CONTENTS

CHAPTER 1

Higher Education Interactive Diagnostic Inventory (HEIDI):
Introduction

1. General introduction

Predictors of first year study success in Psychology at the Open University

The study of psychology has been flourishing for several years at the Open University of the Netherlands (OUNL). On a yearly basis about 1200-1500 student enroll, making the department of psychology of the OUNL with about 7000 students the largest department of psychology in the Netherlands, measured by its student population (Zelfstudie Psychologie OUNL, 2000).

This popularity, however, has its drawbacks. One of the most problematic aspects is the low success ratio of first year students. Among these students, who have to pass two exams based on Gleitmans *Introduction To Psychology* (1993), the success ratio is about 30% (Zelfstudie Psychologie OUNL, 2000). This is much lower than comparable ratio's in higher education (HOOP, 1996).

A possible explanation for this low success ratio is that more than half of the students who enroll for Psychology at the OUNL in fact are not apt for the academic study of psychology. One could call this unrealistic optimism from the viewpoint of enrolling students. This unrealistic optimism is triggered for a great deal by the fact that there are no formal requirements regarding previous education for an academic study at the Open University. Opposed to the regular universities, which require High school (VWO), anybody can enroll at the OUNL. This means that many of the first year students may lack necessary academic capacities, a fact they are unaware of because they have no feedback on previous educational achievements. Furthermore, students at the OUNL lack the social and study support so familiar in normal academic environments. This makes studying at the OUNL a highly individual affair, with study success being largely dependent on individual student factors. The question then arises, what are these individual factors, and in what way are they related to or responsible for study success? Being able to provide an answer to the question above can be of great help for enrolling students. They could obtain at a prognosis of their chances of (first year) academic achievement. The aim of this thesis then is twofold:

1) Making an inventory of relevant and robust predictors for first year academic success for the study of psychology at the OUNL.

2) Creating simple, reliable and valid measurement instruments which can be implemented in a web based diagnostic tool called *Higher Education Interactive Diagnostic Inventory* (HEIDI).

Together the fulfillment of these aims should create a basis for realistic feedforward about one's chances of passing the first academic year of the study of psychology at the OUNL.

The use of the internet grants an efficient and effective processing and feedback of prediction results (Buchanan & Smith, 1999). Potential students can log on to the HEIDI site and complete the measurement instruments as given. With the feedback from their scores, a more realistic decision can be made: am I fit for this study? What are my chances? Where lie my deficiencies, and can I do something in advance to improve my chances?

It is essential to bear in mind that prediction from the measurement instruments is optional; there is no such thing as a cut off score, below which students are not allowed to enroll. So the results from the HEIDI database are not intended to be used in the way the Graduate Record Examination (GRE) is being used for United States university selection policy (Dawes, 1973; Wood & Wong, 1992; Briel, O'Neill, & Scheunemann, 1993). Together, the results of the predictive scales form a 'free diagnostic advice', on which students themselves make a decision whether or not to enroll. The sole purpose of the feedback mechanism is to generate a more realistic basis for study enrollment. Otherwise the Open University would be compelled to change its name.

A feedback mechanism in the way mentioned above would serve the following purposes:

-Serving and utilizing individual student interests

Individual students are in need of adequate information about their chances of study success. Not providing this information means that about 60% of 1200 students

7

spend about a total of € 286,000[1] on spilt study efforts, leaving them with the frustration and disappointment of an unattained personal and societal goal.

-Serving and utilizing educational and academic interests

At the macro level, there will be a greater utility of the study of Psychology at the OUNL. Also, the level of graduates may improve, a fact which should, with the future of psychology in mind, not be disregarded. Aiken, West, Sechrest & Reno (1990) report on the dangers of basic methodological knowledge no longer transferred to future generations of students. (At the OUNL the statistical methodological courses have the lowest success rate).

In addition, it is likely that students enrolling on an advice basis are better fit for the study of Psychology, so they will spend less study time before graduating. (The average study time for students who complete the study of psychology at the moment is nine years).

Implementation of HEIDI can have profound consequences for higher educational practice. Students can individually estimate their chances in a desired field of academic study. HEIDI provides potential students with the equivalent of a diagnostic advice, given by a study counselor. But HEIDI's advice is objective, based on actuarial statistical prediction (Meehl, 1996). If the variables incorporated in the primary study appear to be relevant for a broad academic application, HEIDI can be used for the complete domain of higher education. Because the characteristics of every potential and enrolling student are automatically processed in the database, HEIDI's norms and predictions remain up to date. Changes in estimated predictive validity will provide higher education with valuable feedback about the demands it sets for (enrolling) students, creating the possibility to adapt its curriculum to the capacities of its target group members. Overall deficiencies could be compensated with special courses and training[2]. As noticed above, HEIDI's function is primarily one of dynamic assessment (Lidz, 1998); it's not about creating a cut off score, but it is primarily for signaling and overcoming deficiencies.

---

[1] 720 students x two introductory courses = 720 x 199 x2 = € 286,560
[2] At the OUNL a basic course in mathematics is provided for to help students overcome difficulties with statistics.

-Serving scientific interests

Up till today it is still unclear which individual factors to what extent contribute to study success in higher education. There is a broad consensus that intelligence (*g*) correlates about .50 with success in education (Neisser, Boodoo, Bouchard, Boykin, Brody, Ceci, Halpern, Loehlin, Perloff, Sternberg & Urbina, 1996). But as Prins (1997) showed, intelligence is a necessary but not sufficient condition for success in higher education. The only dissonant concerning the intelligence paradigm is probably Sternberg (1989), who claims that the role of 'g' intelligence is greatly overrated. Sternberg is a strong advocate of what he calls 'practical intelligence' (Sternberg, Wagner, Williams & Horvath, 1995). But there is still no firm consensus on the construct and predictive validity of practical intelligence, and empirical studies on this subject are rare and unconvincing[3] (Schmidt & Hunter, 1993).

Research in the personality domain (Barrick & Mount, 1991; Okun & Finch, 1998; Hurtz & Donovan, 2000) shows some explanatory power for certain personality variables concerning work and study success (e.g. conscientiousness), while also self efficacy (Bandura, 1997), the belief in one's own competencies, may be of importance. But overall it is still unclear which personality characteristics contribute to study success (for an overview, see De Raad & Schouwenburg, 1996).

Dochy (1992, 1993) emphasizes the importance of relevant prior knowledge for attaining success in a specific domain of study. This could be a relevant aspect, because, in view of the large drop-out figure among first year students of Psychology at the OUNL it is very well conceivable that enrolling students start with 'false' knowledge and expectations of the academic field of psychology.

