevant in the discussion of racial and ethnic group differences. Black-White differences in IQ exist, for example (see Table 7.1). Chapter 4 showed that they are the most pronounced on verbal knowledge tests (see Table 7.1). If verbal comprehension develops nonlinearly, small differences in initial conditions, e.g. small language deficits in early development or differences in language use (vocabulary), can have large consequences for the further acquisition of knowledge, hence for development of intelligence and scholastic and academic achievement, especially when selection to educational systems is based largely on knowledge and when transmission of knowledge (teaching) takes place verbally. In principle, Black-White mean group differences in later IQ can thus be the result of small initial differences in verbal comprehension or language use. To date, the actual causes are unknown, however, although some researchers (Rushton & Jensen, 2010a) have argued that certain empirical results imply Black-White differences are genetic.

As mentioned Mean Black-White differences in IQ are the most pronounced on verbal knowledge tests. These tests also happen to demonstrate the highest heritabilities and *g* loadings (see above). This effect has been interpreted as a genetic effect (Rushton & Jensen, 2010a). In Chapter 4 we demonstrated that the reasoning behind this interpretation was invalid. We showed analytically that group differences can be the most pronounced on the most heritable and the most *g* loaded tests even when the origin is purely environmental. Moreover, as Chapter 3 showed, the most heritable, most *g* loaded tests are also the most cultural loaded. The finding that Black-White differences are the most pronounced on the most culturally loaded subtests are problematic for Rushton and Jensen's (2010a) biological *g* theory, because this theory predicts the opposite: the mean group differences are most pronounced on the least culturally loaded tests.

Cultural loaded knowledge tests are often referred to as crystallized tests, in line with Cattell's investment theory of fluid and crystallized intelligence. Crystallized tests often load highly on a factor that is separate from nonknowledge tests (cognitive processing tests). For this reason, such factor (abbreviated Gc) is often interpreted as 'crystallized intelligence'. In addition, crystallized intelligence is interpreted as the result of the 'investment' of fluid intelligence (reasoning). From our review of the fluid-crystallized theory (Chapter 5), we argued that 1) crystallized intelligence is purely a statistical summary, and that 2) the investment hypothesis of fluid intelligence does not explain why crystallized tests should load on a factor other than the factor fluid intelligence (Gf); 3) a second influence is necessary to dislocate the factor crystallized intelligence from the factor fluid intelligence. Also from the review, we concluded that this second causal influence is most likely verbal comprehension or education. Supported by results from factor analyses, we concluded that once cultural, educational and age differences (hence developmental differences) are taken into account, factor Gc is equivalent to verbal comprehension. We maintained that if researchers find a separate Gc factor, this is due to sample heterogeneity with respect to developmental differences.

Likely, such sample heterogeneity also leads to the separation of *g* and Gf. In homogeneous samples they will be equivalent (see Table 7.1), as we found.

Chapter 6 was a theoretical chapter. We addressed in detail whether and to what extent one of the most intriguing empirical findings poses a problem for current theories of general intelligence That is, we asked ourselves whether *g* theories, investment theories, sampling theories, and reciprocal interaction theories explain the finding that the most cultural dependent cognitive abilities (crystallized abilities) are the most *g* loaded and most heritable. We concluded that (in isolation) they do not. (by implication, the reviewed theories do not explain how Black-White differences become the most pronounced on the most culturally dependent, most heritable, most *g* loaded subtests). To explain this effect, better theory is required. An adequate theory of general intelligence certainly needs to incorporate the developmental perspective.

In Chapter 7 we first summarized the previous chapters. Next we presented an integrated model of general intelligence. It was a mutualism model (van der Maas et al., 2006) that incorporated the main idea of investment theory (individual differences in cognitive processes - fluid abilities - give rise to differences in knowledge and skills - crystallized abilities) and Dickens & Flynn's (2001; Dickens, 2008) social multiplier. In line with the mutualism theory (van der Maas et al., 2006), we assumed that cognitive processing benefits from knowledge acquisition. In the integrated theory, an underlying *g* (Spearman, 1904; Carroll, 1993; Jensen, 1998) was absent. Genetic correlations among limiting capacities can be present, but were taken to be the result of what we denoted genetic sampling (Thompson, 1951; Bartholomew et al. 2009; Anderson, 2001; Penke et al., 2007; see Chapter 6) and not as due to general genetic effects (Kovas & Plomin, 2006). The integrated theory accounted for the fact that individual differences are the most pronounced on the most culturally dependent subtests, which are the most heritable and the most *g* loaded.

General Conclusions and Discussion

The main aim of this thesis was to reinvigorate the development of an adequate theory of intelligence by providing a model that accounts for both cognitive growth and (heritable) interindividual differences in intelligence. The model had to be able to explain salient findings in intelligence research, such as a correlation between intelligence subtests' heritability coefficients and their loadings on the (statistical) general factor of intelligence (current theories of intelligence in isolation do not account for those relationships). To this end, we aimed to integrate hypotheses from different theories of intelligence.

In the last chapter we presented an integrated model of general intelligence. The integrated model accounted for the fact that individual differences are the most pronounced on the most culturally dependent subtests, which are the most heritable and the most *g* loaded. The effect was due to differences in gene-environment effects across cognitive abilities. From our simulation studies with the integrated model, we concluded that the relations among *g* loading, heritability, and cultural load will be positive when the society fosters crystallized abilities rather than fluid abilities.

The simulations also showed that the general factor of intelligence does not have to represent a realistic (e.g. biological) common cause of individual differences. In our integrated model it is a merely a statistical entity, i.e., an informative summary descriptive. However, as such, general intelligence certainly has utility. Specifically, we consider it to be similar to 'general health'. Like general health is an informative summary descriptive of physical functioning, '*g*' is an informative summary descriptive of cognitive functioning.

Nevertheless, a distinction between *g* as a statistically entity, and *g* as interpreted as a realistic, common cause of individual differences, is not only important theoretically, but also empirically. Consider genetic association and linkage studies of intelligence, for example. So far, the search for genes for general intelligence has met with relatively little success (Deary, Johnson, & Houlihan, 2009; Plomin and Spinath, 2004; Chabris et al., *in press*). The alternative theories of Dickens and Flynn (2001; Dickens, 2008) and of van der Maas et al. (2006), in which the general factor of intelligence is a statistical entity originating in reciprocal beneficial interactions among cognitive processes or abilities, are able to provide a plausible explanation of this lack of success. In these theories, there are no direct genetic influences general to all abilities. If general intelligence is

indeed the outcome of such interactions, the search for specific 'genetic influences on *g*' is a questionable undertaking (Dolan, Kan, van der Maas, 2008; van der Sluis, Kan & Dolan, 2010), in particular when (1) the genetic determinants of these processes are under mutation selection balance, which means that the process of natural selection reduces genetic variance but cannot deplete all of it because new variation - due to mutations - is continually reintroduced (Penke et al., 2007), and (2) the interactions are nonlinear (e.g. Molenaar, Boomsma, & Dolan, 1993; Kan, Ploeger, Raijmakers, Dolan & van der Maas, 2010).

