IFAU – INSTITUTE FOR
LABOUR MARKET POLICY
EVALUATION

The relation between fluid intelligence and the general factor as a function of cultural background:

a test of Cattell's investment theory

Ann Valentin Kvist
Jan-Eric Gustafsson

WORKING PAPER 2007:23

The Institute for Labour Market Policy Evaluation (IFAU) is a research institute under the Swedish Ministry of Employment, situated in Uppsala. IFAU's objective is to promote, support and carry out scientific evaluations. The assignment includes: the effects of labour market policies, studies of the functioning of the labour market, the labour market effects of educational policies and the labour market effects of social insurance policies. IFAU shall also disseminate its results so that they become accessible to different interested parties in Sweden and abroad.

IFAU also provides funding for research projects within its areas of interest. The deadline for applications is October 1 each year. Since the researchers at IFAU are mainly economists, researchers from other disciplines are encouraged to apply for funding.

IFAU is run by a Director-General. The institute has a scientific council, consisting of a chairman, the Director-General and five other members. Among other things, the scientific council proposes a decision for the allocation of research grants. A reference group including representatives for employer organizations and trade unions, as well as the ministries and authorities concerned is also connected to the institute.

Postal address: P O Box 513, 751 20 Uppsala
Visiting address: Kyrkogårdsgatan 6, Uppsala
Phone: +46 18 471 70 70
Fax: +46 18 471 70 71
ifau@ifau.uu.se
www.ifau.se

ISSN 1651-1166

The relation between fluid intelligence and the general factor as a function of cultural background: a test of Cattell's investment theory[*]

by

Ann Valentin Kvist[♣] and Jan-Eric Gustafsson[♦]

June 27, 2007

Abstract

According to Cattell's (1987) Investment theory individual differences in acquisition of knowledge and skills are partly the result of investment of Fluid Intelligence (*Gf)* in learning situations demanding insights in complex relations. If this theory holds true *Gf* will be a factor of General Intelligence (*g*) because it is involved in all domains of learning. The purpose of the current study was to test the Investment theory, through investigating effects on the relation between *Gf* and *g* of differential learning opportunities for different subsets of a population. A second-order model was fitted with confirmatory factor analysis to a battery of 17 tests hypothesized to measure four broad cognitive abilities The model was estimated for three groups with different learning opportunities (N = 2358 Swedes, N = 620 European immigrants, N = 591 non-European immigrants), as well as for the total group. For this group the *g*-*Gf* relationship was 0.83, while it was close to unity within each of the three subgroups. These results support the Investment theory.

Keywords: Structure of intelligence, Cattell's Investment theory, Fluid Intelligence, General Intelligence
JEL-codes: J24, J61, C3

---

[*] The authors would like to thank Erik Mellander, Berit Carlsted, Jan te Nijenhuis and an anonymous referee for valuable comments

[♣] Göteborg University, email: ann.valentin.kvist@ped.gu.se
[♦] Göteborg University, email: jan-eric.gustafsson@ped.gu.se

Table of contents

1 Introduction

Ever since Spearman (1904, 1927) introduced his “Theory of Two Factors”, issues concerning the structure of human intelligence have been the focus of attention of much research. While there certainly are differences in opinion regarding a wide range of issues, consensus has been achieved that a hierarchical representation of the structure of cognitive abilities is required to capture the complexities of the phenomenon (e. g., Carroll, 1993; Gustafsson, 1988; Jensen, 1998; Johnson & Bouchard, 2005a, b; Messick, 1992). The currently most widely accepted hierarchical model is the Carroll (1993) “Three-Stratum Model”. Since this model may be regarded as an extension of the Cattell and Horn “Gf-Gc” model (see e. g., Horn & Cattell, 1966) it is also referred to as the Carroll-Horn-Cattell (CHC) model (McGrew, 2005).

The CHC model includes factors of three degrees of generality (Carroll, 1993; McGrew, 2005). At the lowest level (stratum I) there are at least some 60 narrow factors, many of which correspond to factors previously identified by Thurstone (1938), Guilford (1967) and other researchers working in the tradition of multiple factor analysis. At stratum II some ten broad factors are identified, and among these, a few are seen as especially prominent, primarily because of the attention they have been given in the research conducted by Cattell and Horn (see, e. g., Cattell, 1963, 1971, 1987; Horn, 1968; Horn & Cattell, 1966). One is *Fluid Intelligence* (*Gf*), which is interpreted as the capacity to solve novel, complex problems, using operations such as inductive and deductive reasoning, concept formation, and classification. Another factor is *Crystallized Intelligence* (*Gc*), which represents individual differences in breath and depth of knowledge of the language, information and concepts of a culture. *Gc* is acquired through education and experience and it primarily reflects verbal knowledge and skills, as well as declarative knowledge in wide areas. Another important factor is *Broad Visual Perception* (*Gv*), which is an ability to generate, retain, retrieve and transform visual images. *Cognitive Processing Speed* (*Gs*) is a broad ability to fluently perform relatively easy or overlearned tasks, particularly when attention and focused concentration is required.

At the third stratum the CHC model includes a factor of General Intelligence (g). This factor relates most highly to complex reasoning tasks while it has lower relations to the stratum II factors involving simpler speeded tasks.

According to Carroll (1993) this suggests that the g-factor involves complex higher-order cognitive processes.

However, even though there is consensus at a general level that such a hierarchical arrangement of factors represents a useful taxonomy of human cognitive abilities, there are substantial differences in opinion concerning fundamental theoretical issues both between the three researchers after which the CHC model has its name, and other researchers (e. g., Johnson & Bouchard, 2005a, b). The most striking locus of differences concerns the need for a *g*-factor at the apex of the hierarchy, and, if such a factor is accepted, the nature of this factor (cf. McGrew, 2005).

1.1 The general factor in hierarchical models of intelligence

According to Carroll (1993) the empirical evidence strongly supports the existence of a *g*-factor, and Jensen (1998), along with many others (e. g., Gustafsson, 1988), has arrived at the same conclusion. However, Horn (see, e. g., Horn & Blankson, 2005; Horn & Noll, 1997) has strongly objected to the idea of a general factor, favouring instead a hierarchical model with broad correlated factors at stratum II. According to Horn & Blankson (2005) the evidence from research on the structure of intelligence is incompatible with the notion that there is a single *g*-factor. They argued, furthermore, that construct validation evidence from several areas of research, such as life-span developmental research "… is counter to a theory that human intelligence is organized in accordance with one common principle or influence." (Horn & Blankson, 2005, p. 53). The main reason for Horn's resistance against a stratum III *g*-factor is that he regards such a factor as a hybrid factor, which is a composite of different stratum II factor. Since the nature of the *g*-factor is determined by the composition of the test battery, it lacks factorial invariance, and *g* is therefore not a meaningful scientific concept (Horn & Noll, 1997).

The problem of the potential non-invariance of the *g*-factor identified by Horn is a serious one indeed, and this problem has been given much attention in the history of research on intelligence. Two main approaches to solving the problem may be identified, which both go back to Spearman (1927). The first approach is to try to identify the essential characteristics of *g* theoretically, and to devise measures which capture these. Spearman (1923, 1927) formulated the "noegenetic laws" which specify that tasks that evoke *g* involve three essential cognitive operations: apprehension of experience, eduction of relations, and eduction of correlates. This provided the theoretical basis for development of the Raven Progressive Matrices Test, which is regarded as the prototypical *g*-test (Jensen, 1998).

The other approach is based on Spearman's (1904, 1927) principle of the "indifference of the indicator" which basically says that one and the same *g*-factor may be estimated from the sum of scores on any varied set of intellectual tasks. This principle follows from Spearman's (1904, 1927) Two-Factor Theory, which states that performance on any single task is influenced by a *g*-factor common to all tasks, and an *s*-factor which is unique to each task and uncorrelated with all other *s*-factors and with *g*. All tests of general intelligence which produce a global, composite, IQ score rely on the principle of the indifference of the indicator. However, while the principle of the indifference of the indicator is guaranteed to be applicable to data which fit a one-factor model, there is no mathematical proof that this principle produces invariant estimates of *g* when applied to data which fit a hierarchical, multidimensional, model.

The label '*g*' is obviously used in quite different meanings. It often refers to a global composite score from heterogeneous tests which is estimated, for example, as a sum of subtest scores (Buckhalt, 2002), as the first principal factor (Jensen, 1998), or as the single factor of a one-factor model (Gignac, 2006). This approach makes the assumption of unidimensionality, but the aggregation of scores makes it reasonable to expect that a factor common to all items of a sufficiently broad and heterogeneous test will dominate the composite score even when the unidimensionality assumption is violated (Gustafsson, 2002). When the *g*-factor is estimated as the apex factor in the CHC model no assumption of unidimensionality is made, but the model is constructed to reflect the actual number of dimensions present in the data.

The Horn criticism thus may be reformulated as a statement that *g* -factors derived from different test batteries are not one and the same. The issue of the non-invariance of *g* has been investigated in a large number of studies, which have been reviewed by Jensen (1998, Ch. 4). He arrived at the conclusion that the *g*-factor is invariant across different methods of factor analysis, and that there is a large amount of consistency of estimates of *g* scores for the same person when the estimates are derived from different test batteries. Johnson *et al.* (2004) even demonstrated estimates of *g* scores from three different batteries to be perfectly correlated.

Carroll (1993, pp. 591–597) examined the degree of consistency in the pattern of *g*-factor loadings obtained in a large number of studies and observed that estimates of factor scores obtained from these studies could be expected to be highly correlated. But he also made the observation that " … the g factor for a given data set is dependent on what lower-order factors or variables are loaded on it" (p. 596). Ashton & Lee (2005) showed that the size of loadings on a first principal factor of a battery of tests was to a large degree influenced by what other tests were included in the battery. It was, in particular, observed that when several *Gc*-tests were included in the analysis, these tests obtained the highest loadings on the *g*-factor, but this was not the case when only a single *Gc*-test was included. This dependence of the nature of the *g*-factor on the composition of the test battery supports Horn's criticism of the non-invariance of *g*.