A subject that has received relatively little attention concerns biographical aspects, such as living situation, effects of earlier education, marital status, children, having a full time job, etcetera. From research mainly in industrial psychology it is known that the effect of biographical variables or biodata on work outcomes can be of a moderate to strong magnitude (Reilly & Chao, 1982). So these aspects could also be of influence on the results of a career in higher education (Dyer, 1987; Federici & Schuerger 1974; Melamed, 1992). Although other predictors like learning styles (Vermunt, 1990)

---

[3] See also the special issue of *Intelligence* (2003, 31) on the subject of Practical Intelligence,

come to mind, we choose to restrict ourselves to the factors mentioned above. The main reason is that they seem student relevant and robust, and that they can be efficiently administered using the internet.

In Figure 1.1 the most prominent individual factors relating to study success in higher education for our study are summarized.
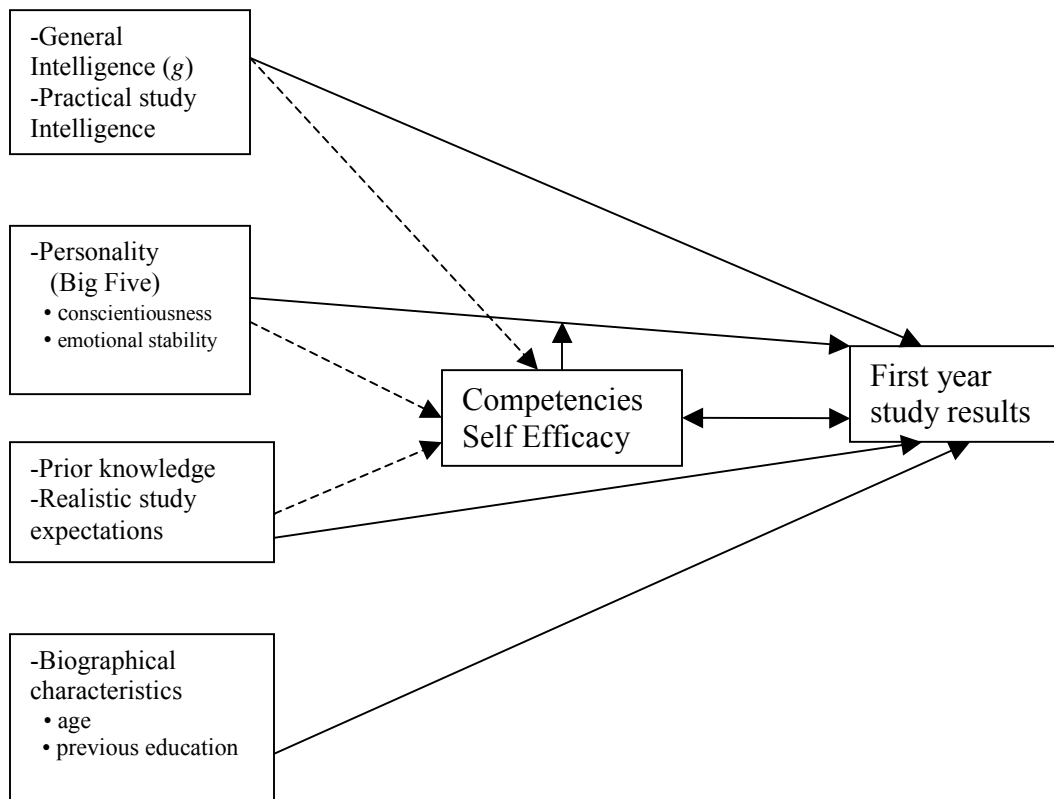

Figure 1.1: Individual factors relating to study success

(Adapted from Schouwenburg, 1998)

The purpose of this study is to construct and apply five groups of independent predictors for first year study success for psychology students of the OUNL. These five groups of predictors are:

(1) Intelligence (*g*) and Practical Study Intelligence;

(2) Big Five personality dimensions, with a possibly study facilitating role for conscientiousness and emotional stability;

(3) Prior Knowledge and realistic study expectations;

(4) Biographical aspects;

(5) Self efficacy as an independent construct and as a possible intervening variable for personality dimensions.

Although Figure 1 suggests direct and indirect effects and mediating/moderating roles (Self Efficacy) as well, these aspects do not form part of the initial studies. The emphasis of the individual studies is on estimating the reliability and predictive validity of the five predictors mentioned above. For practical reasons, and from a viewpoint of construct validity, some studies will be combined. If possible, hypotheses will be formulated for the individual studies, but the main character of each study is explorative.

If the instruments show sufficient psychometric and predictive qualities they will be incorporated in the HEIDI web based inventory. Once installed, HEIDI will function as a continuous data base of relevant student characteristics. Longitudinal studies will be possible, gaining insight in relations between predictors, their incremental validity, and their importance across the academic curriculum. Future research can also show if the HEIDI instruments predict academic success for other studies apart from psychology. Additional predictors, like vocational scales and learning styles can be included on demand, making tailor made advice for potential students of higher education a realistic option.

CHAPTER 2

The effects of numerical-verbal and practical intelligence

(Submitted for publication: *Journal of Educational Psychology*)

2. The effects of numerical-verbal and practical intelligence

Recently, at the Open University of the Netherlands (OUNL) several studies were initiated to gain a more complete understanding of what factors constitute study success for first year psychology students at the OUNL (Bloemers, 1999). At the moment (2002), about 60% of the first year psychology students of the OUNL drop out within one year. Possible causes for this drop out are deficiencies in study skills, information processing capacities, available study time and an underestimation of the contents of the academic curriculum. In four successive studies the role of intelligence, personality, biographical characteristics, prior knowledge and study expectations and self efficacy for first year study success will be investigated. The results of these studies should lead to a web based diagnostic study inventory called Higher Education Interactive Diagnostic Inventory (HEIDI), (Bloemers, 1999). The main purpose of HEIDI is to facilitate a more meaningful and successful choice of study. Students can profit from a realistic and solid advice, giving them feedback, or more important, feedforward on their chances of study success in a given domain of study. The faculty, for a substantial part accountable on its percentage of graduates, will profit from an increase in graduates.

The aim of this first study is to tackle the role of intelligence as an explanatory factor for first year study success. Intelligence will be approached from a multiple perspective, looking both at the role of 'classical' numerical-verbal intelligence, as well as the role of more recently developed constructs like 'practical intelligence' (Sternberg, 1985, 2000). Hypotheses stated are that both numerical-verbal and practical intelligence will manifest themselves as having a significant influence on first year study results of psychology students at the OUNL.