Cross-sectional factor analysis cannot discriminate between *g* theories and alternative theories of general intelligence. In addition, both *g* theories and alternative theories can explain the facts that the general intelligence is (highly) heritable and has (strong) biological correlates. So these facts will not discriminate between different theories either. Developmental analyses on the other hand, may have the potential to discriminate between theories of intelligence. For example, group factors of intelligence are more or less genetically independent initially, but become increasingly genetically interrelated as development unfolds (Hoekstra et al., 2007). This observation is consistent with mutualism, but is not readily explained by *g* theories or sampling theories. Developmental effects are thus important in the understanding of individual differences. We hope future (behavior genetic studies) studies into intelligence will use longitudinal modeling. Together with formal, dynamical system modeling, this will enrich intelligence research.

The main points of this thesis are as follows. First, although it is still not possible to determine whether a realistic, underlying *g* is present or not, we can conclude that current *g* theories are inadequate in explaining certain salient empirical findings. Next to the individual differences perspective they have, they need a developmental perspective. The role of the dynamic interplay between genetic and environmental variables that occurs during development needs to be explicated. Second, formal modeling is important in intelligence research. Using the mutualism model can help researchers to study combined effects. The main advantage of the mutualism model is that it is able to combine the intraindividual differences and interindividual differences scientific perspectives on intelligence, as we advanced.

SUMMARY IN DUTCH / SAMENVATTING IN HET NEDERLANDS

Betere Theorie

Dit proefschrift benadrukt het belang van de ontwikkeling van betere theorie in intelligentie-onderzoek. Daarnaast illustreert het dat het innemen van expliciete wetenschappelijk filosofische standpunten ten aanzien van intelligentie, en de variabelen die daarmee samenhangen, empirische resultaten helpt te begrijpen. Wanneer adequate theorie en expliciete wetenschapsfilosofische standpunten ontbreken, is het uiterst moeilijk goed overzicht te krijgen van de intelligentieliteratuur. Het ontbreken van goed overzicht is onwenselijk daar het wetenschappelijke vooruitgang op het gebied van intelligentieonderzoek kan hinderen.

Waarom het verkrijgen van een goed overzicht over de intelligentieliteratuur zo moeilijk is werd uiteengezet in de inleiding. Eerst werd uitgelegd dat het begrip intelligentie aangeeft 'hoe goed informatie verwerkt wordt' en dat - als we ons beperken tot de intelligentie van de mensen - 'hoe goed informatie verwerkt wordt' doorgaans bepaald wordt aan de hand van psychometrische tests, volgens een ordinale schaalindeling. Daarna werd duidelijk gemaakt dat zonder theoretische koppeling van psychometrische intelligentie aan een kwantitatieve eigenschap (of meerdere kwantitatieve eigenschappen), psychometrische intelligentie niets meer is dan een rangorde op intelligentietestscores (Bartholomew, 2004), of dit nu ruwe scores betreft of scores op daaruit geconstrueerde variabelen. Het probleem dat vervolgens ontstaat, is dat verschillende rangorden gemaakt kunnen worden, waardoor op basis van dezelfde gegevens de ene onderzoeker (terecht) zou kunnen concluderen dat intelligentie groeit, en dus verandert tijdens de ontwikkeling, terwijl de andere onderzoeker (eveneens terecht) zou kunnen concluderen dat intelligentie juist stabiel en onveranderlijk is (zie Figuur 1.1). Welke conclusie in praktijk getrokken zal worden, is grotendeels afhankelijk van het wetenschappelijk perspectief van de onderzoeker. Met wetenschappelijk perspectief wordt hier bedoeld een intra-individueel perspectief, zoals gewoon is in de ontwikkelingspsychologie, of een inter-individueel perspectief, zoals gewoon is in de differentiële psychologie. Benadrukt werd dat de oorzaken van inter-individuele verschillen compleet anders kunnen zijn dan de oorzaken van intra-individuele verschillen. Op grond hiervan werd beargumenteerd dat indien we intelligentie goed willen begrijpen, verklaren, en modelleren, we beide perspectieven nodig hebben. Het interindivuele perspectief is nodig om de covariantiestructuur van intelligentietestscores te verklaren; het intra-individuele perspectief is nodig om cognitieve groei te verklaren.

Opgemerkt werd dat in het merendeel van studies naar intelligentie intra-individuele verschillen buiten beschouwing worden gelaten, en dat de blik slechts is gericht op (de statistische beschrijving) van inter-individuele verschillen aan de hand van factormodellen. Door deze eenzijdige blik schieten intelligentietheorieën mogelijk te kort in het adequaat verklaren van empirische resultaten. Evaluatie van zulke theorieën en modellen in het licht van de ontwikkeling van intelligentie werd een belangrijk doel van dit proefschrift. Een ander belangrijk doel was de ontwikkeling van een geïntegreerde intelligentietheorie en van een formeel, dynamisch systeemmodel voor intelligentie dat kan dienen om (zowel oude als nieuwe) hypothesen te toetsen.

De inleiding bevatte ook een overzicht van de belangrijkste factormodellen en theorieën voor psychometrische intelligentie. Deze zullen nu worden behandeld.

Overzicht van de belangrijkste factormodellen en theorieën voor psychometrische intelligentie

De huidige, dominante intelligentietheorieën en -modellen leunen in grote mate op de ideeën van Spearman (1904) die poogde de positieve intercorrelaties tussen verschillende cognitieve tests te verklaren. Deze correlaties waren volgens hem het resultaat van de invloed van een gemeenschappelijke, maar niet geobserveerde variabele, die hij de algemene intelligentiefactor noemde, en die hij afkortte tot *g* (van het Engelse '*general* intelligence'). De correlaties zijn echter niet volmaakt doordat (1) ook testspecifieke variabelen (*s,* van het Engelse '*specific*') van invloed zijn, en (2) meetinstrumenten niet perfect zijn, zodat een deel van de ongedeelde variantie toe te

schrijven is aan meetfouten. De specifieke variabelen werden verondersteld statistisch onafhankelijk te zijn van zowel elkaar als van *g*. Om de *g*-theorie te staven, ontwikkelde Spearman de statistische techniek factoranalyse.