Jensen (1998, pp. 90–91) estimated the *g*-loading of IQ-tests from the average intercorrelation between different tests, and concluded that the average *g* loading is 0.83–0.88. While these estimates are high, they show that there also are other sources of variance in the scores obtained from IQ tests.

It may, thus, be concluded that even though there is considerable empirical support for the possibility to estimate an invariant *g*-factor, there also are indications that this generalization may not be without its limitations.

The question about non-invariance of the *g*-factor has been approached in yet another way, which may be regarded as an extension of Spearman's theoretically based approach to the identification of *g*. Undheim (1981) and Gustafsson (1984) argued that the characteristics of the *g*-factor as described by Spearman (1904, 1927) agree so well with the characteristics of the *Gf*-factor as described by Horn & Cattell (1966), that *g* and *Gf* should be considered to be one and the same factor. The equality of *g* and *Gf* also has been demonstrated empirically in a series of studies in which a higher-order *g*-factor has been shown to have a perfect relationship with the *Gf*-factor (e. g., Gustafsson, 1984, 1988, 1994, 2002; Undheim, 1981; Undheim & Gustafsson, 1987). Since *Gf* is identified in an invariant manner, it follows that *g* too is invariantly defined as an apex factor in the CHC model.

Horn & Blankson (2005, p. 53) rejected this line of reasoning, arguing that *Gf* does not account for the interrelationships among other variables indicative of intelligence. However, if *Gf* is equivalent to a stratum III *g*-factor in the CHC model which accounts for the intercorrelations among the stratum II factors this statement is incorrect. This issue thus could and should be determined on the basis of empirical research.

While the *g* = *Gf* relationship has been observed in many other studies as well (e. g., Keith, 2005; Reynolds & Keith, 2007), all attempts at replication have not been successful. Carroll (1993) reanalyzed the Gustafsson (1984) data, and failed to find the perfect relationship between *g* and *Gf*. One reason for this may be that Carroll (1993) relied on exploratory factor analysis, and with this technique he failed to identify the Inductive factor, which in turn caused him difficulties separating *Gf* and *Gv*. However, in another study of the relationship between *Gf* and *g*, Carroll (2003) used confirmatory factor analysis (CFA), without quite being able to show the identity. It thus must be concluded that the empirical support for the equivalence between *Gf* and *g* is strong, but not unanimous. The results presented by Carroll (1993) show, however, that *Gf* is the stratum II factor which has the highest loading on the stratum III *g*-factor.

Johnson & Bouchard (2005a, 2005b) challenged the entire CHC framework, through taking their starting point in Vernon's (1961) hierarchical model. The Vernon model is an extension of the Spearman (1904, 1927) Two-Factor Model in such a way that in between the *g*-factor and the specific factors Vernon introduced two major group factors (*v:ed* and *k:m*), and below these a large set of minor group factors. Johnson & Bouchard (2005a) used CFA to fit three alternative hierarchical models to a very large test battery. One of the models was derived from the Cattell-Horn *Gf*-*Gc* model, another from the Carroll model, and the third from Vernon's model. The results showed that the Vernon model fitted better than the other two models. Through adding memory and image rotation factors Johnson & Bouchard (2005a) showed that fit could be improved further. Given that this model is a descendent of Spearman's model it may be expected that in this model too the *g*-factor would have its strongest relation to factors representing reasoning. A very strong fourth-order *g*-factor was indeed found, which had a correlation of 0.99 with a third-order perceptual factor which primarily had relations to complex spatial and reasoning tests. Parenthetically it may also be noted that in the model derived from Carroll (1993) there was a loading of 1.00 of the second-order *Gf*-factor on *g*, while for all the other factors the relation was lower.

It would carry too far to discuss here the details of how the Johnson & Bouchard (2005a) model compares with the other hierarchical models. Suffice it to conclude that this model too supports the interpretation of the *g*-factor as a non-verbal reasoning factor.

However, strong opposition also has been voiced against the idea that the *g*-factor is a non-verbal reasoning factor, and instead it has been argued that measures of verbal abilities are better indicators of *g* (e. g., Gignac, 2006; Robinson, 1999). One of the bases for this argument is the observation that the verbal subtests (e. g., Vocabulary and Information) in the Wechsler batteries have the highest loadings on the first principal factor (Gignac, 2006). However, this need not necessarily be because these tests have higher *g*-loadings. If verbal tests are over-represented in the test battery this may cause the *g*-factor to be verbally biased, because the *g*-factor is confounded with a verbal group factor.

Gignac (2006) devised a procedure which aims to control for such bias. In the so called single-trait correlated uniqueness (STCU) CFA procedure a one-factor model is fitted first, and any further common variance is accounted for through allowing pair-wise covariances of the residuals of groups of tests. Gignac (2006) applied this procedure to five datasets, in most of which different Wechsler test batteries had been used. The residuals showed that the STCU CFA procedure reduced the difference between factor loadings for verbal and non-verbal groups of tests. Correcting factor loadings for unreliability of the tests tended in some cases to further reduce the difference. However, even so there still remained a difference in favour of the verbal tests. On the basis of these results Gignac (2006) concluded that the crystallized subtests are the best indicators of *g*, and he rejected the Gustafsson (1984) conclusion that *Gf* equals *g*, arguing that this result was caused by a methodologically flawed selection of tests.

There are, however, several problems with Gignac's (2006) analyses and conclusions. As was acknowledged by Gignac (2006, p. 43) the lack of fluid intelligence subtests within the batteries analyzed made this study less than ideally suited to investigate the differences in *g*-loadings for crystallized and fluid subtests. Ashton & Lee (2006) also showed that the STCU CFA procedure cannot be relied upon to produce unbiased estimates of *g*-loadings, because alternative models with equally good fit may be fitted to the same data. They furthermore demonstrated that different results were obtained with large test batteries than with small batteries. Applying the STCU CFA procedure to two large batteries they concluded that there was no difference in the size of the *g*-loadings for verbal and reasoning tasks.

The results reported by Ashton & Lee (2006) could be interpreted as showing that the hypothesis that verbal subtests are the best measures of *g* is incorrect, and also that the hypothesis that *g* equals *Gf* is incorrect. This conclusion may be premature, however, because there is yet another problem with the STCU modelling procedure that may cause it to produce biased results. The problem is that this procedure does not take into account the fact that the amount of truly unique variance may be different for different tests, and for different categories of tests. The STCU CFA procedure captures any covariance between the test residuals, but in addition there may be systematic test-specific variance which adds to the reliability of the test. In CFA models this variance is typically confounded with the random error variance, but in certain situations these sources of variance may be separated. One way to do that is to split each test into half-tests, which are both entered into the CFA model. Mårdberg & Carlstedt (1998) fitted such a model to a battery of tests measuring verbal, spatial and reasoning abilities. The results showed the unique components of variance to be substantially larger for the spatial and reasoning tests than for the verbal tests.

The test uniqueness attenuates the estimated loading of the test on *g*, but it does not influence the estimated loading of a lower-order factor on a higher-order factor. This circumstance may explain the seeming contradiction that there is a perfect relation between *g* and *Gf*, while at the same time the tests measuring *Gf* do not load more highly on *g*, or even lower, than do tests measuring crystallized abilities.

To summarize the discussion so far it may be concluded that there is considerable evidence in favour of a *g*-factor as an apex factor of the hierarchical model of the structure of intelligence. The available evidence also provides some, but far from unanimous, support for the idea that the *g*-factor is a non-verbal reasoning factor. Thus, more empirical evidence is needed to settle the issue about the nature of the *g*-factor. However, in order to arrive at a deeper understanding we also need stronger theory. There is reason, therefore, to go somewhat more deeply into a theoretical model that aims to explain the nature of the *g*-factor.

1.2 The Investment theory

Cattell is often ascribed the same negative position with regard to the *g*-factor as is taken by Horn. However, this is incorrect because there is a profound difference between the positions of these two researchers with regard to the meaningfulness of introducing a higher-stratum *g*-factor. While Horn rejected such a factor as meaningless because it is non-invariant, Cattell (1971, 1987) was not hostile to the idea that there is a *g*-factor. Cattell (1987) argued that according to conventional wisdom it would be expected that the stratum II factor *Gc* should have a loading of unity on the stratum III *g* factor. However, four early third-order factor analytic studies involving *Gf* and *Gc* along with personality variables demonstrated that *Gf* had higher loadings on *g* than had *Gc*. Cattell interpreted the third-stratum general factor to be the "historical" g_f factor, and he asked "How is one to explain this tendency of the historical g_f (i. e., $g_{f\ (h)}$) to load g_f more than it does g_c?" (p. 138). The answer proposed by Cattell was the Investment theory.

The Investment theory postulates that in the development of the individual there is initially a single, general, relation-perceiving ability which is connected with the maturation of the brain. This ability, which was labelled *Gf* by Cattell, is thus primarily associated with genetic factors and neurological functioning. It can be applied to any sensory, motor or memory area, and Cattell argued that a child's rate of learning of different tasks (e. g., spatial, numerical, conceptual) depends on this general ability. In particular the child's:

> … rate of learning in fields demanding insights into complex relations – and these fields include especially the problems of reading, arithmetic, and abstract reasoning with which he struggles at school – will depend appreciably on his level of fluid intelligence (though motivation, goodness of teaching, etc., will still play their part, as with the acquisitions of low relational complexity). (Cattell, 1987, p. 139).

Thus, through practice and experience children develop knowledge and skills and according to the Investment theory these developed abilities (i. e., *Gc*) are influenced by *Gf* and by effort, motivation and interest, and also by previous levels of *Gc*. The reason why *Gf* is a general ability is that:

> … in all kinds of relation-eduction in new material requiring fluid ability, the child high in one manifestation will be high in another, and from correlations rooted in such observations eventually we obtain the fluid ability factor. But as a result of the fluid ability being *invested* in all kinds of complex learning situations, correlations among these acquired, crystallized abilities will also be large and positive, and tend to yield a general factor. (Cattell, 1987, p. 139).