First we will give a brief overview of developments in the field of intelligence and 'practical intelligence'. Theoretical and empirical issues are highlighted, with an emphasis on the practical use of these different 'forms' of intelligence for academic success.

Intelligence

After a century of theorizing and research, the term 'intelligence' still remains in the centre of attention. Having undergone various 'pendulum swings' (Plomin & Petrill, 1997), from a more genetic psychometric side (e.g. Boring, 1923) illuminating constituent factors, to a more environmental practical side emphasizing daily use, the pendulum now somewhat has come to rest in a status quo. This status quo means that there is considerable agreement on the value of the 'psychometric' interpretation of intelligence (Gottfredson, 1998; Jensen, 1998, Carroll, 1993; Neisser et al., 1996), but some researchers, mostly to be found in Sternberg spheres, keep searching for alternate forms and explanations (Sternberg, Wagner & Okagaki, 1993; Sternberg, 1995).

What various forms of intelligence exist? There are numerous introductions on intelligence (Sternberg & Detterman, 1986; Ceci, 1993; Tomic & Van der Molen, 1997). Goff and Ackerman (1992) give a useful summary, differentiating between three major models of intelligence:

(1) The two factor model of intelligence (Spearman, 1904, 1927);

(2) The multiple factor model (Thurstone, 1938; Guilford, 1967)

(3) The hierarchical model (Horn & Cattell, 1966; Carroll, 1983).

(1) The two-factor model of intelligence.

Generally, Spearman (1904) is assumed to be the father of 'general' intelligence, also known as *g*. Spearman showed two factors to be at the root of intelligence test measures: a general factor called *g* and more specific factors called *s*. Factor analyzed measures on various intelligence tests show a high intercorrelation or first principal factor. For explaining the (relatively small) remaining variance, specific independent factors have to be assumed, such as spatial intelligence. Although criticized by Thurstone (1938), Guilford (1967) and Gardner (1983), Spearman's intelligence theory has been supported empirically on many occasions. Taking the amount of empirical evidence (Jensen, 1998), it nowadays makes no longer sense to deny the fact that (*g*) intelligence test scores point to differences in performance on real life tasks, as encountered in work and education (Jensen, 1998; Schmidt & Hunter, 1998, Gottfredson, 1997). In particular the work of Schmidt and Hunter in Industrial and Organizational psychology has demonstrated that

the importance of *g* increases with situational complexity (g-loadedness; Hunter, Schmidt & Judiesch, 1990; Gottfredson, 1997).

(2) The multiple factor model of intelligence

The multiple factor approach to intelligence is represented by Thurstone (1938), Guilford (1967) and Gardner (1983). Thurstone and Gardner both recognize seven independent forms of intelligence, although they differ highly on the nature of these seven forms of intelligence. Guilford created his *structure of intellect model*, and claims, in its most recent version, the existence of 180 independent intelligence abilities (Guilford, 1988). Although some intelligence abilities have shown to be more or less independent of *g* (e.g. spatial insight), Goff and Ackerman observe that the multiple factor approach to intelligence has grown to '….unwieldy complexity (Humphreys, 1962; McNemar, 1964)' (Goff & Ackerman, 1992, p. 538). This is mainly due to lack of empirical evidence for the claimed intelligence factors or forms of intelligence.

(3) The hierarchical model of intelligence

While authors in the multiple intelligence 'school' tried to broaden and articulate the concept of intelligence to explain differences in various forms of daily behavior, authors of the hierarchical approach try to reduce the amount of (possible) intelligence factors by looking for more parsimonious models. In words of Goff and Ackerman:

> 'These models consider common factors and the relations among them. Thus, each narrow ability is subsumed by a higher order ability that is greater in breadth. For instance, abilities such as verbal fluency, verbal comprehension, and verbal reasoning might be narrow ability components of a broader general verbal ability.' (Goff & Ackerman, 1992, p. 538).

Since Cattell's (1963), Horn & Cattell's (1966) and Carroll's (1993) factor analytic studies, general intelligence or *g* is generally viewed as consisting of a fluid component *(Gf)* and a crystallized component *(Gc). Gf* points to flexibility in information processing, general (nonverbal) reasoning and being able to solve new problems. *Gc* represents domain specific *g*-environment interactions, more pointing to accumulated (experiental) verbal and conceptual knowledge. In Carroll's (1993) factor analytic study, a hierarchical three stratum model of intelligence emerged as the most sophisticated refinement of the *g*

approach. Within Carroll's model, *g* as a third order factor is at the highest level of the hierarchy. Underneath *g* reside *Gf*, *Gc*, Visual Perception, Perceptual speed, Learning and Memory, Knowledge and Achievement and Ideational Fluency as second order factors. These second order intelligence factors emanate in various specific third order factors such as verbal comprehension and inductive reasoning. Broad acceptance for Carrolls model has come from various sources (Neisser et al., 1996). According to Te Nijenhuis, Evers and Mur (2000), hierarchical models form the best validated models within the field of psychometric intelligence.

Practical Intelligence

The concept of 'practical intelligence' is inherently linked to the work of Sternberg who takes a special place among researchers on intelligence. Although probably best fitted within the 'multiple intelligences' approach, Sternberg developed an alternative theory of intelligence (Sternberg, 1985, 1996). Sternberg criticizes the superiority of *g* mainly from its measurement context (psychometric tests); most prominent in his writings is thus an aversion with traditional test based measurements of *g*. According to Sternberg *g* might be a successful predictor in an academic (test) environment, but it may be far less successful when looking at the context of every day behavior. Sternberg therefore broadened the concept of intelligence to what he calls the 'triarchic theory of intelligence' (Sternberg, 1985, 1996). In the definitive form of this theory, intelligence is made up of three basic forms: academic or analytical intelligence, practical intelligence and creative intelligence. Together, these forms should cover what Sternberg calls 'successful intelligence' (Sternberg, 1996). Successfully intelligent people in Sternberg's vision capitalize on their effective intelligent capacities, thereby compensating their intelligence weaknesses.