Spearman's *g*-theorie (alsook het gebruik van factoranalyse) heeft van meet af aan veel kritiek te verduren gekregen terwijl onderzoekers in de loop der tijd meerdere valide alternatieve verklaringen voor de positieve correlaties tussen test scores hebben geopperd (Thomson, 1919; 1951; Bartholomew, Deary en Lawn, 2009; van der Maas et al., 2006; Dickens, 2008). De consensus is tegenwoordig dat Spearman's *g*-theorie in overeenstemming is met de gevonden positieve correlaties tussen intelligentietestscores, maar dat factoranalyse niet kan aantonen dat een theorie (dus ook die van Spearman) juist is. Factoranalytische resultaten op zichzelf zijn statistische overzichten van de data. Om aan de factoren een verdere betekenis te hechten, bijvoorbeeld als reële bron van gemeenschappelijke variantie, is theorie vereist (zie bijvoorbeeld Borsboom, Mellenbergh, en Van Heerden, 2003).

Hoewel factoranalyse niet kan aantonen dat een theorie juist is, kan ze een theorie wel verwerpen, namelijk als de theorie geformuleerd is als statistisch model en dit model niet adequaat bij de data past. Uiteindelijk gebeurde dit met de theorie van Spearman. Deze werd verworpen omdat 1-factor modellen over het algemeen te eenvoudig zijn om de statistische relaties in de data goed te verklaren: Na uitpartialisatie van de variabele *g,* blijven bepaalde testscores positief gecorreleerd, wat impliceert dat bepaalde specifieke factoren in Spearman's model niet statistisch onafhankelijk zijn, zoals werd verondersteld. Omdat 1-factor modellen onvoldoende verklaringen bieden, zijn uitgebreidere factormodellen voor intelligentie ontwikkeld. In studieverslagen wordt echter vaak niet duidelijk gemaakt of de factoren in dergelijke modellen dienen te worden beschouwd als reële oorzakelijke bronnen van (co)variantie, bijvoorbeeld als beperkte capaciteiten in psychologische of biologische systemen, of als louter statistische constructen die bepaalde covarianties samenvatten.

Zoals eerder vermeld, om de factoren in een factormodel een zinnige interpretatie als reële bronnen van covariantie te geven is koppeling van het factormodel aan theorie nodig. Een van bekendste intelligentietheorieën is Cattell's (1963, 1987) investeringstheorie aangaande 'vloeibare' en 'gekristalliseerde' intelligentie. Opgemerkt dient te worden dat deze theorie een van de weinige intelligentietheorieën is die rekening houdt met zowel de covariantiestructuur van intelligentie als de ontwikkeling van cognitieve vaardigheden. Onder meer vanwege dit feit is Cattell's investeringstheorie een van de invloedrijkste theorieën in de intelligentieliteratuur.

In investeringstheorie is vloeibare intelligentie een redeneercapaciteit. De ontwikkeling van deze capaciteit gaat samen met de rijping van de hersenen. Individuele verschillen in vloeibare intelligentie zijn uiteindelijk grotendeels toe te schrijven aan individuele verschillen in genetische opmaak. Het verwerven van kennis en vaardigheden, de zogeheten gekristalliseerde vaardigheden, wordt sterk beïnvloed door deze capaciteit. De gekristalliseerde vaardigheden worden niet alleen geacht het resultaat te zijn van 'investering' (gebruik) van vloeibare intelligentie tijdens het aanleren van de vaardigheden, maar ook van motivatie, hoeveelheid genoten onderwijs en cultuur. Verschillen in gekristalliseerde vaardigheden tussen mensen uit verschillende culturen reflecteren in belangrijke mate verschillen in onderwijs en cultuur. Echter, binnen een cultuur, vooral als gecorrigeerd wordt voor verschillen in genoten educatie, zijn individuele verschillen in gekristalliseerde vaardigheden een goede afspiegeling van individuele verschillen in de factor vloeibare intelligentie.

Eén van de gemeenschappelijke factoren in Cattell's bifactormodel voor intelligentie representeert de capaciteit vloeibare intelligentie (afgekort to Gf). De andere factor, te weten gekristalliseerde intelligentie (afgekort door Gc), vat de gemeenschappelijke variantie tussen de gekristalliseerde vaardigheden samen. Omdat gekristalliseerde vaardigheden afhankelijk zijn van vloeibare intelligentie, modelleerde Cattell de factoren Gf en Gc als positief gecorreleerd. Net als Spearman's *g*-model, bleek Cattell's Gf-Gc-model echter te eenvoudig (zie bijvoorbeeld Carroll, 1993).

Tegenwoordig wordt aangenomen dat psychometrische intelligentie ongeveer 70 positief gecorreleerde specifieke cognitieve vaardigheden omvat (McGrew & Flanagan, 1998). Een intelligentiebatterij meet meestal een kleiner aantal van deze vaardigheden, die vertegenwoordigd worden door eerste-orde factoren, maar als men alle cognitieve vaardigheden zou meten en de intercorrelaties tussen alle eerste factoren verder zou factoranalyseren, zou een model met 8 tot 10

positief gecorreleerde tweede orde factoren statistisch gezien een bevredigend resultaat opleveren, zo is de verwachting (Cattell, 1987; Carroll, 1993; McGrew, 2009; Horn & Blankson, 2005). Op dit tweede orde niveau, zijn de twee meest geaccepteerde modellen voor intelligentie, namelijk het uitgebreide Gf-Gc model van Cattell en Horn en het drie-stratum model van Carroll, vrijwel identiek (McGrew, 2009). De interpretatie van de tweede orde variabelen in deze modellen is vaak in termen van eigenschappen van afzonderlijke (neuro-)cognitieve systemen (zie bijvoorbeeld Carroll, 1993). Individuele verschillen in deze eigenschappen worden vervolgens beschouwd als het resultaat van genetische verschillen enerzijds en verschillen in omgevingsinvloeden anderzijds (Jensen, 1998; Plomin et al., 2008).

De tweede-orde variabelen in deze 2 modellen zijn positief gecorreleerd. Het is mogelijk de covariantiestructuur tussen deze tweede-orde factoren verder te factoranalyseren, zoals Carroll (1993) heeft gedaan in honderden datasets. Aan de hand van de resultaten poneerde Carroll het bestaan van een enkele bron van covariantie tussen de tweede orde variabelen, welke in zijn model vertegenwoordigd wordt door een overkoepelende factor op derde-orde niveau. Carroll's model kan worden beschouwd als een model waarin Spearman's oorspronkelijke *g* wordt geherintroduceerd terwijl rekening wordt gehouden met de correlaties tussen bepaalde specifieke factoren (*s*, zie boven) die blijven bestaan na uitpartiallisatie van *g*. In tegenstelling tot Carroll, poneerden Horn en Cattell geen algemene factor. Horn (Horn & Noll, 1997) beschouwde een dergelijke factor als niets meer dan een statistische samenvatting van de covariantiestructuur tussen de tweede orde variabelen; Cattell (1987) was van mening dat de algemene intelligentiefactor (*g*) en de vloeibare intelligentiefactor (Gf) dezelfde variabele voorstelde, en betrekking hadden op hetzelfde cognitief systeem.