The Investment theory thus can provide an explanation for the observation that *Gf* tends to have a perfect relationship with *g*. It is worth noting that even though the Investment theory has been much discussed and investigated, this particular aspect of the theory has not been a focus of attention. One reason for this may be that the link between the effects of *Gf* on individual differences in acquisition of knowledge and skill in different areas and *Gf* as a general, higher-order, factor is not immediately apparent. However, this link is easy to understand once it is realized that a higher-order factor exerts an influence on a greater number of manifest variables than does a factor below it. Thus, the *g*-factor has a wider breadth of influence than other factors of intelligence, but it does not necessarily exert a particularly strong influence on performance on any single task (see Coan, 1964; Gustafsson, 2002; Humphreys, 1962).

Thus, according to Cattell's line of reasoning, the *Gf* factor develops into a general factor because it influences acquisition of knowledge and skills in different domains. For example, most new words are learned by inferring their meanings from the contexts in which the words are embedded (Lohman, 2004). Similarly, Landauer & Dumais (1997) argued that most knowledge development occurs through inductive inference of partial information encountered in different contexts. In support of Cattell's reasoning, this makes vocabulary tests and other tests of knowledge reflect the efficiency of past reasoning processes. But if different subgroups within a population have had different opportunities to acquire the knowledge tested, for example because the language of the test is the mother tongue for some, and the second or third language for others, the simple relationship between *Gf* and amount of knowledge acquired will break down. This should apply not only to acquisition of verbal knowledge and skills, but also to development of abilities in other domains, such as the spatial one.

This suggests a way to test both the Investment theory and the hypothesis that *g* equals *Gf*, namely through investigating the effect on the relation between *Gf* and *g* of differential learning opportunities for different subsets of a population. From the Investment theory follows the prediction that within populations which are homogeneous with respect to learning opportunities there should be a perfect relationship between *Gf* and *g*, while for populations which are composed of subgroups who have had different learning opportunities, the relation between *Gf* and *g* should be lower. This implies that the validity of the Investment theory may be tested through investigating the strength of the relationship between *Gf* and *g* in homogeneous and heterogeneous populations. A similar suggestion was made by Carroll (1996, p. 16).

2 Method

For the purpose of our study we need a fairly large number of subjects with different cultural backgrounds, who have been subjected to the same testing procedure. Such a group was made available to us through the Swedish National Labour Market Board. The agency offers vocational training to persons who are unemployed or at risk of becoming unemployed. In selecting candidates a procedure called “RA”, which is a Swedish acronym for “Directed Aptitude Testing”, is used. The procedure includes the use of traditional psychometric tests, and it has been administrated to a large number of job applicants over the years 1993–2003. Over time and over applicants to different training courses the tests have differed somewhat, but a core of 15-20 tests have been used with great frequency. These form the basis of the analyses in this study.

2.1 The test battery

Most of the tests used in the RA procedure are based on tests designed to measure the seven Primary Mental Abilities of Thurstone (1938). Tests were imported from the United States and adapted for use in Sweden, or they were developed in Sweden on the same schema. The tests will be interpreted within the framework of the CHC model, which entails attributing the Thurstone abilities to the stratum II abilities using the Carroll (1993) and McGrew (2005) findings.

The following tests were included:

(i) *Raven’s Standard Progressive Matrices.* This test was developed to measure the eductive component of g as defined by Spearman (1927). The test requires the completion of a matrix pattern, and is non-verbal (Raven, Court & Raven, 1998). The test consists of 60 items. It is expected to be influenced primarily by *Gf*, even though a small relationship with *Gv* (Gustafsson, 1984; Lynn, Allik & Irwing, 2004) is also expected.

(ii) *Aros Number Series.* Originally constructed by Thurstone the object of the test is to measure mathematical-inductive ability. The test consists of number series, and the object is to identify the mathematical basis of the series and then add the next two numbers in the series. There are 20 such items. The test is expected to be influenced primarily by *Gf*.

(iii) *USTM Number Series.* The principle is the same as for “Aros Number Series”, but the items are less complicated and the subjects are asked to add a single number only. The test has 38 items, and is expected to be influenced primarily by *Gf*.

(iv) *WIT Numbers.* The WIT tests were developed in Sweden on the basis of the Thurstone model. In WIT Numbers the task is to create a mathematical statement from given numbers, using simple arithmetic principles. An example is “2; 2; 4”. Here the correct answers are anyone of: “2+2=4”; “4-2=2”; “2x2=4”; or 4/2=2”. The test has 20 items, and is expected to be influenced primarily by *Gf*.

(v) *R16A.* This test consists of mathematical tasks, where the problems are presented in written form. An example is: “Per had 3 apples and Anders had 7 apples. How many more apples did Per have, as compared to Anders?” The test has 28 items. It is primarily expected to be influenced by *Gf*, but also by *Gc*, because of the verbal instructions and since it presupposes some mathematical knowledge.

(vi) *Instructions II.* A number of instructions with verbal, numerical and spatial content demanding working memory capacity are to be carried out. The output is a written statement, a number, a drawing, or a combination of these. An example is: “If there are more than 50 centimetres to a meter, then underline “No”. If this is not the case, then circle “100”. ” The test has 39 items and is expected to be influenced primarily by *Gf*, and to some extent by *Gc*.

(vii) *SP2A.* Simple drawings illustrate different technical situations, such as heating systems, vehicles in motion, or electric circuits. A written statement poses three alternative outcomes, and the correct one should be indicated. An example is: ”Which pair of scissors would you use to cut wire?” The illustration shows scissors with different proportions. The test has 45 items and is expected to be influenced by *Gc* and *Gv*.

(viii) *DLS Reading*. This is a test of reading speed and reading comprehension in Swedish intended for grades 7–9. In the text parentheses are inserted, each containing three words or expressions. Only one fits the content of the story, and this should be underlined. An example is: "The largest herd living animal of the tundra is the musk ox. …The musk ox is well equipped for life (in the desert *on the tundra* in the forest), and is often forced to dig up its feed from under the snow." The test has 34 items, and is expected to be influenced primarily by *Gc*.

(ix) *WIT Antonyms*. Part of the WIT battery, WIT Antonyms is a vocabulary test. On each line five words are presented. The subject should find the two that are antonyms. An example is "Beautiful *Old* Sad Fast *Young*". The tests has 29 items, and is expected to be influenced primarily by *Gc*.

(x) *WIT Puzzle*. Part of the WIT battery, the WIT Puzzle is a test of two-dimensional spatial ability. The subject should indicate the parts that together with a given figure form a square. The test has 20 items and is expected to be influenced primarily by *Gv*.

(xi) *Aros Metal Folding*. The drawing of a sheet of metal is indicated with solid and dotted lines. The solid lines should be imagined cut, and the dotted folded into a sharp crease. This creates a figure that should be indicated among four choices. The test thus requires the mental transformation of two-dimensional figural representations into three-dimensional ones. The test has 40 items, and is expected to be influenced primarily by *Gv*.

(xii) *Wire*. This is a manual test of spatial ability. The subject is presented a large two-dimensional figure made of coarse wire, and should reproduce this, but on a smaller scale, using a straight piece of wire 75 centimetres long, and with a diameter of 1 millimetre. The result is graded directly on a stanine scale. It is expected to be influenced primarily by *Gv*.

(xiii) *Stockholm box*. A manual test of spatial ability. The subject is presented with mechanical models, and a box with mechanic parts, that should be assembled to copy the models. A point is given for each correct part and the maximum score in the rescaled version used here is 21.3. This test is expected to be influenced primarily by *Gv*.

(xiv) *Crawford Pins.* The task is to enter thin metal pins into small holes with the use of tweezers. A maximum of 36 can be obtained. The test is expected to be influenced primarily by *Gv*.

(xv) *P-numbers, P-words, P-figures.* These are tests of perceptual speed, with numerical, verbal and figural content, respectively. The subject is presented with two columns, each with 5 groups of numbers, letters or figures. The numbers consist of four digits, e.g. “2212”. The letter combinations are made up of three letters, e.g. “Hhp”. In the left hand column one, two or three items are indicated by being crossed out. The task is to cross out the matching items in the right hand column. The maximum score is 150 on each. The tests are expected to be influenced primarily by the general speed factor, Gs.

2.2 Subjects

The subjects in this study were all registered at the Employment office and participated in the RA procedure in connection with their employment officer suggesting vocational training. The group consists of N = 3570 subjects, of whom 86.1 % were men and 13.9 % were women. The predominance of men could be caused by selection of the employment officers, but most likely it is the result of self selection, based on the type of training the courses offered. Many of these are oriented towards traditionally male areas of work.

The ages of the subjects range from 18 to 60, with a mean of 33.6 years and a standard deviation of 8.8. A little more than 10 % of the subjects have a coded disability, typically involving aspects of mobility. The disability is often the reason for the need to change area of work.

In connection with the RA procedure all subjects were interviewed about their school backgrounds. The country where the applicant had received the basic schooling (approximately the first nine years) was registered. This measure was chosen as an alternative to citizenship or country of birth, since primary schooling was considered to have a more direct influence on learning opportunities, even though it is realized that this too is a crude measure. For the purpose of our study countries were grouped into larger entities.

Those who received their primary schooling in another country than Sweden were regarded as immigrants. This group consisted of persons with a multitude of reasons for migrating to Sweden. The immigrants (N = 1211) had spent a mean of 8.2 years in Sweden, with a range from 1 to 31 years (sd = 5.5).

Language background is likely to affect performance on the test battery, because this has affected the opportunity to acquire knowledge and skills focused upon in the tests. Furthermore, the instructions in the RA procedure are given in Swedish. Being a native Swedish speaker with a Swedish school background is thus an advantage in the test situation. However, cultural background also influences the development of other than verbal aspects of intelligence (Sternberg & Kaufman, 1998), and it affects the experiences of and attitudes towards psychometric testing. Thus, in Western societies tests of different kinds are abundantly used, and being subjected to tests is a fairly common, if not an altogether relaxed, experience. In other cultures the concept of psychometric testing may be more or less unknown, which could make the testing situation more obscure. Familiarity with this kind of procedure differs with cultural and educational background, as does the experience of its validity. Cultural concepts and emotional values attached to working fast versus working with accuracy may differ. Such cultural factors are likely to affect test performance over and above the influence of language.