According to Sternberg 'practical intelligence' manifests itself in problems concerning daily behavior. These problems have the following characteristics (Sternberg, Wagner, Williams & Horvath, 1995):

-They are not yet (verbally) formulated

-They have personal value

-No information about the solution is given

-They are related to everyday experience

-They are ill-defined

-They imply more than one 'good' answer

-They can be handled in more than one way

The points above can best be taken as a description and critique of test problems stated in multiple choice format, showing strict conditions under which reliability, validity and fairness of such a test emerge. But criticizing standard psychometric measures of intelligence, as Sternberg does, does not prove the existence or the implied 'added value' of something like practical intelligence. Criticizing the multiple choice razor as the only criterion for real world success is one thing, but postulating a new form of intelligence behind the everyday behavior of people just from a dislike of conventional intelligence test formats is quite something else. Sternberg and colleagues contaminate the concept of what might be called 'maximal behavior' with 'typical behavior' (Goff & Ackerman, 1992; Hofstee, 2001). Their definition of practical intelligence as '…the ability to understand one's environment and to use this knowledge in figuring out how best to achieve one's goals' (Williams, Blythe, White, Sternberg & Gardner, 1996), where the ability refers to '….the ability involved when intelligence is applied to real world contexts' (Sternberg et al., 2000) does in fact no more than saying that g, mixed with certain personality features, is an important aspect of successful every day behavior (especially e.g. in taking frequent multiple choice exams).[4]

Standard intelligence tests are designed to measure 'maximal' behavior, which is a manifestation of one's capabilities to solve problems within a given time limit. This can be contrasted with measurements of 'typical' behavior, were the individual is free to express his or her own opinion on a stated problem. Personality questionnaires normally take this latter form. Measurements of maximal behavior show greater predictive power of real life criteria, such as job performance and study success (Schmidt & Hunter, 1998). But maximal behavior predictions, apart from not being perfectly reliable and valid, will leave some variance unexplained, meaning that there is room for typical (and g-crystallized) behavior to move in. But this kind of behavior should not be confounded

---

[4] For an extended discussion on the concept of 'practical intelligence' and its empirical backing, see the special issue of *Intelligence* (2003, 31).

with general intelligence. Research in the field of personality on the Big Five shows that only the fifth factor, known as Openness, Intellect or Intellectual Autonomy (Hendriks, 1998) has a low but significant correlation with intelligence measures.

Empirical support for the concept of practical intelligence has been scarce up till now (e.g. Denny & Palmer, 1981; Cornelius & Caspi, 1987; Scribner, 1984; Ceci & Liker, 1984). More concrete evidence comes from a recent meta-analysis by McDaniel, Morgeson, Finnegan, Campion and Braverman (2001). These authors investigated the predictive validity of what they call 'Situational Judgment Tests', tests assessing the capacity to solve practical job related problems. According to the authors, the *Tacit Knowledge Inventory for Managers* (Wagner & Sternberg, 1991) resembles in fact the concept of situational judgment. The overall validity of these instruments was .34, with predictive studies showing lower overall validity (.18) than concurrent studies (.35). The relationship of Situational Judgment Tests with cognitive ability was relatively high, being .46. This latter aspect could be due to the fact that in situational tests the items are verbally stated, suggesting a relationship with verbal intelligence.

The main question in the studies mentioned above is the issue whether behavioral outcomes ('intelligent behavior') should be taken as proof of some alternate form of intelligence (practical/tacit), or are manifestations of acquired (situational) knowledge, that is, a manifestation of *Gc* (for an extensive discussion, see Sternberg & Wagner, 1993; Schmidt & Hunter 1993). Sternberg et al., claim an independent identity for an alternative form of intelligence ('practical intelligence'), because behavioral outcomes of this form of behavior according to them show low to zero correlations with existing psychometric test (*Gf*) results. But this phenomenon can be very well explained within the hierarchical intelligence paradigm: it just illustrates highly specific *Gc* variance, probably in combination with or due to the fact that *Gc* accumulates, whereas *Gf* decreases during the lifespan (Horn, 1982).

The most obvious conclusion in the conceptual debate on the existence of 'practical intelligence' may be that meaningful results on psychometric intelligence tests require a solid western based education as a sine qua non. However, not being educated in a conventional way, or having a low educational level does not mean that people are stupid in any information processing way. Motivation and practice, as well as the

accumulation of 'tacit knowledge' may well make up for what Polanyi (1967) called 'personal knowledge'. The result can be a highly effective amalgam of domain specific behavior, as studies by Sternberg and others (Sternberg, 2000) have demonstrated (see McDaniel, Morgeson et al., 1991, for a methodological critique). These results however are no proof that there is any alternate form of intelligence, apart from *g*. They indicate that (low) *g* measures predict little about one's (cognitive) performances in a *highly specific* domain, showing that there is some specific situational variance left in people's *g* estimates. But *g* denotes the aggregation of cognitive functioning over an (infinite) number of situations, which is quite something else. The whole argument on situation specific variance unexplained by *g* scores can be compared to the discussion on the stability of personality traits, where aggregation was the crucial factor in showing stability (Epstein, 1977). *G* is a cross situational aggregate, and thus by definition there is room for situation specific non-*g* variance.

To resolve this conceptual entanglement of the culturally specific meaning of any measurement of intelligence, the solution lies in measurements of intelligence on a more basic (and more) culture free biological level, such as reaction/inspection time and speed of information processing (Jensen, 1998).

Intelligence and study success

The general opinion is that *g* is a fairly good predictor of the various tasks of life (Gordon, 1997; Gottfredson, 1997). Being able to complete some form of education is probably one of the most universal of these tasks. In their task force article on intelligence, Neisser and others (1996) estimate the correlation between IQ and study results at about .50. Data from various studies on the Graduate Record Examination, widely used by United States Universities as an admission instrument (Kuncel, Hezlett & Ones, 2001) corroborate this finding. Others, like De Klerk (1983), have reported even higher correlations. In a study concerning psychology freshmen, Elshout and Veenman (1990) found highly intelligent students to have a better learning performance. This study showed sustained durability across the curriculum, especially for complex learning tasks. Furthermore, highly intelligent students in this study showed a better learning strategy and more discipline.

Other researchers are more critical of general intelligence as being the dominant factor responsible for study success. Meyer (1999) and Sternberg (1996) point to social aspects and practical skills in relation to study success. Their arguments, however, relate to the specific curriculum being measured as well: a high correlation between the score on psychometric intelligence tests and study results should for a large part result from similarities in predictor-criterion content, form and required skills. From the viewpoint of fairness, reliability and validity, multiple choice instruments have shown their superiority, making the discussion on criterion-predictor similarities somewhat obsolete.