In de afgelopen decennia hebben onderzoekers gewerkt aan integratie van het uitgebreide Gf-Gc model en het drie-stratum model (zie McGrew, 2009). Dit geïntegreerde model staat bekend als het Cattell-Horn-Carroll (CHC) model voor intelligentie. De vraag wel of niet een algemene factor te extraheren, blijft hierbij onderwerp van discussie. Belangrijk hier te vermelden is dat deze discussie niet zozeer gebaseerd is op statistische overwegingen maar op theoretische: Een statistische algemene factor kan geëxtraheerd worden, maar of deze factor een reële variabele voorstelt, is de vraag. In het licht van deze discussie zijn alternatieve verklaringen voor de positieve correlaties tussen intelligentietestscores dus belangrijk.

De Inbreng van een Ontwikkelingsperspectief: Mutualisme

Een (met het oog op de ontwikkeling van intelligentie) belangrijke alternatieve verklaring voor de bevinding dat intelligentietestscores positief met elkaar correleren, is dat deze correlaties het resultaat zijn van de dynamische ontwikkeling van cognitie, zoals beschreven in het mutualismemodel voor intelligentie (van der Maas et al., 2006). Deze alternatieve verklaring speelt een voorname rol in de conceptualisering van intelligentie in dit proefschrift (zie Hoofdstuk 1, 6, 7, en Appendix B).

In het mutualismemodel voor intelligentie wordt aangenomen dat de groei in cognitie grotendeels autonoom en zelfregulerend is, maar dat die groei wordt begrensd door genetisch en omgevingsbeïnvloede beperkte capaciteiten. Echter, er bestaan multipele cognitieve processen. De groei van die processen wordt - volgens het oorspronkelijke mutualismemodel - begrensd door processpecifieke (unieke), (statistisch) onafhankelijke capaciteiten. Een andere aanname is dat de cognitieve processen elkaar tijdens hun ontwikkeling wederzijds en overwegend positief beïnvloeden. Dat wil zeggen, de groei van het ene cognitieve proces stimuleert de groei van het andere cognitieve proces en andersom. Als gevolg hiervan heeft de capaciteit die de groei van het ene cognitieve proces begrenst een indirect effect op de groei en het uiteindelijke niveau van functioneren van andere cognitieve processen.

Doelen van het Proefschrift

Zoals boven al kort werd vermeld, is het doel van dit proefschrift uiteindelijke tweeledig geworden. Eén doel was om de ontwikkeling van een adequate theorie van intelligentie nieuwe impuls te geven. Hiertoe werd getracht een model te ontwikkelen dat zowel de cognitieve groei verklaart alsmede het bestaan van genetisch beïnvloedde interindividuele verschillen in intelligentie, en dat resultaten geeft

die in overeenstemming zijn met de belangrijkste empirische resultaten uit intelligentieonderzoek. Een voorbeeld van een dergelijk resultaat is de aanwezigheid van een positieve correlatie tussen intelligentiesubtests' erfelijkheidscoëfficiënten en hun factorladingen op de (statistische) algemene intelligentiefactor. Bij de ontwikkeling van een model dat tot dit resultaat zou moeten leiden, stuitten we op enkele theoretische vraagstukken die niet volledig of niet afdoende worden behandeld in de intelligentieliteratuur. Een tweede doel werd daarom om deze theoretische vraagstukken nader te behandelen. De vraagstukken hebben met name betrekking op de interpretatie van de (hoge) erfelijkheidscoëfficiënten van de variabelen in intelligentiemodellen, hetzij direct, hetzij indirect.

Werkhypothese

Aanvankelijk hebben we als werkhypothese het CHC model aangenomen, omdat dit thans het meest uitgebreide en meest geaccepteerde factormodel voor intelligentie is, en omdat het model van Cattell en Horn en het model van Carroll tot de meest onderzochte modellen behoren. We veronderstelden dat de tweede orde factoren in het CHC model (inclusief Gf en Gc) individuele verschillen in unieke cognitieve systemen representeren, waarvan de groei beperkt wordt door genetische en omgevingsbeïnvloedde capaciteiten. Een reële, onderliggende, algemene factor werd echter niet geponeerd, daar een zo een factor tot dusverre niet is geïdentificeerd en omdat we geloven dat de positieve correlaties tussen het functioneren van de cognitieve systemen toe te schrijven zijn aan wederzijds positieve interacties tussen variabelen binnen die systemen (zie Figuur 1.8). De positieve interacties werden geacht zich voor te doen gedurende de cognitieve ontwikkeling, zoals in het mutualismemodel. We namen verder aan dat individuele verschillen in het functioneren van deze systemen in principe adequaat gemeten kunnen worden door intelligentietests, zonder uit te sluiten dat in werkelijkheid intelligentiesubtestscores het functioneren van meer dan één systeem weerspiegelt (zie Figuur 1.8).