In a broad classification it can be expected that immigrants from Western countries are more adequately prepared and have a more familiar relation to the tasks involved in the test battery than immigrants from non-Western countries. In this material immigrants from the European group of countries, with the addition of USA, Australia and New Zeeland, were therefore brought together in one category, The remaining non-European group consisted largely of immigrants from the Middle East and northern Africa. The analysis thus focuses on three groups of subjects: Swedish non-immigrants (N = 2358), European immigrants (N = 620) and Non-European immigrants (N = 591).

The applicants were asked about their educational backgrounds, and even though this information may not be perfectly reliable, it should be useful for judging the comparability of the groups with respect to level of education. Among the Swedish non-immigrants 43 % had 12 or more years of theoretical education, while for European immigrants and Non-European immigrants the corresponding figures were 36 % and 43 %, respectively. The three groups thus were quite similar with respect to level of education.

2.3 Analytical procedures

The data collected from the RA procedure were available in a data base. No single participant took every test in the test battery. Instead different subgroups of participants were administered different subsets of tests. The test battery used in each case was put together by the psychologist responsible for the RA procedure, primarily on the basis of what training program was applied for. This procedure resulted in incomplete data, which, however, can be analyzed with the missing data modelling procedures developed by Muthén, Kaplan & Hollis (1997) (see Roberts *et al.*, 2000, for an interesting example).

The missing data procedure applied makes the assumption that the data is 'missing at random' (MAR), which implies that the procedure yields unbiased estimates when the missingness is random given the information in the data. This is a much less restrictive assumption than the assumption that the data is 'missing completely at random'. Even though we cannot guarantee that the assignment of tests to applicants yields data that completely satisfy the MAR assumption, the fact that there are high interrelations among observed variables which are exchangeable indicators of a limited set of latent variables implies that there is much information in the data, which should allow for good possibilities to satisfy the MAR assumption.

The tests listed above were used as manifest variables in a confirmatory factor analysis model. A higher-order model was fitted to the data for the whole group, and this model was then tested on the three subgroups of subjects. The modelling was done with the Mplus Version 3 program (Muthén & Muthén, 2004), under the STREAMS 3.0 modelling environment (Gustafsson & Stahl, 2005).

Even though the three-level CHC-model served as the conceptual framework, the model was set up as a higher-order model with factors at two levels, the stratum II factors being identified as first-order factors. The reason for this was that the test-battery included too few tests to allow identification of the stratum I factors. At the level of first-order factors *Gf* thus was set to relate to tests requiring general novel problem solving capacity; *Gc* was set to relate to tests measuring verbal knowledge and skills such as reading speed and vocabulary, and also to tests measuring numerical skills; *Gv* was hypothesized to relate to tests involving two-and three-dimensional tasks, as well as some manual tests of spatial skills; and *Gs* was hypothesized to relate to tests of perceptual speed and accuracy. On the second-order level the stratum III general factor, *g*, was hypothesized to relate to all the stratum II latent variables. The hypothesized model is presented in *Figure 1*.

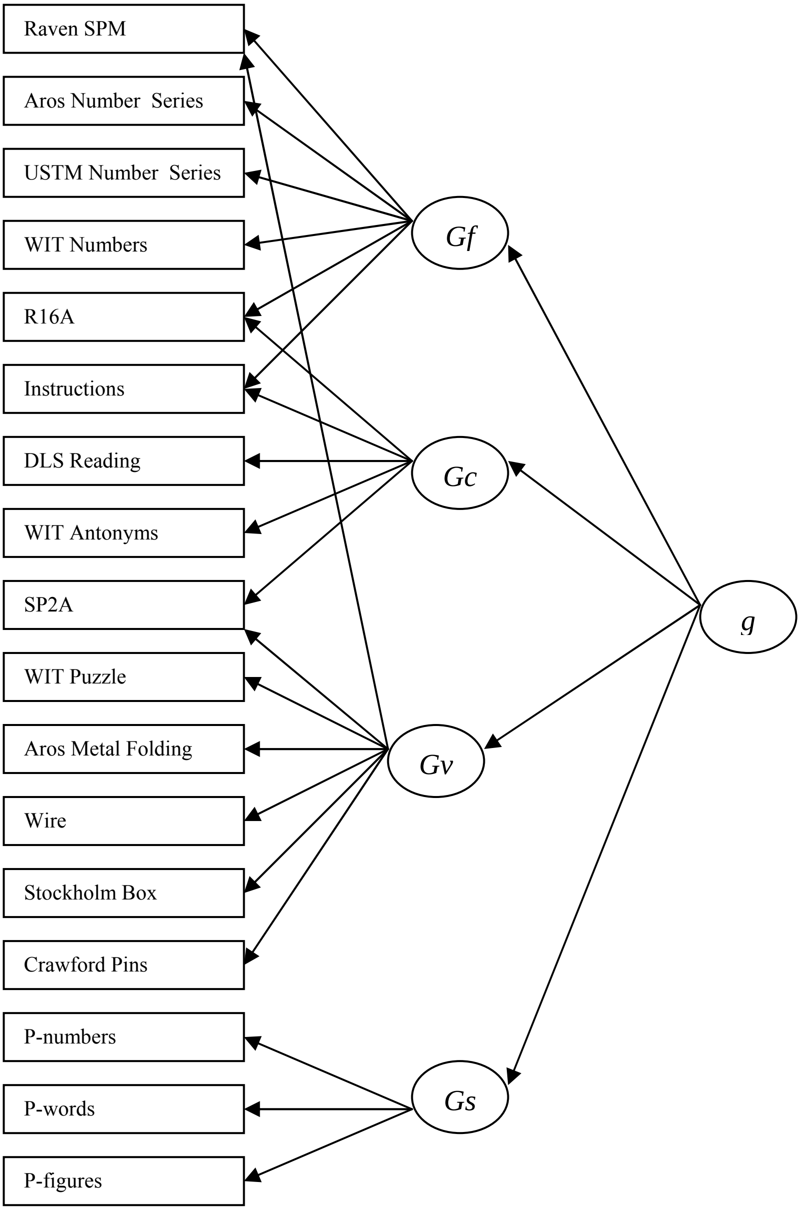

Figure 1 The hypothesized model of relations between the manifest and latent variables

Even though the multiple-group CFA procedure was used as the main analytical tool, an additional analysis was conducted with the method of correlated vectors developed by Jensen (1998), which involves computing a correlation between the factor loadings of the tests and the observed differences in means. This method was originally developed as a tool for investigating the hypothesis that the black-white performance difference on cognitive tests in the U. S. may be accounted for by a group difference in *g*, which Jensen refers to as the Spearman hypothesis. This method has been widely adopted as a method for investigating factors associated with the *g*-factor, and it has been applied to test if the Spearman hypothesis can account for performance differences between other groups, such as non-immigrant and immigrant groups. Using the method of correlated vectors te Nijenhuis & van der Flier (1997, 2003) concluded that performance differences between immigrants and non-immigrants in the Netherlands can be explained by the Spearman hypothesis in combination with language bias in tests with a strong verbal component.

However, the method of correlated vectors has been criticized on methodological grounds. Ashton & Lee (2005) observed that the *g*-loadings, and therefore the outcome of the analysis, varied as a function of which tests were included in the factor analysis. Dolan (2000) and Dolan, Roorda & Wicherts (2004) argued that the method of correlated vectors suffers from several weaknesses, which primarily are due to the fact that the method is not based on an explicit and testable model. Instead they argued that multiple-group CFA should be used to investigate the nature of group differences in cognitive performance. Thus, for both substantive and methodological reasons it is interesting to compare the results from the multiple-group CFA and the method of correlated vectors.

3 Results

In the first step of the analysis, descriptive results for the total group of subjects, as well as for the three sub-groups of subjects, are presented, and in the second step results from the model fitting are reported. In the third and final step results from application of the method of correlated vectors are reported.

3.1 Descriptive results

Table 1 presents descriptive statistics for the total group of participants.

Table 1 Descriptive statistics for the manifest variables, whole group (N = 3570)

Test label	*Mean*	*sd*	*N*
Raven	47.36	7.53	1229
Aros Number Series	12.70	3.41	1532
USTM Number Series	23.59	7.50	470
WIT Numbers	10.33	3.02	2254
R16A	12.78	5.05	1943
Instructions	23.22	9.03	3302
DLS Reading	20.37	9.03	2764
WIT Antonyms	14.89	4.96	2816
SP2A	30.01	7.69	2743
WIT Puzzle	12.13	3.96	2922
Aros Metal Folding	21.76	6.83	2461
Wire	5.89	1.59	2267
Stockholm Box	12.33	3.35	1397
Crawford Pins	26.03	5.23	1433
P-Numbers	80.58	18.28	2325
P-Letters	62.68	14.71	2326
P-Figures	60.78	13.69	2324

Most of the tests in the battery were taken by more than 50 % of the participants, and for some tests there is data for almost all subjects (e. g., Instructions, WIT Puzzle and WIT Antonyms). For a few tests, the number of participants was smaller (e. g., USTM Number Series and Raven). The proportion of cases having observations on each possible combination of tests also was satisfactory, even though for one combination of tests (Crawford Pins/USTM Number Series) there were no observations. As has been shown by Kaplan (1995) a lack of observations for some combinations of variables does not threaten the usefulness of the missing-data modeling procedure.

Table 2 presents descriptive statistics for the three subgroups of cases.