Practical intelligence and study success

There are few studies that focus on the relation of practical intelligence, c.q. tacit knowledge, and study success. In one study on the relation between tacit knowledge and study success, Williams and Sternberg (see Sternberg, Wagner & Okagaki, 1993) correlated a tacit knowledge study questionnaire with an academic index and an adaptation index. The academic index was a composite, consisting of Grade Point Average (GPA) and Scholastic Aptitude Test (SAT) scores. The adaptation index consisted of a composite of subjective experiences about well being at school and self perception of happiness. Williams and Sternberg found no significant correlation between the two indexes. Several items of the tacit knowledge test showed a substantial multiple correlation with the academic index and the adaptation index, with R^2's of .43 and .63. In a related study, the 'Practical Intelligence for School Program' (Williams, 1996) a positive correlation was found between practical intelligence and study results. Practical intelligence in this study was based on social skills, study skills and self management.

The conclusion is that even the proponents of practical intelligence do not and cannot doubt the predictive power of *g* in explaining study success. But studies on 'practical intelligence' show that specific practical skills and insights can explain some of the variance in study success. An important role seems to be the curriculum content. Success in normal, highly *g* loaded academic contexts, is best predicted by conventional intelligence tests. But no academic curriculum has 100% *g* saturation, so there is room for specific abilities and skills as well.

At the OUNL, a substantial part of the curriculum consists of processing large parts of verbal written information, the results being tested with multiple choice exams. So there is a high *g* loadedness for both study contents and criterion measurements. Further upward in the curriculum, more practical and personalized tasks have to be passed. At entrance, students have a mean age of about 35, so as practical intelligence seems to increase with age (Cornelius & Caspi, 1987), these students should show some evidence of it. The predictive value of this practical intelligence for first year study results is another issue. But it is important to find out whether non-*g* factors can contribute to study success. The OUNL is an 'open' university; about 30% of the enrolling students have a relatively low education (below high school level). So it is our aim to find out if study results of this specific group, being probably low on *g* measures (with a downward perspective), can be boosted and/or compensated in some way.

To answer our main question, 'What is the influence of *g* and practical intelligence on the study results of first year OUNL psychology students' we decided:
- To use two existing separate measures of high *g* loaded conventional intelligence tests, a verbal and a numerical reasoning test.
- To construct a specific OUNL 'practical study intelligence questionnaire'.

Hypotheses
Hypotheses are that the verbal and numerical intelligence tests scores will show a significant positive relationship with study success, being measured by the two first year exam results. For the practical study intelligence questionnaire we expect a low to negative correlation with the two conventional intelligence tests, and a low but significant positive correlation with first year exam results.

<p align="center">**Method**</p>

Instruments

  Conventional intelligence tests

Two standard academic intelligence subtests were chosen from what is known as the Drenth Series (Drenth & Hoolwerf, 1976), a Verbal Aptitude Test (analogies; VAT'69 analogieën) and a Numerical Aptitude Test (number series; NAT'70 cijferreeksen). The Drenth series is a widely used intelligence test battery in the Netherlands for estimating academic potential. We chose to administer the analogies and the number series, as being straight measures of fluid (g) intelligence. Both tests require cognitive operations similar to the criterion behavior: processing and reasoning with large amounts of verbal and numerical information.

  The VAT '69 analogies consists of 40 multiple choice items. Every item is an analogy with the first and the last word left out: '……is to B as C is to…..'. Both open places have five independent alternatives, in order to minimize the chance that right answers can be obtained by guessing. The test time for the VAT analogies is 40 minutes. The NAT '70 number series consists of 26 multiple choice items, with five alternatives. Participants have to complete a given series of numbers with the next number. The test time for the NAT is 30 minutes.

  Both the NAT and the VAT are speed and power tests. Apart from a time limit, the items increase in difficulty. The instruction states clearly that it is not necessary to finish all the items; it is just the objective to give as many 'good' answers as possible within the time limit.

  The practical study intelligence questionnaire (PSIQ)

As every practical intelligence questionnaire is domain specific, we constructed a questionnaire especially for first year OUNL psychology students. In constructing the PSIQ we followed guidelines from Sternberg (1993, 1995, 2000). Practical study intelligence was defined as 'The ability to adapt to the demands of the study of Psychology or to adapt the study to one's own abilities in such a way that a maximal study facilitation and study result are realized'.

<p align="center">22</p>

Analogous to the construction of the Tacit Knowledge Inventory for Managers (TKIM), (Wagner & Sternberg, 1991) we located a number of specific potential domain or study problem fields[5]):

- Communication with the OUNL/study counselors/mentors
- Planning (when and how long to study), exam planning etc.
- Processing of subject material (how to handle difficult text sections)
- Exam preparation
- The handling of study tasks
- Profiting from contacts with other students
- Exam strategies
- Curriculum planning extra
- Additional problems

These potential problem fields were presented to 23 senior students of psychology, asking them to generate specific and relevant problems for each problem field. At the same time, they were asked to give possible solutions to their formulated problems, effective as well as ineffective ones. Based on this material we constructed a draft questionnaire of eight problem situations. For each problem situation we created four to seven possible solutions. Because of the difference in numbers of solutions the total number summed op to 41 solutions. The effectiveness of each solution had to be judged on a five-point Likert scale.

The next step consisted of presenting the draft questionnaire to people familiar with first year study problems. Eight student counselors answered the following questions regarding the problem situations and adjacent items:

-Are the study problems relevant for first year students and non ambiguous stated?

-Are the presented solutions realistic?

-Do you have any suggestions concerning the problems and/or solutions?

Based on the feedback given by the study counselors a second version of the PSIQ was constructed. This version consisted of 8 problem situations, covered by 41

---

[5] All materials regarding the construction of the PSIQ, as well as the PSIQ, can be obtained from the author. Address: wim.bloemers@ou.nl

items which had to be scored on a five point scale. The number of items for each situation varied from four to six items. An example:

Situation D.

Very soon you will have to take your exam in Psychology 1B. You already failed twice on this exam and you have no idea if you will succeed this time.
Please score each of the alternatives below on exam preparation on a 1-5 scale.

1	2	3	4	5
very ineffective	ineffective	neutral	effective	very effective

1.	You copy the two practice exams at the study centre and you start practicing these. If you can pass these practice exams you trust that you are well prepared.
2.	You start studying the subject material again in detail, because the last exam contained a lot of detail questions.
3.	You take the two practice exams and you check on which chapters you made the most mistakes. You spend extra time on these chapters.
4.	You get into contact with students who already passed this exam and you ask them about their experiences.
5.	You keep to the study goals as written in the subject material. If you master these study goals, your comprehension of the subject material is adequate.