Samenvatting per Hoofdstuk

Hoofdstuk 2 (gepubliceerd als Kan, Ploeger, Raijmakers, Dolan, & van der Maas, 2011) behandelde problemen omtrent de interpretatie van de latente genetische en omgevingsvariabelen in gedragsgenetische modellen in het algemeen. Vanzelfsprekend heeft dit gevolgen voor de interpretatie van de latente genetische en omgevingsvariabelen in factormodellen van intelligentie: De latente variabelen en hun invloeden worden niet gemeten, maar afgeleid, met behulp van genetisch informatieve onderzoeksdesigns (bijvoorbeeld tweeling- en adoptiestudies). Bovendien worden de relaties tussen de variabelen gemodelleerd als lineair. We onderzochten wat de geschatte relatieve bijdragen van genetische en omgevingsvariabelen aan de totale variantie in de geobserveerde variabelen waren indien in werkelijkheid de onderliggende mechanismen niet-lineair waren. Onze hypothese was dat de schattingen dan niet correct zijn, omdat niet-lineariteit een additionele, onafhankelijke bron van variantie vormt (Molenaar, Boomsma, & Dolan, 1993). Om deze hypothese te onderzoeken gebruikten we een dynamisch systeemmodel waarin wederzijdse interacties tussen variabelen plaatsvinden. Het mutualisme model kon niet gebruikt worden, omdat dit model geen niet-lineaire termen bevat. Als alternatief gebruikten daarom het twee-cellen-model van van Oss & van Ooyen (1997) (dit model is beschreven in Appendix A). De twee cellen in het model kunnen worden geïnterpreteerd als subsystemen van werkgeheugen, zodat interindividuele verschillen in de parameters kunnen geïnterpreteerd worden als interindividuele verschillen in werkgeheugencapaciteit. Computersimulaties met het niet-lineaire model leverden resultaten op die overeenkomen met empirische bevindingen in intelligentieonderzoek: Ten eerste, de werkelijke onderliggende bronnen van variantie bleken moeilijk te detecteren. Ten tweede, de geschatte relatieve bijdrage van de latente genetische variabelen aan de totale variantie, d.w.z., de erfelijkheid van de geobserveerde variabelen, nam toe over de tijd (en de geschatte relatieve bijdrage van de latente omgevingsvariabelen nam af). We concludeerden dat voorzichtigheid is geboden bij het interpreteren van hoge erfelijkheidscoëfficiënten als betekenend 'in relatief grote mate genetisch beïnvloed'. De coëfficiënten kunnen bijvoorbeeld overschat zijn als gevolg van de aanwezigheid van niet-lineariteit in ontwikkelingsprocessen.

Meer voorzichtigheid is geboden bij de interpretatie van hoge erfelijkheidscoëfficiënten. Zoals uitgelegd in Hoofdstuk 3, betekent een relatief hoge geschatte erfelijkheid van de ene cognitieve vaardigheid ten opzichte van die van de ander niet dat omgevingsinvloeden en culturele invloeden relatief minder belangrijk zijn. Integendeel, uit de empirie blijkt dat indien we ons concentreren op individuele intelligentiesubtests dat hoe groter de mate is waarin culturele aspecten onderscheid maken tussen mensen, hoe hoger de geschatte erfelijkheid is. Subtests die individuee verschillen in (sterk cultureel afhankelijke) kennis en vaardigheden meten (bijvoorbeeld tests die woordenschat meten), vertonen hogere erfelijkheidscoëfficiënten dan subtests die individuele verschillen in (minder cultureel afhankelijke) cognitieve processen meten. Met het oog op investeringstheorie betekent dit dat gekristalliseerde vaardigheden hoger erfelijk zijn dan indicatoren van vloeibare intelligentie. De oorspronkelijke theorie voorspelde echter het tegenovergestelde: Vloeibare intelligentie is in grotere mate erfelijk dan gekristalliseerd vaardigheden. We stelden voor dat de verklaring ligt in verschillen in gen-omgevings covariantie: Omdat omgeving met name selecteert op gekristalliseerde vaardigheden en niet zozeer op vloeibare intelligentie is (of wordt) gen-omgevingscorrelatie groter voor gekristalliseerde vaardigheden dan voor vloeibare intelligentie. Hierdoor worden in standaard gedragsgenetische modellen de erfelijkheidscoëfficiënten van gekristalliseerde vaardigheden in relatief grotere mate overschat. Uit het onderzoek bleek ook dat subtests die gekristalliseerde vaardigheden meten de hoogste ladingen op de algemene intelligentiefactor laten zien. 'Algemene intelligentie' (als statistisch construct) lijkt dus meer op 'gekristalliseerde intelligentie' dan op 'vloeibare intelligentie'.

In Hoofdstuk 4 werden vloeibare en gekristalliseerde intelligentie in verband gebracht met *groepsverschillen* in IQ scores. We onderzochten Rushton en Jensen (2010)'s redenering dat als groepsverschillen (bijvoorbeeld, raciale of etnische verschillen) het grootst zijn op de subtests die de hoogste erfelijkheidscoëfficiënten en *g*-ladingen vertonen, dat dit impliceert dat de bron van de groepsverschillen genetisch is. Deze redenering is gebaseerd op Jensen (1998)'s (methodologisch zwakke) 'methode van gecorreleerde vectoren'. We concluderen dat zelfs als onderzoekers deze methode als valide methode accepteren, voorgaande conclusie toch invalide is, omdat die conclusie gebaseerd is op drogredenatie, namelijk op *een conclusie trekken uit het gevolg van die conclusie*. We tonen analytisch aan dat groepsverschillen in intelligentie ook het grootst kunnen zijn op de subtests die de hoogste erfelijkheidscoëfficiënten en *g*-ladingen vertonen indien groepsverschillen geheel toe te schrijven is aan verschillen in omgeving. Daarnaast laten we empirisch zien dat raciale (blank-zwart) verschillen het meest uitgesproken zijn op de meest cultureel geladen subtests (gekristalliseerde vaardigheden). Deze resultaten zijn in strijd met eerdere uitspraken van Rushton en Jensen (zie bijvoorbeeld, Rushton & Jensen, 2005). Eerder voorspelden en beweerden zij dat deze verschillen juist het kleinst zijn op de meest cultureel geladen tests. De dominante intelligentietheorieën hebben dus moeite de empirische resultaten te verklaren omdat deze aangeven dat groepsverschillen in *g* (of Gf), het meest uitgesproken zullen zijn op de minst cultureel geladen tests (welke vloeibare intelligentie meten).

Hoofdstuk 5 (waarvan gedeelten zijn gepubliceerd als Kan, Kievit, Dolan, & van der Maas, 2011) betreft de nadere interpretatie van gekristalliseerde intelligentie. Uit een overzicht van de investeringstheorie, blijkt dat gekristalliseerde intelligentie dient worden opgevat als een constructivistische variabele, dat wil zeggen als een statistisch samenvatting (van hoeveelheid kennis). Deze interpretatie wijkt af van de interpretatie van de CHC factoren als representerend individuele verschillen in onderliggende cognitieve capaciteiten. We stelden voor om factor Gc te schrappen uit het CHC-model. We lieten zien dat dit rechtmatig kan worden gedaan aangezien in een heranalyse van een representatieve dataset waarop het CHC-model is gebaseerd, de factor Gc - zoals voorspeld - overbodig was daar deze statistisch equivalent was aan *verbale capaciteit*. In de analyse bleek *g* eveneens overbodig, daar deze variabele statistisch equivalent was aan de variabele vloeibare intelligentie (Gf). We stelden voor om ook *g* als verklarende variabele uit het CHC-model te verwijderen. De grootte van de correlatie tussen Gf en *g* is waarschijnlijk een functie van steekproefheterogeniteit met betrekking tot culturele achtergrond (Kvist & Gustafsson, 2007). In homogene groepen met betrekking tot cultuur, onderwijs en ontwikkeling, zullen de correlaties tussen enerzijds Gc en verbale capaciteit en anderzijds tussen *g* en Gf (redeneervermogen) naderen tot een waarde van +1.