Table 2 Descriptive statistics for the tests for the subgroups

Test label	*Swedish non-immigrants N = 2358*			*European immigrants N = 620*			*Non-European immigrants N = 591*		
	Mean	*sd*	*N*	*Mean*	*sd*	*N*	*Mean*	*sd*	*N*
Raven	48.61	6.45	776	46.49	7.03	235	43.82	10.03	218
Aros Number Series	12.90	3.31	1110	12.72	3.72	218	11.59	3.42	204
USTM Number Series	25.48	6.90	259	21.82	7.55	106	20.70	7.64	104
WIT Numbers	10.37	3.00	1510	10.67	2.94	355	9.87	3.10	388
R16A	14.01	4.89	1217	11.54	4.68	344	9.97	4.49	381
Instructions	27.21	6.72	2203	17.29	7.99	581	12.91	6.58	517
DLS Reading	24.26	7.19	1870	14.58	7.34	449	9.87	5.19	445
WIT Antonyms	16.77	4.42	1890	11.33	3.80	465	10.71	3.30	460
SP2A	33.26	6.11	1763	26.46	6.84	495	21.79	5.69	484
WIT Puzzle	13.11	3.65	1918	11.22	3.77	511	9.29	3.72	492
Aros Metal Folding	23.33	6.33	1702	20.30	6.49	388	16.07	5.98	371
Wire	5.92	1.56	1488	5.98	1.53	378	5.69	1.73	400
Stockholm Box	12.92	3.37	875	11.71	3.11	241	11.02	3.02	280
Crawford Pins	26.18	5.28	1033	25.91	5.10	227	25.34	5.14	173
P-Numbers	83.70	18.10	1506	77.19	17.29	439	72.13	16.75	379
P-Letters	64.90	14.70	1507	60.79	14.53	439	56.02	12.49	379
P-Figures	63.30	13.41	1507	57.69	13.45	438	54.28	12.17	378

The proportion of cases within each of the subgroups who received the different tests was roughly the same. It may be observed, however, that there were substantial differences in level of performance between the three different groups, the Swedish non-immigrants generally performing highest, and the non-European immigrants performing lowest. The performance differences were largest for the tests hypothesized to measure *Gc*, and somewhat smaller for the tests measuring *Gf* and *Gv*. The smallest group differences were observed for the tests measuring psychomotor skills (Wire, Crawford Pins).

3.2 Modeling results

In the first modeling step the hypothesized model was fitted to the data for the whole group. This model had a fit which was not quite acceptable ($\chi^2 = 1182.96$, df = 110, $p < 0.00$, RMSEA 0.053, with a 90 % confidence interval of 0.052–0.055, SRMR 0.068). After three modifications of this model (addition of a path from *Gv* to P-Figures, and addition of paths from *Gf* to WIT Puzzle and to Aros Metal Folding), the fit was judged acceptable, even though the overall goodness-of-fit test was still highly significant ($\chi^2 = 814.16$, df = 107, $p < 0.00$). This, however, was due to the large sample size, as shown by a Root Mean Square Error of Approximation (RMSEA) value of 0.043, with a 90% confidence interval of 0.040–0.046. The Standardized Root Mean Residual (SRMR) value was 0.044, which also indicates good fit. The standardized loadings of the manifest variables on the first-order factors are presented in *Table 3*.

Table 3 Standardized factor loadings for the tests, whole group

Test label	*Gf*	*Gc*	*Gv*	*Gs*
Raven	0.59		0.24	
Aros Number Series	0.81			
USTM Number Series	0.87			
WIT Numbers	0.66			
R16A	0.60	0.29		
Instructions	0.17	0.82		
DLS Reading		0.83		
WIT Antonyms		0.84		
SP2A		0.53	0.43	
WIT Puzzle	0.38		0.58	
Aros Metal Folding	0.26		0.71	
Wire			0.43	
Stockholm Box			0.73	
Crawford Pins			0.33	
P-Numbers				0.80
P-Letters				0.91
P-Figures			0.28	0.70

The sizes of the loadings for the different tests generally agreed very well with expectations. One exception was the Instructions test, which was hypothesized to have a major loading on *Gf*, and a minor loading on *Gc*. However, for this test *Gc* was found to account for the largest part of the variance, while the contribution from *Gf* was smaller. Performance on this test thus seems to be more dependent on acquired knowledge than on reasoning ability.

The model also estimated the relations between the four first-order factors and the second-order *g* factor. The loadings were found to be 0.83, 0.80, 0.55 and 0.61 for *Gf*, *Gc*, *Gv* and *Gs*, respectively. Even though the *Gf* factor had a loading which was marginally higher than the loading for *Gc*, this loading was far from unity. A formal test of the hypothesis that the loading of *Gf* on *g* is 1.0 had to be rejected ($\Delta\chi^2 = 57.76$, $\Delta df = 1$, $p < 0.00$). It must therefore be concluded that the data for the total group of subjects did not support the hypothesis of equivalence between *Gf* and *g*.

In the next step of modeling the model was fitted to each of the three subgroups of cases separately, which yielded the model fit values presented in *Table 4*.

Table 4 Model fit estimates for the three subgroups

	Swedish non-immigrants	*European immigrants*	*Non-European immigrants*
χ^2	553	211	166
Df	107	104	106
RMSEA	0.042	0.041	0.031
90 % conf. Interval	0.039-0.046	0.033-0.049	0.021-0.040
SRMR	0.048	0.060	0.066

The model fitted excellently within all three groups, with RMSEA- and SRMR-values similar to those observed for the total group of cases. It may be observed that there was a slight variation in numbers of degrees of freedom over groups. This is due to the fact that in separating the material into three subgroups there were a few more combinations of tests without any observations, and the number of these varied for the three subgroups.

Table 5 presents the estimates of the standardized relations between the manifest variables and the first-order factors.

Table 5 Standardized factor loadings for the tests for the subgroups

Test label	*Gf*			*Gc*			*Gv*			*Gs*		
	SNI	*EI*	*NEI*	*SNI*	*EI*	*NEI*	*SNI*	*EI*	*NEI*	*SNI*	*EI*	*NEI*
Raven	0.54	0.70	0.66				0.25	0.23	0.21			
Aros Number Series	0.81	0.79	0.80									
USTM Number-Series	0.84	0.90	0.82									
WIT Numbers	0.73	0.62	0.68									
R16A	0.61	0.58	0.45	0.22	0.34	0.48						
Instructions	0.22	0.28	0.19	0.71	0.72	0.70						
DLS Reading				0.67	0.74	0.68						
WIT Antonyms				0.79	0.76	0.73						
SP2A				0.34	0.58	0.49	0.53	0.36	0.40			
WIT Puzzle	0.36	0.45	0.34				0.58	0.48	0.58			
Aros Metal Folding	0.28	0.34	0.13				0.68	0.63	0.70			
Wire							0.48	0.47	0.45			
Stockholm Box							0.75	0.65	0.68			
Crawford Pins							0.33	0.21	0.56			
P-Numbers										0.80	0.77	0.76
P-Letters										0.90	0.92	0.92
P-Figures							0.30	0.25	0.28	0.72	0.71	0.66

Note. The groups are Swedish non-immigrants (SNI), European immigrants (EI), and Non-European Immigrants (NEI)

The sizes of the loadings were generally highly similar over the three subgroups, and they also agreed very well with the results obtained for the pooled group of cases (see *Table 3*). There were a few exceptions, however.

Comparing the results for the three subgroups with the results for the total group of cases it can be noted that the loadings on the tests measuring *Gc* (i. e., Instructions, DLS Reading, and WIT Antonyms) were lower within all three subgroups than they were within the pooled group. This is likely to be due to the fact that there were large differences in level of performance on the Gc-tests between the three subgroups, and these differences appeared as individual differences when the subgroups were pooled.

The technical test SP2A, where each problem is presented with a picture and a text, was in all groups a measure of both *Gc* and *Gv*. However, this test had its highest loading on *Gv* for the Swedish group, while for the immigrant groups the highest loadings were observed for *Gc*. One reason for this may be that the written instruction makes the test more *Gc*-loaded for those who do not have Swedish as their first language. A similar pattern of differences was observed for the mathematical test R16A. This test had loadings on both *Gf* and *Gc*, and for the Swedish and European groups the highest loading were observed for *Gf*, while for the Non-European group loadings on *Gf* and *Gc* were equal.

For the manual dexterity test Crawford Pins there was a more pronounced influence from *Gv* in the non-European group than in the other groups. However, for the other *Gv*-tests the standardized loadings on *Gv* generally were highly similar over the three groups.

The main conclusion from this comparison of relations between manifest variables and the stratum II factors is that the relations generally were highly similar. However, in some cases the demands for Swedish in tests designed to measure *Gv* and *Gf* seemed to make these tests measure *Gc* to a higher extent in immigrant groups.

Table 6 presents the loadings of the first-order factors on the second-order *g*-factor.

Table 6 Standardized factor loadings of first-order factors on the g - factor and χ2 tests of the hypothesis that the loading is unity

	Swedish non-immigrants		*European immigrants*		*Non-European immigrants*	
Factor	*r*	χ^2	*r*	χ^2	*r*	χ^2
Gf	0.98	0.32	0.99	0.02	0.98	0.04
Gc	0.80	73.92	0.67	57.17	0.63	30.94
Gv	0.37	556.36	0.49	57.42	0.38	67.90
Gs	0.58	514.46	0.56	123.83	0.44	88.00

Note. The χ^2 statistics all have 1 df, and values larger than 3.84 are significant at the 5 % level.

The standardized loadings for the three subgroups were quite similar to one another, but there were some striking differences compared to the results obtained in the analysis of the pooled group of cases. For all three groups the observed loadings of *Gf* on *g* were so high that they cannot be regarded as being different from unity. Statistical tests of the hypothesis that the four first-order factors had a perfect relation with *g* are also presented in *Table 6.* For *Gf* this hypothesis could not be rejected for any of the three groups, while for all the other factors it was rejected for all groups. Thus, in contrast to the analysis of the pooled group of cases the results from the analysis of the three subgroups provide strong support for the hypothesis that *Gf* equals *g*.

The models considered so far have, for simplicity, been fitted within one of the subgroups at a time. However, such one-group models do not allow estimation of latent variable means and they do not allow statistical tests of differences model parameters over groups. Therefore a series of three-group models were also fitted, with the primary purpose of investigating group differences in latent variable means. The means are more easily interpretable within a model with correlated first-order factors than in a higher-order model, so the model compared over the three groups was the oblique model with four correlated stratum II factors (i. e., *Gf*, *Gc*, *Gv*, and *Gs*).