The last step in the construction was presenting the draft version of the PSIQ to a panel of experts to create norm scores. The draft questionnaire was sent to all 18 academic psychology student counselors of the OUNL, asking them to score the PSIQ. All 18 counselors returned the questionnaire. For keying each item, the (mean) preferential expert score was computed. Then we computed the absolute difference from the mean expert score on all items for all experts. A reliability analysis on these scores resulted in a Cronbach's α of .73, a satisfactory value. The total test score was the sum of all (absolute) deviation scores, a higher score indicating a lower 'practical study intelligence' (the higher a subject's score, the greater the deviation from the mean expert score). The PSIQ scores of our participants were computed in the same way.

Criterion scores

As criterion scores we used the scores on the two first year introductory exams, Psychology 1A and 1B. These two exams consist of 40 multiple-choice questions, each listing four alternatives. The cut-off score is at 27 items correct. The administration of these two exams is computerized, allowing individual choice of exam date and location (any location of an OUNL study centre). The computer randomly generates questions for the exam from an item bank of about 500 items that cover the educational goals of the

course. Immediate results with feedback on errors are given to examinees. However, for the final analyses, no results on reliability of the exam scores are available because every individual exam differs. The SYS (student administration system) only shows number of exam trials, acquired mark and examination date.

Respondents

The NAT and the VAT were administered to 103 first year students of the OUNL. Administration took place after the introductory classes at four OUNL locations. Before administering the two tests, participants were given a brief explanation on the aim of the project. A total of 94 respondents returned data on both instruments and these were used for further analysis. This first sample contained 17 males (18%) and 77 females (82%). The mean age was 35.9 years (range 20-70; sd 9.4). The mean age of the males was 44 years (sd 10.4), mean age of females was 34 years (sd 8.2). Eighty six percent of the participants had at least HAVO (high school).

All 94 participants with data on the two intelligence tests were sent the PSIQ by mail. Of these, 40 returned the PSIQ, a response rate of 37.6%. The definitive sample for analysis on the combined data consisted of 40 participants. From these 40 participants, 29 participated in the 1A exam and 24 in the 1B exam.

Procedure

Administering the NAT and the VAT did not show to be easy. Considering the need of strict objective and fair conditions, it was decided to approach students enlisted at major locations of the OUNL, asking them to score the two tests after having visited an introductory lesson. Test time for the two tests, including an introduction on the project, took 1.5 hours, a considerable investment in time and energy for the participants (this is the main reason why nine participants dropped out). All 94 participants who completed data on both tests were sent their individual scores with a short explanation by mail.

The PSIQ was sent to the 94 remaining participants by mail, including an addressed and stamped return envelope. All respondents who returned the questionnaire (40) were given feedback on their scores by mail.

Results

Table 2.1 lists the descriptives for predictor and criterion measures.

Table 2.1:

Descriptives of the NAT, VAT, Exams 1A and 1B and the PSIQ (shortened version).

Measure	Items	Mean score	sd	α
NAT	26	9.44	4.05	.71
VAT	40	12.43	6.54	.86
PSIQ *	32	26.7	5.44	.71
Exam 1A	40	6.80	1.91	-
Exam 1B	40	6.85	1.75	-

*The PSIQ descriptives refer to the shortened version

Scores on the NAT and VAT equal mean norm scores for Higher Vocational Education (HAVO). The mean PSIQ item score for our participants is less than one scale point from the expert score, reflecting considerable 'practical study intelligence'.

The next step consisted of computing predictor-criterion correlations. Because initial results on the PSIQ showed an α of only .41 we used a shortened version, consisting of 32 items, with an α of .71. The NAT and the VAT showed acceptable α's, respectively .71 and .86.

Table 2.2 lists the correlations of the NAT, the VAT and the PSIQ shortened version results for the two first year examinations. Background variables (sex, age and previous education) were included in the analysis as well, as they could reveal useful information.

As Table 2.2 shows, the score on the VAT seems a consistent and significant predictor for both first year exam results. The score on the NAT is significant only for the first exam result. The score on the PSIQ has no predictive value for the scores on either exam. The small correlation between the PSIQ and the VAT suggests that verbal intelligence could explain to some extent the PSIQ scores. Previous education shows a significant relationship with both intelligence test measures and with the first exam result. The NAT-VAT correlation of .44 is comparable with correlations mentioned in the Test Manual (Drenth & Hoogwerf, 1970).

Table 2.2:

Intercorrelation matrix of NAT, VAT, PSIQ-shortened, exams Psychology 1A and 1B

	NAT	1A result	1B result	Age	Sex	PSIQscore[1]	Previous ed.[2]
VAT	.44** (94)	.41** (65)	.34* (46)	-.13 (94)	-.19 (94)	-.21 (40)	.38** (87)
NAT	--	.33** (65)	.11 (46)	-.02 (94)	-.34** (94)	-.14 (40)	.31** (87)
1A result		--	.66** (46)	-.12 (65)	-.08 (65)	.04 (30)	.38** (65)
1B result			--	-.04 (46)	.04 (46)	-.14 (25)	.17 (41)
Age				--	-.41** (94)	-.21 (40)	-.08 (87)
Sex					--	.24 (40)	-.16 (87)
PSIQ score[1]						--	.14 (37)

*$p<.05$; **$p<.01$; N between brackets; [1]Correlations should be reversed;
[2]Previous education was recoded ranging from 1-7; 1=lowest, 7=highest education.

Subsequent multiple regression analysis (Table 2.3) did not reveal any more surprising information. Leaving out the PSIQ result (n would become too small), we did two multiple regressions with the Exam 1A and 1B results as criterion variables and the VAT and NAT scores, sex, age and previous education as predictor variables. The model specifications are given in Table 2.3. Taking into account the magnitude of some of the correlations, the results from the multiple regression analyses are disappointing, with only a minimal amount of variance explained. For the 1A exam result, none of the predictors had a significant ß value. For the 1B exam result, the VAT listed a significant ß value (B=.125, se=.054, ß=.473, t= 2.332, p<.03). Because both examinations showed very high mean p values (.85 and .88) we did some final analyses tot test the assumption that there could be relevant predictor differences between students having taken one or two exams and those who did not, and between students who actually passed the first and the second exam and those who did not. All exam data where dichotomized in 'taken' versus 'not taken' and 'passed' versus 'not passed' and point-biserials were computed. Table 2.4 lists the outcomes of the correlation analyses.

Table 2.3:

Multiple regression results of exam criterion scores on the NAT, VAT, PSIQ (short.), sex, previous education and age.