In Hoofdstuk 6 richtten we ons op de vraag of de huidige intelligentietheorieën (*g* theorie, investeringstheorie, 'sampling' theorie, en wederzijdse-interactie-theorieën) in staat zijn de intrigerende bevinding te verklaren dat de meest cultureel afhankelijk intelligentietests de hoogste erfelijkheidscoëfficiënten en de hoogste *g*-ladingen vertonen. Geconcludeerd werd dat ze dat (op zichzelf) niet kunnen. Dit impliceert dat de besproken theorieën niet verklaren hoe raciale groepsverschillen het meest uitgesproken zijn op de meest cultureel afhankelijke, hoogst erfelijke, hoogst *g*eladen vaardigheden.

Het proefschrift eindigde met een algemene discussie (Hoofdstuk 7), waarin we een geïntegreerd model van algemene intelligentie presenteren, gebaseerd op het mutualisme model van der Maas et al (2006). Het model omarmt het voornaamste idee van investeringentheorie, namelijk dat individuele verschillen in cognitieve processen - vloeibare vaardigheden - leiden tot verschillen in kennis en vaardigheden - gekristalliseerde vaardigheden. Dickens & Flynn's (2001, Dickens, 2008) idee van een sociale vermenigvuldiger is eveneens opgenomen. Dit idee houdt in dat in onze cultuur mensen die relatief goed zijn in bepaalde cognitieve vaardigheden grotere kansen krijgen om deze en andere cognitieve vaardigheden verder te ontwikkelen. Verder, conform de mutualismetheorie (van der Maas et al., 2006), veronderstelden we dat cognitieve verwerking profiteert van het verwerven van kennis. In het geïntegreerde model ontbreekt een reële *g*. Genetische correlaties tussen de limiterende capaciteit kunnen aanwezig zijn, maar worden geïnterpreteerd zijnde het resultaat van wat we noemen genetische sampling (Thompson, 1951; Bartholomew et al. 2009; Anderson, 2001; Penke et al., 2007; zie Hoofdstuk 6) en niet als het gevolg van algemene genetische invloeden (Kovas & Plomin, 2006). Simulaties met het model gaven resultaten die in overeenstemming zijn met de bevinding dat individuele verschillen het meest uitgesproken zijn op de meest cultureel afhankelijk subtests, welke de hoogste erfelijkheidscoëfficiënten en hoogste *g*-ladingen vertonen. Het effect werd toegeschreven aan verschillen in gen-omgevings-effecten tijdens de cognitieve ontwikkeling. Meer specifiek, de erfelijkheidscoëfficiënten van gekristalliseerde vaardigheden bevatten in grotere mate gen-omgevingscorrelatie dan de erfelijkheidcoëfficiënten van vloeibare vaardigheden (zoals we in Hoofdstuk 3 hypothetiseerden).

Algemene Conclusie en Discussie

De belangrijkste conclusies van dit proefschrift zijn als volgt. Ten eerste, hoewel het nog steeds niet mogelijk is te bepalen of een realistische, onderliggende *g* aanwezig is of niet, kunnen we opmaken dat de huidige *g* theorieën ontoereikend zijn in het verklaren van belangrijke empirische bevindingen uit gedragsgenetisch intelligentieonderzoek. Naast het individuele-verschillenperspectief dat ze reeds bezitten, behoeven ze een ontwikkelingsperspectief. De rol van de dynamische wisselwerking tussen genetische en omgevingsfactoren die aanwezig is tijdens de cognitieve ontwikkeling dient expliciet te worden gemaakt en ook als zodanig te worden gemodelleerd. Ten tweede, het gebruik van formele modellen is belangrijk in intelligentieonderzoek. Gebruik van zulke modellen kan wetenschappers helpen bij onderzoek naar gecombineerde effecten. Het afsluitende hoofdstuk gaf een voorbeeld van dit soort formeel onderzoek, door specifieke hypotheses in het mutualisme model in te brengen die gezamenlijk de relaties tussen subtests′ *g*-ladingen, culturele ladingen, en erfelijkheidscoëfficiënten zouden moeten verklaren en dat ook deden. Ten derde, het mutualisme model voor intelligentie is uitermate geschikt voor dit soort formeel onderzoek, daar het mutualismemodel het (gebruikelijke) individuele verschillenperspectief combineert met het (veelal ontbrekende) ontwikkelingsperspectief, zoals we in de introductie bepleitten. Op deze manier zullen bevindingen uit de intelligentieliteratuur hopelijk beter worden begrepen. Zoals de geschiedenis van de wetenschap laat zien, leidt beter begrip eindelijk tot betere voorspellingen en tot effectievere interventies. Beter begrip staat gelijk aan betere theorie.

ACKNOWLEDGEMENTS/DANKWOORD

Het is zover: mijn proefschrift is afgerond en dat zonder al te veel vertraging! Mocht u het niet weten, dat laatste is een stuk bijzonderder dan het eerste. Ik ben dan ook menigmaal trotser op het doorstaan van het intensieve promotietraject dan op het uiteindelijke product. Vier jaar lang je gedachten moeten laten gaan over meer dan 100 jaar literatuur is een pittige kluif, zeker als die literatuur uitblinkt in het hanteren van vage begrippen en het gebruik van suboptimale statistische methoden, en waarin bovendien persoonlijke meningen en politieke standpunten het wetenschappelijk onderzoek in grote mate beïnvloeden (lees: in de weg staan). Over kluif gesproken... als je wilt, kun je je volledig stukbijten op het onderwerp intelligentie, weet ik nu. Mijn persoonlijke uitdaging was om wel te bijten, maar toch niet zo hard en lang dat ik geen tijd of ruimte meer had om iets op papier te zetten.