In the first step a model (Model 1) was fitted in which each and every parameter was constrained to be equal over groups. As may be seen in *Table 7* this model fitted poorly, indicating that there were differences between the groups for one, several or all of the model parameters.

Table 7 Fit statistics for multiple-group models

Model	*Description*	χ^2	*Df*	*RMSEA*	$\Delta\chi^2$	*Δdf*
1	Everything constrained	6833	445	0.110		
2	Free latent variable means	3744	437	0.080	3089	8
3	Free manifest variable intercepts	3306	411	0.077	438	26
4	Free residual variances in manifest variables	3035	377	0.077	271	34
5	Free variances/covariances among latent variables	1186	359	0.044	1848	18
6	Free factor loadings	967	319	0.041	220	40

Next a series of models was therefore fitted in which the constraints on model parameters over groups were successively relaxed. In Model 2 the constraints on the latent variable means were relaxed, and since this model fitted considerably better than Model 1 it may be concluded that the differences between the groups with respect to the latent variable were highly significant. In Model 3 the constraints on the manifest variable intercepts were relaxed, which implies that differences were allowed between the groups with respect to manifest variable means over and above what is accounted for by the latent variables. For this model a slightly improved fit was obtained. In Model 4 the constraints of equality over groups with respect to the residual variances of the manifest variables were relaxed. This too brought about a slight improvement of fit. In Model 5 the constraints on the covariances for the latent variables were relaxed, which caused a considerable improvement of fit. In Model 6, finally, the equality constraints on the factor loadings were relaxed, which also caused a slight improvement of fit. In Model 6 no constraints of equality remain over groups, and as may be seen in *Table 7* this model fitted excellently.

Some of the differences between the groups have already been commented upon above, so here only the differences in latent variable means will be focused upon. These were estimated from a model in which the observed variable means were added to Model 6. This model fitted well ($\chi^2 = 1191.43$, df = 345, $p < 0.00$, RMSEA = 0.045, with a 90% confidence interval of 0.043-0.048) and estimates are presented in *Table 8*.

Table 8 Estimated latent variable means for the sub-groups

	Gf	*Gc*	*Gv*	*Gs*
Swedish non-immigrants	0.00	0.00	0.00	0.00
European immigrants	-0.14	-1.84	-0.63	-0.31
Non-European immigrants	-0.40	-3.38	-1.28	-0.72

Note. The latent variable means for the Swedish non-immigrants have been set to 0, to identify the differences between latent variable means over groups.

The means are expressed in terms of standard deviation (sd) units. The means for the Swedish non-immigrants have been set to zero, so the estimates shown for the other groups are differences in latent variable means as compared to the reference group of Swedish non-immigrants.

It must be observed that the three sub-groups are not representative samples of any well-defined populations so simple generalizations cannot be made. However, in this case the profile of performance over the four latent variables is of the greatest interest, and these may be meaningfully compared over the groups. The largest differences between the groups were observed for *Gc*, for which factor the Non-European immigrants had a level of performance 3.4 sd units below the Swedish group, while the European immigrants performed 1.8 sd units below. The smallest differences were observed for *Gf*, and for this factor the Non-European group only performed .4 sd unit below the Swedish non-immigrants. The pattern of differences observed for *Gc* and *Gf* indicates that the two immigrant groups have had less opportunity to acquire the knowledge and skills measured by the *Gc* tests than have the Swedish non-immigrants. The pattern of differences for *Gs* was similar to that observed for *Gf*, while for *Gv* the differences were of an intermediate size. The fact that there were quite substantial differences also with respect to *Gv* indicates that spatial-figural knowledge and skills also are culturally determined.

3.3 Results from the method of correlated vectors

The Spearman hypothesis, as formulated by Jensen (1998), basically states that group differences in performance on cognitive tests are a function of the *g*-loading of the tests. However, the results from the multiple-group CFA analyses do not lend any support to the hypothesis that the locus of differences between the Swedish non-immigrant group and the immigrant groups is in the *g*-factor, if this factor is taken to be the same factor as the *Gf*-factor. The major source of differences rather is the *Gc*-factor. It is, therefore, of great interest to investigate which results are obtained with the method of correlated vectors.

Table 9 presents the basic ingredients of this analysis. The first two columns present the standardized group differences between the Swedish non-immigrants and the European and Non-European immigrants, respectively. These differences were computed as Cohen's d (i. e., the observed mean difference divided by the pooled within-group standard deviation). The column labelled g presents standardized factor loadings from a one-factor CFA model, which was estimated from a three-group model in which the observed variable means were allowed to vary, but all other parameters of the measurement model were constrained to be equal over the group. This model had poor fit (□2 = 8238.88, df = 448, $p < 0.00$, RMSEA = 0.121, with a 90 % confidence interval of 0.119 – 0.123), because it imposes a one-dimensional model on multidimensional data. The next four columns present estimated factor loadings from a so called nested-factor model (Gustafsson & Balke, 1993), in which four orthogonal factors were fitted to data. One factor (*Gf=g*) was related to all tests, while the others (*Gc*, *Gv* and *Gs*) were residual factors which were related to subsets of tests in the same manner as in the higher-order CFA model presented above. It should be observed that in this model there was no *Gf*-factor for the *Gf* subset of tests, which is due to the fact that the equivalence of *Gf* and *g* causes this residual factor to disappear (see Gustafsson, 2001, 2002). The four-factor model was estimated in the same manner as the one-factor model, and it had acceptable fit ($\chi^2 = 1804.73$, df = 428, $p < 0.00$, RMSEA = 0.052, with a 90% confidence interval of 0.050-0.055).

Table 9 Standardized mean differences and factor loadings from a one-factor and a four-factor model with nested factors

Test label	*Δ SNI -EI*	*Δ SNI - NEI*	*g*	*Gf=g*	*Gc*	*Gv*	*Gs*
Raven	0.29	0.74	0.50	0.65		0.26	
Aros Number Series	0.25	0.44	0.51	0.80			
USTM Number Series	0.31	0.70	0.61	0.84			
WIT Numbers	-0.07	0.12	0.37	0.67			
R16A	0.47	0.84	0.64	0.81	0.13		
Instructions	1.40	2.07	0.85	0.77	0.47		
DLS Reading	1.49	2.09	0.72	0.51	0.50		
WIT Antonyms	1.32	1.46	0.72	0.61	0.39		
SP2A	1.05	1.85	0.68	0.47	0.29	0.46	
WIT Puzzle	0.50	1.06	0.57	0.58		0.55	
Aros Metal Folding	0.57	1.12	0.58	0.52		0.65	
Wire	-0.04	0.15	0.20	0.24		0.29	
Stockholm Box	0.39	0.66	0.36	0.26		0.62	
Crawford Pins	0.08	0.26	0.16	0.16		0.25	
P-Numbers	0.34	0.64	0.41	0.46			0.65
P-Letters	0.26	0.60	0.44	0.50			0.75
P-Figures	0.39	0.67	0.45	0.46		0.26	0.63

Note. The groups are Swedish non-immigrants (SNI), European immigrants (EI), and Non-European Immigrants (NEI).

Table 10 presents the correlations among the standardized mean differences and the estimated factor loadings. The correlations between the mean differences and the *g*-factor loadings estimated from the one-factor model amounted to 0.84 for the SNI-EI difference and to 0.86 for the SNI-NEI difference, so these results provided strong support for Spearman’s hypothesis. However, the correlations with the factor loadings estimated from the four-factor model provided a different pattern of results. There were no significant correlations between the *Gf=g* factor and the mean differences, while there were strikingly high correlations between the mean differences and the *Gc* factor (0.94 and 0.88, respectively).

Table 10 Correlations between the difference in mean scores for European immigrants (EI) and non-European immigrants (NEI) as compared to the Swedish non-immigrant (SNI) group with factor loadings estimated from a one-factor model and a nested-factor model with four factors.

	SNI-EI	*SNI-NEI*	*g*	*Gf=g*	*Gc*	*Gv*	*Gs*
SNI-EI	1.00						
SNI-NEI	0.97*	1.00					
g	0.84*	0.86*	1.00				
Gf=g	0.22	0.24	0.68*	1.00			
Gc	0.94*	0.88*	0.74*	0.21	1.00		
Gv	-0.11	0.00	-0.21	-0.51*	-0.30	1.00	
Gs	-0.20	-0.21	-0.21	-0.17	-0.27	-0.23	

*Correlation is significant at the 0.05 level (2-tailed).

These results thus indicate fairly good agreement between the results of the method of correlated vectors and the multiple-group CFA analysis, when the factor loadings from the well-fitting four-factor model were used in the analysis. However, the correlated vectors analysis yielded incorrect results when it was based upon factor loadings estimated from the one-factor model. The reason for this is that the *g*-factor estimated from the one-factor model is biased in such a way that the loadings for *Gc*-tests are overestimated, while the loadings for *Gf*-tests are underestimated.

4 Discussion and conclusion

The main aim of the current study was to test a prediction derived from Cattell's Investment theory, namely that the *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 that this relationship would not hold in heterogeneous populations where subgroups differ with respect to opportunity to learn. Using a set of data consisting of Swedish non-immigrants, European immigrants, and non-European immigrants who had been tested with a Swedish test-battery, clear-cut support was obtained for the prediction: the relationship between *Gf* and *g* was only 0.83 when all subjects were treated as a single group, but it was unity within each of the three subgroups of cases. This result provides support for the Investment theory, and for the hypothesis that *Gf* is equivalent to *g*. However, the results of this study also imply that the hypothesis of *Gf*-*g* equivalence only holds true when the subjects have had approximately equally good, or equally poor, opportunities to develop the knowledge and skills measured.

The results from the current study also provide a possible explanation for why some studies have failed to establish the equivalence of *Gf* and *g*. Thus, in studies based upon heterogeneous populations the perfect relation cannot be expected to appear, even though a high relationship between *Gf* and *g* is expected. For example, in the study by Carroll (2003) previously referred to, which failed to find the perfect relation between *Gf* and *g,* the matrices analyzed were pooled across the ages from kindergarten to adulthood, and this may have caused a population heterogeneity which prevented the perfect relation to appear. These data could be reanalyzed with the data organized into homogeneous age groups to test this hypothesis.