Criterion variable: Exam 1A result (n=58)

R	R2	Adj. R2	Se estimate	R2 change	F change	df1	df2	Sig. F
.50	.25	.17	1.76	.25	3.38	5	52	.01

Predictors: (Constant), NAT, VAT, previous education, age, sex

Criterion variable: Exam 1B result (n=41)

R	R2	Adj. R2	Se estimate	R2 change	F change	df1	df2	Sig. F
.41	.17	.05	1.65	.17	1.45	5	35	.23

Predictors: (Constant), NAT, VAT, previous education, age, sex

Table 2.4:

Point-biserial correlations between exams taken/passed versus not taken/not passed and predictor variables (N between brackets)

Predictors	Exam 1A (not) taken	Exam 1A (not) passed[1]	Exams 1A/1B (not) taken	Exams 1A/1B (not) passed[1]
VAT	.18 (94)	.38** (65)	.25* (94)	.30* (46)
NAT	.11 (94)	.46** (65)	.19 (94)	.23 (46)
PSIQ (short.)[2]	-.15 (40)	-.17 (30)	-.18 (40)	-.39 (24)
Sex	-.07 (94)	-.04 (65)	-.15 (94)	.02 (46)
Age	-.16 (94)	-.09 (65)	-.11 (94)	.09 (46)
Previous Education	.02 (87)	.23 (58)	.19 (87)	.02 (41)

*$p< .05$; **$p< .01$ [1]Participants who did not take the exam where coded as missing.
[2]Correlation should be reversed; a higher score on the PSIQ means a greater deviation from the expert score.

Table 2.4 shows that none of the predictor variables had a significant influence on 'taking the first exam' versus 'not taking the first exam'. The two intelligence measures differentiated significantly between students from the first group (taking the first exam) who passed this first exam, versus those who didn't. The verbal intelligence test further differentiates significantly between students who took the first and second exam versus students who did not, and between students who passed these two exams versus those who did not. No other predictors show a significant influence, although the relatively high correlation of the PSIQ score with passing exam 1B is noteworthy.

Composite intelligence measure

Because the object of this study was the creation of efficient and valid predictors for first year study success of psychology students at the OUNL, we did some additional analyses to explore the predictive power of a composite intelligence measure. Administration of this composite measure would have to take no more than 30 minutes, guaranteeing sufficient student participation. Capitalizing on high item-total correlations, a composite 25 item intelligence measure was constructed ($\alpha=.92$). The composite contained 10 items from the NAT ($\alpha= .86$) and 15 from the VAT ($\alpha= .87$). The correlation between these shortened versions was .66, showing some perspective as a potential *g* measure.

Because at the moment of analysis only results on the Psychology 1A exam were available, the correlation between the composite instrument score and the 1A exam was computed. The correlation was .77, with separate correlations for the VAT items component of .48 and for the NAT items of .79. Although the fact that capitalization on coincidental scores may have played a role in the analysis above (the definitive analysis contained only 20 participants), a composite instrument could be an effective predictor, both from the psychometric and the efficiency point of view.

Discussion

The aim of this study was to compare the predictive validity of numerical and verbal intelligence and practical intelligence for first year study results. Practical study intelligence showed no predictive power. Verbal intelligence proved to be the most important predictor for passing the first year introductory exams. A composite, more *g* resembling, intelligence measure indicated that general intelligence is probably an effective predictor for first year study success. This conclusion supports study outcomes on the role of general intelligence on time bounded high complexity tasks (Hunter, 1980) and on the predictive validities of the GRE verbal, analytical and numerical scales (Kuncel, Hezlett & Ones, 2001).

The role of practical intelligence in predicting first year study success remains unclear from our study. As an instrument measuring typical behavior, criterion

correlations within the .20 range should be possible (Barrick & Mount, 1991; Hurtz & Donovan, 2001). The most probable cause for the lack of predictive validity of the PSIQ besides the small N is the fact that the criterion measures used for first year study success are measures of 'maximal' performance. Our PSIQ measures showed considerable SD and reliability, indicating the existence of differences in practical intelligence in our sample. To find out if practical intelligence does play a role in academic achievement, more 'typical' behavior should be included in the criterion scores. Working with longitudinal multiple criterion measures, including more individual study tasks could clarify this issue. Another possibility is that practical intelligence, as a measure of 'typical behavior' correlates with certain aspects of personality, like conscientiousness, emotional stability and autonomy. Personality dimensions show relatively low but stable correlations with work performance criteria (Barrick & Mount, 1991; Hurtz & Donovan, 2001), and initial research on the effects of personality on academic performance seems to yield comparable results (Wolfe & Johnson, 1995; Okun, & Finch, 1998). We did find however a small correlation between the PSIQ scores and the results on the VAT and the NAT, suggesting that verbal and to a lesser extent numerical intelligence may overlap with our measure of Practical Intelligence (see also McDaniel et al., 2001). To clarify the concept of practical intelligence a large scale study should be undertaken, including practical intelligence, g, and personality measures as predictors and multiple study criteria, including maximal and personalized, more typical behavior measures. Such a study could shed some light on the factor structure of practical intelligence and could reveal in what kind of situations it eventually plays a role in the academic curriculum.

Limitations of our study

This study had some severe limitations, most notably the small number of respondents on the various measures. With regard to the construction of the PSIQ, the use of the expert panel can be questioned. Although we followed instructions from Sternberg (1995) in constructing this instrument, the panel consisted solely of academic student counselors. The question is whether this group had enough 'ecological' insight into first year student problems. Another possible distortion could be the fact that both the expert panel, and the students may have scored the PSIQ from a 'socially desirable' point of view. Future

studies on PSIQ keying should also include successful students in the expert panel and the design of the instrument should be aimed at capturing as much 'real' behavior as possible.

Regarding the operationalization of criterion and predictor scores, one critical remark could be that with the use of two standard intelligence tests and *g*-like criterion measures, we capitalized on high *g* predictor-criterion similarity. But this brings forward the question 'what kind of criterion would you like'? There should be (inter)national standards for curricula in Psychology, guaranteeing a professional and competent level of graduates. One necessary element of graduates is general knowledge of the field of psychology. This knowledge is best tested by a sample of objective questions covering the whole domain of psychology. Also from efficiency, objectivity and fairness points of view, multiple choice tests in an academic curriculum seem inevitable, creating for students a *g*-like situation where speed and handling complexity are important determinants.