Mijn dank gaat daarom eerst uit naar mijn promotor, professor Han van der Maas. Han, je hebt me eens verteld, volgens mij tijdens mijn sollicitatiegesprek, dat je me als aio aan wilde nemen omdat ik mezelf inderdaad in een probleem vast kan bijten en niet gauw loslaat. Ik nam toen aan dat zulk bijten nodig was voor het aio-project, of omdat het nodig is in intelligentieonderzoek in het algemeen. Later kwam ik erachter dat je waarschijnlijk (nog) een ander doel voor ogen had: Het bewaren van sommige studenten (bijters) voor de wetenschap. Ik heb je leren kennen als iemand die persoonlijke ontwikkeling en wetenschappelijke vorming van je studenten en medewerkers hoog in het vaandel hebt staan en dat sterk stimuleert. Ik vermoed zelfs dat je dat nog belangrijker vindt dan wetenschappelijke output (ik denk dan vooral aan aantal publicaties). De aanwijzingen die ik daarvoor heb bestaan uit het feit dat ik uiteindelijk zeer vrijgelaten werd in mijn onderzoek (of dat zo ervoer); dat ik nooit het gevoel heb gekregen dat ik *per se* jouw ideeën per se op jouw manier op moest schrijven; dat je me toestond liever 1 doorwrocht artikel te schrijven dan - laten we zeggen - 4 oppervlakkige. Daar zou ik (toen) niet gelukkig van worden, me dunkt. Je gaf me daardoor een gevoel van vrijheid, en blijkbaar had ik dat nodig. Hiervoor ben ik je ongelofelijk dankbaar (onderschat het niet!). Wat ik ook waardeer is dat je wél ingreep als dat noodzakelijk was. Je liep dan quasi-nonchalant mijn kamer binnen en vroeg: “Hoe gaat het? Als je vastloopt, omdat er beslissingen genomen moeten worden en je aarzelt, laat me het dan weten. Ik ben namelijk heel goed in beslissingen nemen.” En zo is het! Een beslissing van jouw kant had ik zeker af en toe nodig, dus dank. Naast een goede aanstuurder, ben je verder een creatieve en snelle denker, en bezit je een open geest, met interesse voor alle aspecten van wetenschap. Het was me een genoegen college bij je te volgen (ik denk nu aan het probleem van het touw om de aarde), je te mogen assisteren bij R-werkcolleges (ik denk nu aan het probleem van de lamp in de gevangenis), iets voor je uit te zoeken (ik denk nu aan je opdracht ‘Zoek eens uit waarom crystallized intelligence een factor is. Waarom is dat nou een factor? Een factor... ik snap het niet!’), en om met je van gedachten te wisselen (ik denk nu aan de problemen rondom de nieuwste versie van sampling theorie). Ik hoop ook in de toekomst nog met je samen te mogen werken.

Mijn dank gaat ook uit naar professor (jaja!) Conor Dolan, mijn copromotor en adviseur. Ik dank je niet alleen voor het lenen van je racefiets tijdens mijn vakantie in de Alpen. Ik dank je ook voor al het eindeloos ge-edit van mijn wazige zinnen. Ik zou zelf allang gek geworden zijn (maar misschien was je dat al)! Dank ook voor al je kritische opmerkingen in de kantlijn. Dank voor het beantwoorden van al mijn vragen, vooral als ik dacht ergens een antwoord op te weten, vervolgens dit antwoord bij je probeerde te checken, en terug kwam met de conclusie dat mijn ‘antwoord’ geheel geen antwoord was, maar een product van mijn ‘idiosyncratische’ geest, zoals je die ooit noemde. Niet alleen op de UvA, maar ook op de VU, zal je die geest (dat spook?) weer tegenkomen. Ten slotte, bedankt dat je die idiosyncratische geest toch altijd serieus wilde nemen. Ik hou van leren, en ik leer het meest als blijkt dat ik fout heb gedacht! Jij weet als geen ander fouten in gedachtegangen op te sporen, fijn dus! Eén ding: Ik vind wel dat je wat minder bescheiden mag zijn in je werk. Dat is toch gewoon goed? Lees je artikel ‘Multi-group covariance and mean structure modeling of the relationship between the WAIS-III common factors and sex and educational attainment in Spain’ nog maar eens. Het is een van mijn favoriete artikelen!

Dr. Jelte Wicherts, mijn tweede copromotor, ook jou ben veel dank verschuldigd. Dit blijkt alleen al uit het feit dat ik eerst 1 copromotor zou hebben, en er met 2 geëindigd ben! Ook blijkt je hulp uit een aantal co-auteurschappen van de artikelen die ik ter publicatie ingediend heb. Wat ik vooral van jou geleerd heb, hoewel misschien een beetje laat, is dat je als (goede) wetenschapper niet alleen goed onderzoek moet afleveren, maar dat je het ook zo moet neerschrijven dat het uiteindelijk gepubliceerd wordt. Dit houdt in, denk ik, dat een onderzoeker naast het bezitten van academische vaardigheden ook sluw moet zijn. De onderzoeker moet rekening houden met het volgende: Wetenschappers zijn mensen, terwijl mensen vinden het niet leuk om ongelijk te krijgen. Blijkbaar vinden wetenschappers het niet leuk om wetenschappelijk ongelijk te krijgen, terwijl ongelijk krijgen eigenlijk de essentie is van wetenschappelijke vooruitgang, me dunkt. Op de een of andere manier ben jij in staat ervoor te zorgen dat de onderzoekers die je een wetenschappelijk pak rammel geeft je vriend willen zijn. Briljant! Wat is je geheim? Kan ik een cursus bij je volgen?

Naast een promotor en twee copromotores, bestaat mijn promotiecommissie uit prof. Paul de Boeck, prof. Dorret Boomsma, prof. Henk Kelderman, prof. Maartje Raijmakers en dr. Denny Borsboom. Ook hen dank ik van harte, niet alleen voor het lezen en beoordelen van mijn proefschrift, maar ook voor de interessante en veelal vruchtbare discussies die ik met de meesten heb mogen voeren. Ook dank voor sommige coauteurschappen, en voor mijn huidige baan!

Over co-auteurschappen gesproken, co-auteurs waren - naast Han van der Maas, Conor Dolan, Jelte Wicherts en Maartje Raijmakers - Rogier Kievit, Sanne Haring en Annemie Ploeger. Dank voor al het werk dat jullie verzet hebben. Zonder jullie geen proefschrift!

Verder gaat bijzondere dank uit naar mijn paranimfen, Raoul Grasman en Peter Halpin, bij sommigen (d.w.z., mij) bekend als Rufus en Dufus. Ik dank jullie bij voorbaat voor de assistentie bij de laatste loodjes en het grootse feest dat komen zal (ik ben benieuwd!). Nog meer dank ik jullie voor de trouwe gang naar de kroeg na het werk en de (vaak wetenschapsfilosofisch getinte) gesprekken die we daar voerden. De beste (althans volgens Dufus)! Dank ook voor de goede vriendschap die ontstaan is tijdens de afgelopen (4) jaren! Biertje? Ik zal voor borrelnootjes zorgen, Raoul! And Peter… I wish you would't be allowed to leave the Netherlands! You know why.