The results of the current study thus indicate that *Gf* is a causal factor in determining individual differences in the full range of knowledge and skills measured by cognitive tests, presumably because *Gf* is involved in at least the early phases of acquisition of knowledge and skills in all domains. To the extent that learning opportunities systematically differ between different subgroups of the population the *Gf* factor will no longer take the role of being the *g* factor. It must be observed, however, that if such differential learning opportunities only affect knowledge and skill represented by a single stratum II ability dimension, the stratum III *g*-factor will still be equal to *Gf*. This is because in this situation only the residual variance of the single ability dimension will be affected by the differential learning opportunities. For the covariances among abilities to be affected in such a way that the $g = Gf$ relationship is disturbed, the differential learning opportunities must influence two or more of the abilities in the model.

As was shown in the analysis of group differences in latent variable means the two immigrant groups had a much lower level of performance on the *Gc* factor. This factor was primarily measured by tests of reading and vocabulary, and it is quite obvious that the immigrant groups had not had the same opportunities to acquire the Swedish language proficiency needed to perform well on these tests as had the Swedish non-immigrants.

It is also interesting to observe that there were quite substantial group differences in level of performance on the *Gv* factor, which indicates that cultural background exerts an influence on visual-spatial performance as well. This is also indicated by the late and prolonged maturation rate of the *Gv* factor (McArdle *et al.*, 2002). This influence may at least partially be mediated by the educational system, as is suggested by a recent study by Cliffordson & Gustafsson (in press), which demonstrated differential effects of high school educational track on development of spatial ability.

While the verbal area is subject of direct training in most cultures, the visual-spatial areas of performance are presumed to develop more spontaneously and indirectly, while the growing individual is engaged in play or other motor activities. Thus it has traditionally been thought of as relatively more "culture free". However, Maruyama (1999) has studied cultural differences with respect to attitudes towards spatial experiencing and processing, and has described distinctly different ways of relating to space and using spatial constructs. Such cultural differences can be assumed to influence also the development of individual abilities.

However, it is necessary to be cautious in making conclusions about cultural differences since the groups investigated are not representative samples from any well-defined cultural groups. Becoming an immigrant involves processes of selection and self-selection, as does the process leading up to an application for a vocational training course. It would thus be of great interest to have the current study repeated on other groups, which should be more clearly defined in this respect and preferably also should be more balanced with respect to gender composition.

While the current study provides support for Cattell's Investment theory it may be noted that the empirical evidence in support of this theory largely has been missing. For example, the hypothesis derived from the theory that *Gf* should have higher heritability than *Gc* has not generally been supported, even though Cattell (1987) reports some studies showing this to be the case. Longitudinal studies investigating cross-lagged effects of *Gf* on *Gc* also generally have failed to identify the hypothesized relations (Gustafsson & Undheim, 1992). It may also be noted that the notion of *Gf* as a biologically and genetically determined ability which has been associated with the Investment theory does not agree with findings of a strong environmental determination of *Gf* as evidenced by the Flynn effect (Dickens & Flynn, 2001), effects of schooling (Cliffordson & Gustafsson, in press) and recent findings of the fluidity of the human brain, particularly in the early years (Blair, 2006). It is obvious that further research is needed to resolve these contradictory and paradoxical results. One interesting approach to be elaborated in this research is the "mutualism" dynamical model developed by van der Maas *et al.* (in press), both as a vehicle to investigate alternative models for possible interrelations between *Gf* and *Gc* in development, and as a general framework for understanding mutual influences among abilities.

One interesting methodological finding of the current study is that the method of correlated vectors was shown to yield incorrect results when factor loadings were estimated with a simplified one-dimensional model, but that results were reasonable when factor loadings were estimated with a well-fitting four-dimensional model. This suggests that the method used for estimating the *g*-factor may be of greater importance than is usually recognized. According to the conventional wisdom of the field very much the same *g*-factor is estimated, whether this is done via a sum of scores on a heterogeneous test battery, via a principal factor or principal component solution, or via a hierarchical factor model. Even though Jensen (1998) favored the latter method, it seems that the first principal factor is the most commonly used method for identifying the *g*-factor. However, as was observed by Ashton and Lee (2005) the first principal factor tends to be biased in favour of *Gc*-tests, as was also the case with the one-factor CFA model fitted here. This effect does not seem to be caused by there being an excessive number of *Gc*-tests in the batteries, but rather by the fact that a larger proportion of the systematic variance in the *Gc*-tests is turned into common variance than is the case for *Gf*-tests. If this hypothesis is correct it implies that much of the research thought to focus on the *g*-factor has in fact focussed on *Gc*.

It may, finally, be noted that there is considerable confusion in the literature concerning the meaning and nature of the *g*-factor. Blair (2006) discussed relations between fluid intelligence and general intelligence, and rejected the idea that these are identical because of an obvious lack of agreement in many studies. However, the *g*-factors investigated in the studies reviewed by Blair (2006) were typically defined by scores on IQ-tests or as the first principal factor, which explains why only relatively low relations were found with measures of *Gf*.

References

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

Ashton, M. C., & K. Lee (2006). “Minimally biased” *g*-loadings of crystallized and non-crystallized abilities. *Intelligence*, 34, 469-477.

Blair, C. (2006). How similar are fluid cognition and general intelligence? A developmental neuroscience perspective on fluid cognition as an aspect of human cognitive ability. *Behavioral and Brain Sciences*, 29, 109-160.

Buckhalt, J. A. (2002). A short history of *g*: Psychometrics’ most enduring and controversial construct. *Learning and Individual Differences*, 13, 101-114.

Carroll, J. B. (1993). *Human cognitive abilities*. Cambridge: Cambridge University Press.

Carroll, J. B. (1996). A three-stratum theory of intelligence: Spearman’s contribution. In I. Dennis, & P. Tapfield (Eds.). Human abilities. *Their nature and measurement*, (pp. 1-18). Mahwah, 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*, (pp. 5-21). Oxford: Elsevier Science Ltd.

Cattell, R. B. (1963). Theory of 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*. New York: North-Holland.

Cliffordson, C., & Gustafsson, J.-E. (in press). Effects of age and schooling on intellectual performance: Estimates obtained from analysis of continuous variation in age and length of schooling. *Intelligence*.

Coan, R. W. (1964). Facts, factors and artifacts: The quest for psychological meaning. *Psychological Review*, 71, 123-140.

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 multigroup confirmatory factor analysis. *Multivariate Behavioral Research,* 35(1), 21-50.

Dolan, C. V., Roorda, W., & Wicherts, J. M. (2004). Two failures of Spearman's hypothesis: The GATB in Holland and the JAT in South Africa. *Intelligence,* 32(2), 155-173.

Gignac, G. (2006). Evaluating subtest 'g' saturation levels via the single-trait correlated uniqueness (STCU) SEM approach: Evidence in favor of crystallized subtests as the best indicators of 'g'. *Intelligence*, 34, 29-46.

Guilford, J. P. (1967). *The nature of human intelligence*. New York: McGraw-Hill

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, Advances in the psychology of human intelligence. Vol. 4, (pp. 35-71). Hillsdale, New Jersey: Lawrence Erlbaum Associates, Inc.

Gustafsson, J.-E. (1994). Hierarchical models of intelligence and educational achievement. In A. Demetriou & A. Efklides (Eds.) Intelligence, mind and reasoning: Structure and development. Elsevier.

Gustafsson, J-E. (2001). On the hierarchical structure of ability and personality. In J. M. Collis, & S. Messick (Eds.) Intelligence and personality. Bridging the gap in theory and measurement (pp.25-42). London: Lawrence Erlbaum Associates, Publishers.

Gustafsson, J-E. (2002). Measurement from a hierarchical point of view. In H. I. Braun, D. N. Jackson, & D. E. Wiley (Eds.) The role of constructs in psychological and educational measurement (pp. 73-95). London: Lawrence Erlbaum Associates, Publishers.

Gustafsson, J.-E., & Balke, G. (1993). General and specific abilities as predictors of school achievement. Multivariate Behavioral Research, 28(4), 407-434.

Gustafsson, J.-E., & Stahl, P.-A. (2005). STREAMS 3.0 Users Guide. Mölndal, Sweden: MultivariateWare.

Gustafsson, J.-E., & Undheim, J. (1992). Stability and change in broad and narrow factors of intelligence from ages 12 to 15 years. Journal of Educational Psychology, 84(2), 141-149.

Horn, J. L. (1968). Organization of abilities and the development of intelligence. *Psychological Review*, *75*(3), 242-259.

Horn, J. L., & Blankson, N. (2005). Foundations for better understanding of cognitive abilities. In D. P. Flanagan, & P. L. Harrison (Eds.) *Contemporary intellectual assessment: Theories, tests and issues (2nd edition)*, (pp. 41-68). New York, NY: Guilford Press.

Horn, J. L., & Cattell, R. B. (1966). Refinement and test of the theory of fluid and crystallized intelligence. *Journal of Educational Psychology*, 57, 253-270.

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

Humphreys, L. G. (1962). The organization of human abilities. *American Psychologist*, 17, 475-483.

Jensen, A. R. (1998). *The g factor. The science of mental ability*. Westport, Connecticut: Praeger Publishers.

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

Johnson, W., & Bouchard, T. Jr (2005a). The structure of intelligence: It is verbal, perceptual and image rotation (VPR), not fluid and crystallized. *Intelligence*, 33, 393-436.

Johnson, W., & Bouchard, T. Jr (2005b). Constructive replication of the visual-perceptual-image rotation model in Thurstone's (1941) battery of 60 tests of mental ability. *Intelligence*, 33, 417-430.

Kaplan, D. (1995). The impact of BIB-spiraling induced missing-data patterns on goodness-of-fit tests in factor analysis. *Journal of Educational and Behavioral Statistics*, 20(1), 69-82.

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 (2nd edition)*, (pp. 581-614). New York, NY: Guilford Press.

Landauer, T. K., & Dumais, S. T. (1997). A solution to Plato's problem: The latent semantic analysis theory of acquisition, induction, and representation of knowledge. *Psychological Review*, 104(2), 211-240.

Lohman, D. (2004). Aptitude for college: The importance of reasoning tests for minority admissions. In R. Zwick (Ed.). *Rethinking the SAT. The future of standardized testing in university admissions*, (pp. 41-56). New York and London: RoutledgeFalmer.

Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. *Intelligence*, 32, 411-424.

Mårdberg, B., & Carlstedt, B. (1998). Swedish Enlistment Battery (SEB): Construct validity and latent variable estimation of cognitive abilities by the CAT-SEB. *International Journal of Selection and Assessment*, 6(2), 107-114.

Maruyama, M. (1999). Self-heterogenization and Cultural Milieu Selection: Two New Directions in Counseling. In P. Pedersen (Ed.), *Multiculturalism as a Fourth Force* (pp. 37-72). Philadelphia, PA: Taylor & Francis Group.

McArdle, J. J., Ferrer-Caja, E., Hamagami, F., & Woodcock, R. W. (2002). Comparative longitudinal structural analyses of the growth and decline of multiple intellectual abilities over the life span. *Developmental Psychology*, 38(1), 115-142.

McGrew, K. S. (2005). The Cattell-Horn-Carroll Theory of Cognitive Abilities. Past, Present and Future. In D. P. Flanagan, & P. L. Harrison (Eds.) *Contemporary intellectual assessment: Theories, tests and issues (2nd edition)*, (pp. 136-181). New York, NY: Guilford Press.

Messick, S. (1992). Multiple intelligences or multilevel intelligences? Selective emphasis on distinctive properties of hierarchy: On Gardner's *Frames of Mind* and Sternberg's *Beyond IQ* in the context of theory and research on the structure of human abilities. *Psychological Inquiry*, 3(4), 365-384.

Muthén, B. O., Kaplan, D., & Hollis, M. (1987). On structural equation modeling with data that are not missing completely at random. *Psychometrika*, 52, 431-462.

Muthén, L., K., & Muthén, B. O. (2004). *Mplus User's Guide*. Third Edition. Los Angeles, CA: Muthén & Muthén.

Raven, J. C., Court J. H., & Raven, J. (1998). *Raven Manual: Section 1 General Overview* (1998 ed.). Oxford: Oxford Psychologists Press Ltd.

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.

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

Robinson, D. L. (1999). The 'IQ' factor. Implications for intelligence theory and measurement. *Personality and Individual Differences*, 27, 715-735.

Spearman, C. (1904). General intelligence: objectively determined and measured. *American Journal of Psychology*, 15, 291-293.

Spearman, C. (1923). *The nature of 'intelligence' and the principles of cognition*. London: Macmillan.

Spearman, C. (1927). *The abilities of man*. London: MacMillan.

Sternberg, R., & Kaufman, J. C. (1998). Human abilities. *Annual Review of Psychology*, 49, 479-502.

te Nijenhuis, J., & van der Flier, H. (1997). Comparability of the GATB scores for immigrants and majority group members: Some Dutch findings. *Journal of Applied Psychology*, 82, 675-687.

te Nijenhuis, J., & van der Flier, H. (2003). Immigrant-Majority group differences in cognitive performance: Jensen effects, cultural effects, or both. *Intelligence*, 31, 443-459.

Thurstone, L. L. (1938). Primary Mental Abilities. *Psychometric Monographs*, no. 1.

Undheim, J. O. (1981). On intelligence II: A neo-Spearman model to replace Cattell's theory of fluid and crystallized intelligence. *Scandinavian Journal of Psychology*, 22, 181-187.

Undheim, J. O., & Gustafsson, J. E. (1987). The hierarchical organization of cognitive abilities: Restoring general intelligence through the use of linear structural relations (LISREL). *Multivariate Behavioral Research*, 22, 149-171.

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

Vernon, P. E. (1961). *The structure of human abilities,(2nd edition).* London: Methuen.

Publication series published by the Institute for Labour Market Policy Evaluation (IFAU) – latest issues

Rapporter/Reports

2007:1 Lundin Daniela ”Subventionerade anställningar för unga – en uppföljning av allmänt anställningsstöd för 20–24-åringar”

2007:2 Lundin Daniela, Eva Mörk & Björn Öckert ”Maxtaxan inom barnomsorgen – påverkar den hur mycket föräldrar arbetar?”

2007:3 Bergemann Annette & Gerard van den Berg ”Effekterna av aktiv arbetsmarknadspolitik för kvinnor i Europa – en översikt”

2007:4 Junestav Malin ”Socialförsäkringssystemet och arbetsmarknaden – politiska idéer, sociala normer och institutionell förändring – en historik”

2007:5 Andersson Christian ”Lärartäthet, lärarkvalitet och arbetsmarknaden för lärare”

2007:6 Larsson Laura & Caroline Runeson ”Effekten av sänkt sjukpenning för arbetslösa”

2007:7 Stenberg Anders ”Hur påverkar gymnasialt komvux löneinkomster och vidare studier?

2007:8 ”Forslund Anders & Kerstin Johansson ”Lediga jobb, arbetssökande och anställningar – den svenska matchningsfunktionen”

2007:9 Kennerberg Louise ”Hur förändras kvinnors och mäns arbetssituation när de får barn?”

2007:10 Nordin Martin ”Invandrares avkastning på utbildning i Sverige”

2007:11 Johansson Mats & Katarina Katz ”Underutnyttjad utbildning och lönegapet mellan kvinnor och män”

2007:12 Gartell Marie, Ann-Christin Jans & Helena Persson ”Utbildningens betydelse för flöden på arbetsmarknaden”

2007:13 Grönqvist Hans & Olof Åslund ”Familjestorlekens effekter på barns utbildning och arbetsliv”

2007:14 Lindqvist Linus ”Uppföljning av plusjobb”

2007:15 Sibbmark Kristina ”Avidentifierade jobbansökningar – erfarenheter från ett försök i Göteborgs stad”

2007:16 Hesselius Patrik & Malin Persson ”Incitamentseffekter och Försäkringskassans kostnader av kollektivavtalade sjukförsäkringar”

2007:17 Eriksson Stefan & Jonas Lagerström ”Diskriminering i anställningsprocessen: resultat från en Internetbaserad sökkanal”

2007:18 Eriksson Robert, Oskar Nordström Skans, Anna Sjögren & Olof Åslund ”Ungdomars och invandrades inträde på arbetsmarknaden 1985–2003”

2007:19 Agerström Jens, Rickard Carlsson & Dan-Olof Rooth ”Etnicitet och övervikt: implicita arbetsrelaterade fördomar i Sverige”

2007:20 Bennmarker Helge, Kenneth Carling & Anders Forslund ”Vem blir långtidsarbetslös?”

2007:21 Edmark Karin ”Strategisk konkurrens i kommuners utgiftsbeslut för barnomsorg, skola och äldreomsorg”

Working Papers

2007:1 de Luna Xavier & Per Johansson “Matching estimators for the effect of a treatment on survival times”

2007:2 Lundin Daniela, Eva Mörk & Björn Öckert “Do reduced child care prices make parents work more?”

2007:3 Bergemann Annette & Gerard van den Berg “Active labor market policy effects for women in Europe – a survey”

2007:4 Andersson Christian “Teacher density and student achievement in Swedish compulsory schools”

2007:5 Andersson Christian & Nina Waldenström “Teacher supply and the market for teachers”

2007:6 Andersson Christian & Nina Waldenström “Teacher certification and student achievement in Swedish compulsory schools”

2007:7 van den Berg Gerard, Maarten Lindeboom & Marta López ”Inequality in individual mortality and economic conditions earlier in life”

2007:8 Larsson Laura & Caroline Runeson “Moral hazard among the sick and unemployed: evidence from a Swedish social insurance reform”

2007:9 Stenberg Anders “Does adult education at upper secondary level influence annual wage earnings?”

2007:10 van den Berg Gerard “An economic analysis of exclusion restrictions for instrumental variable estimation”

2007:11 Forslund Anders & Kerstin Johansson “Random and stock-flow models of labour market matching – Swedish evidence”

2007:12 Nordin Martin “Immigrants’ return to schooling in Sweden”

2007:13 Johansson Mats & Katarina Katz “Wage differences between women and men in Sweden – the impact of skill mismatch”

2007:14 Gartell Marie, Ann-Christin Jans & Helena Persson “The importance of education for the reallocation of labor: evidence from Swedish linked employer-employee data 1986–2002”

2007:15 Åslund Olof & Hans Grönqvist “Family size and child outcomes: Is there really no trade-off?”

2007:16 Hesselius Patrik & Malin Persson “Incentive and spill-over effects of supplementary sickness compensation”

2007:17 Engström Per & Patrik Hesselius “The information method – theory and application”

2007:18 Engström Per, Patrik Hesselius & Malin Persson “Excess use of Temporary Parental Benefit”

2007:19 Eriksson Stefan & Jonas Lagerström “Detecting discrimination in the hiring process: evidence from an Internet-based search channel”

2007:20 Agerström Jens, Rickard Carlsson & Dan-Olof Rooth “Ethnicity and obesity: evidence of implicit work performance stereotypes in Sweden”

2007:21 Uusitalo Roope & Jouko Verho “The effect of unemployment benefits on re-employment rates: evidence from the Finnish UI-benefit reform”

2007:22 Edmark Karin “Strategic competition in Swedish local spending on childcare, schooling and care for the elderly”

2007:23 Valentin Kvist Ann & Jan-Eric Gustafsson “The relation between fluid intelligence and the general factor as a function of cultural background: a test of Cattell’s investment theory”

Dissertation Series

2006:1 Hägglund Pathric “Natural and classical experiments in Swedish labour market policy”

2006:2 Savvidou Eleni “Technology, human capital and labor demand”

2006:3 Söderström Martin “Evaluating institutional changes in education and wage policy”

2006:4 Lagerström Jonas “Discrimination, sickness absence, and labor market policy”

2006:5 Johansson Kerstin “ Empirical essays on labor-force participation, matching, and trade”