Ja, nu wordt het lastig.... Ik dank zovelen... wie als eerste te noemen? Degene die het langst naast me hebben gezeten, mijn kamergenoten! Rogier (alweer), een Mac gebruik ik niet meer, maar ik heb nog wel net zoveel schermpjes openstaan! Zien? Ik dank je voor de vele discussies, hetzij wetenschappelijk, hetzij filosofisch, hetzij anderzijds. Andere kamergenoten van de afgelopen jaren: Marjan, Jij bent vast degene die mijn achterhoofd het best kent! Ik hoop dat je na mijn vertrek nog even hebt kunnen genieten van het (terecht) ingepikte plekkie. Gaat het goed met Karlijn? Sanja, zullen we even bijpraten, morgen? Je bent toch in de buurt.... and … isn't it chocolate time? De rekentuinders, Marthe, Sharon, en 'de Gebroeders'. Lang heb ik niet in de hoekkamer gezeten, maar toch ben ik daar vaker (per ongeluk) opgesloten (omdat ik vaak haast IN mijn computer zit en dus onzichtbaar ben) dan in de kamer die ik erna lang bewoonde. Dank voor het terugkomen. En veel meer, hoor!

Sommigen ben ik dankbaar omdat ik zonder hen geen promovendus *geworden* zou zijn. Ten eerste is dat dr. Hans Phaf, zonder wie ik nooit naar een promotieplaats gesolliciteerd zou hebben. Dank voor het overhalen! Klaas Visser, dank voor je openheid en dat ene laatste zetje. Docenten, met name Peter Molenaar, dank voor de inspiratie. Medewerkers en patiënten/cliënten van het ZMC dank ik voor het meedenken - Lot Holleman in het bijzonder. Medewerkers van Electrisol B.V., vooral Richard, Jan, en Laurens, dank voor de flexibiliteit. In dit rijtje passen ook diverse personen op de discussiefora en sociale netwerken die ik vaak bezocht voor (en ook tijdens) mijn aioschap. Namen zal ik niet noemen, maar voel je vooral aangesproken als je wilt! Auch danke ich gerne Monika. Du verstehst ja alles…. Mijn dank gaat (zelfs) uit naar het Centraal Bureau voor de Statistiek, voor het ontdekken wat ik niet wil!

Sommigen ben ik dankbaar omdat ik zonder hen geen promovendus *gebleven* zou zijn (denk ik): Collega's op de UvA, voor het creëren van een klimaat dat ik wél wil; (destijds) toekomstige collega's van de VU, niet alleen voor het gezelschap tijdens de BGA congressen, maar ook voor het delen van mijn grootste interesse, de gedragsgenetica; studenten, voor nog meer inspiratie! Dylan, bedankt voor het gezamenlijk afleggen van de eerste loodjes (bij Biologie). Mariska (en groep), dank voor de beste cursus ooit! Ik dank ook graag degene met de meeste allround kennis van de afdeling psychologische methodenleer: Ineke! Dank voor alle organisatorische hulp.

I thank professor. W. T. Dickens for discussion and for inspiration.

Naast het aioschap bestaat er zoiets als ontspanning. Zeer belangrijk! Allereerst denk ik dan aan mijn medebandleden, Peer (Poir of Pwaar) en Dimitri (Dimi, de trommelaer). Dank voor alle lawaai, lyrics, lol, en loopjes, maar nog meer voor de trouwe jaren van vriendschap. Dat laatste geldt ook voor mijn (andere) makkers van 'de Bertrand': Marijn (Nijn), Dennis (Dnuz), Tim (Wim of Bim), Ernst (Ernie), en Jasper (pappa Wapsie). Vooral de inmiddels traditionele vakanties naar de Alpen/Dolomieten met jullie - en met Stijn, Rick, Boris, en/of Bas - doen me deugd! Niet alleen vanwege de voortreffelijke hefe Weizen en Schnitzel (Noordkant), en overheerlijke wijn, pizza en pasta (Zuidkant), of vanwege de prachtige uitzichten (beide kanten) die jullie me laten voorschotelen. Zelfs het vervoeren van jullie stinkende mennekes is fijn, zolang dat maar in één van de VTJ-9's gebeurt, he! Doe de groeten aan Gerdien, Ramona, Wendy, Jasmijn, en Johanneke; Dimitri aan Renate. Ook hen dank ik voor de gezelligheid, en gastvrijheid die ze al jaren brengen.

Allan en Pieter. Hier schieten woorden tekort.... Vrienden van het eerste uur. Dank, ook aan jullie families.

Bijzondere dank gaat uit naar de volgende personen, die een grote en/of bijzondere vormende invloed hebben gehad op wie ik uiteindelijk ben geworden. Willeke (ik toast op eeuwige vriendschap!), Pim, Monique, en hun families. Ich danke Ieva.

Mijn familie. Marretje, Bert, Gré, Erwin, Marike, Co, Wil, Robert-Jan, Angelique, John, Thera, Richard, Jeanette, Kees, Fenna, John-Paul, Patricia, Bep, Jan, Irene, Sylvia, Robert, Ron, Jola, Jorrit, Lieke; Lucia ook! Ik hoop dat jullie gezien hebben dat mijn proefschrift is opgedragen aan mijn beide opa's. Ik had hen graag willen vertellen dat ik ze dankbaar ben voor wat ze mij gegeven hebben, bijvoorbeeld (toen ik nog een klein Kees-Jannetje was) in de vorm van zelfgemaakt speelgoed. Daarom heb ik nu eens iets voor *hen* in elkaar geknutseld. Ik mis ze!

Mijn ouderlijk gezin, ten eerste mijn liefste broer en mijn liefste zus: Dank voor het vinden van mijn liefste 'schoonzus' Ingrid en liefste 'zwager' Remco. Zonder jullie zou ik niet mijn liefste nichtjes Robin en Julie en liefste neefjes Jurre en Floris hebben gehad. Carel en Marileen, dank ook dat ik zoveel jaren met jullie dag in dag uit heb doorgebracht, met alles wat daarbij hoort, van vreugde tot verdriet (en natuurlijk weer terug). Dit geldt uiteraard voor mijn ouders, Gerard en Mary. Ook jullie zijn de liefste! Jullie hebben me laten weten trots op me te zijn. Daarvoor dank. En.. vergeet niet: Ik ben minstens zo trots op jullie! Bedankt ook voor de benodigde taaiheid, die ik vast van jullie geërfd heb.

Last but not least. Domna! As we use to say to each other: Me loves u! :-* Thank you for entering my life at the time my focus was primarily on work, so I believe. Thanks to you, my focus shifted towards life (which includes work of course, but also much more) - a life I hope to share with you, my despinis, for much longer! Nuff said… Housy? ;)

Zo.. hehe!