It is evident that respondent drop out played a major role during the research process. For the final analysis, less than 30 participants of the original group of 104 remained. This phenomenon seriously hampered statistical inference, and may have caused significant restriction of range in predictor and criterion scores. However, closer inspection of predictor scores showed no restriction of range in NAT-VAT scores when comparing the total group with the group who did the Psy 1A exam. Means and standard deviations remain virtually unchanged. There is more reason to suspect restriction of range in the criterion scores. Both the 1A and 1B exams show only a small SD, with about 80% of the scores in the 6-9 range (on a ten point scale). We observed earlier that from the high p-values of our exam data (p.33) it seems likely that students only do exams when they are almost certain to succeed. Many students never reach the level at which they have the subjective certainty to succeed, causing a high drop-out rate and a severe restriction of range in the criterion scores. Encouraging students to take exams and supplying more adequate feedback on study progression could bring more students to take exams. A compact interactive web supported sample of critical study questions, frequently updated, could give students the necessary support in mastering the exam subjects and increasing self confidence, necessary for taking the exams.

Concerning drop-outs a fruitful strategy could be to approach these students individually, and simply ask them what caused them to drop out. This 'open question' approach might reveal causes of study drop-out that remain undetected in statistical research.

Implications for HEIDI

General intelligence or *g,* as measured with the NAT and the VAT proved to be a valid predictor for psychology 1A and 1B exam results in this study. The implication is that to give (potential) students a reliable and useful first year study advice, a general intelligence estimate should be part of such an advice and thus included in the HEIDI website. This *g* measure should be compact and efficient, because students do not seem willing to invest much time and energy in such a measure.

Practical intelligence in this study showed no predictive value for first year study success. It is our opinion that until the concept and the predictive power of practical intelligence are clarified, no advice on this subject should be given to our students.

As stated earlier, at the OUNL the Psychology 1A and 1B exams will be synthesized into one single multiple choice exam, putting more emphasis on *g* as a decisive capacity for study success (complexity of and pressure on the exam situation will increase). So, a serious advice on *g* as an information processing capacity seems warranted. It is advisable to refer to this compact general intelligence estimate solely as a *study subject processing capacity,* as it is not the intention to hamper students in other activities they will undertake when choosing not to start a study of Psychology at the OUNL because of a low *g* score.

CHAPTER 3

The effects of biodata and Big Five Personality characteristics

(Submitted for publication: *Journal of Higher Education*)

3. Effects of biodata and personality characteristics on study success

Up till now, research within the HEIDI project has shown verbal and numerical intelligence to be of considerable importance for study success (chapter 2). The present study focuses on two groups of variables also expected to show significant influence on first year study success: Biographical aspects and Big Five personality variables. Personality variables or 'dimensions', as they are called within the broad Big Five framework, can be seen as the second cornerstone of differential psychology. As stated in the introduction (chapter 1), predictive validities of various Big Five dimensions for job related tasks generally are found to be around .20. Although research on the influence of personality variables on academic success is still limited, the potential of this perspective has been laid out by e.g. Messick (1967) and more recently by De Raad en Schouwenburg (1996). Empirical data on the subject so far seem to be in line with general findings from Industrial and Organizational Psychology (Holder & Wankowski, 1980; Okun & Finch, 1998).

Biographical variables, or biodata, are mainly used in Industrial and Organizational spheres, and show moderate to sometimes high predictive validities for various job categories. Application of biodata for educational purposes has also been scarce and has scored very limited success (Melamed, 1992).

The primary aim of our present study is to estimate the importance of both biographical and personality characteristics for the first and second exam results, psychology 1A and psychology 1B. Relations between personality and biographical aspects will also be discussed.

Before coming to our actual study, we will first elaborate on the theoretical and empirical backgrounds of biodata and personality, with their possible relation to academic performance as a final outlook.

Biodata

It is important to distinguish, within the HEIDI context, between questionnaires with biographical items and personality characteristics. Both questionnaires contain items

34

concerning behavior and experience. But in biographical, or bio-questionnaires, respondents are questioned about verifiable objective behavioral facts, such as age and marriage (Mumford & Owens, 1987). In personality questionnaires individuals are questioned by means of hypothetical situations and interpretations of behavior, trying to get a grip on what is called 'typical behavior' (Hofstee, 2001). Asher (1972) summarizes the above by speaking of 'hard' versus 'soft' items. According to Asher, questions of bio-questionnaires vary on a continuum with extremities denoting the hard and soft ends. The hard end would then denote a 'true' bio-questionnaire and the soft end a 'true' personality questionnaire. Mael's (1991) distinction between objective and subjective items is analogous to Asher's 'hard' and 'soft' distinction.

The distinction above is important when looking at the use of bio- and personality questionnaires, trying to avoid a position where a questionnaire ends up somewhere in between, resulting in a muddy and unidentifiable concept of both soft and hard personality characteristics. Moreover, to discern the predictive validities of bio- and personality questionnaires, the two underlying concepts should show sufficient discriminating validity.

After having established a more definitive grip on the concept of biodata, we will now take a closer look at the history and development of biodata questionnaires.

History and development of bio-questionnaires

Around 1895, within the insurance business, the need for an effective selection system issued in a standardized method to evaluate the life history of applicants for the position of insurance agent. Goldsmith (1922) introduced analysis and weighting procedures for scoring biographical items. He also changed the original application intake sheet to a Weighted Application Blank (WAB). The WAB is constructed on the assumption that 'the best predictor of future behavior is past behavior'. During the following years various articles emerged, concentrating on improved empirical methods for biographical item analysis and improved weighting procedures (Kenagy & Yoakum, 1925; Manson, 1925).

The application of bio-questionnaires was extrapolated to various other occupational fields (Viteles, 1932; Owens, 1976). The recent multiple choice format of

biographical items was developed during World War II, where it showed promising results in the military context (Levine & Zachert, 1951; Parish & Drucker, 1957). After this, successful implementation for personnel selection and governmental functions was only a matter of time. From 1960 on the WAB saw its name changed into the more popular 'biodata' or biographical questionnaire. Up to the present day, biodata questionnaires are founded on the method known as empirical scaling. This means that items are selected and weighted based on their predictive validity towards a desirable criterion (Nickels, 1994).

While the predictive validity of empirically selected and weighted items seemed indisputable (Mumford & Owens, 1987), there was, up till the 1970's, no discussion why biographical items show high predictive validity. But halfway the sixties, empirical scaling came under attack. The main argument was the lack of psychological theory behind 'blind' correlations:

> 'Empirical keys select items solely on the basis of their ability to predict complex criterion performances, item content tends to be complex and lack psychological meaningfulness' (Guion, 1966).

Three alternative methods were developed to capture this lack of psychological meaningfulness in the construction and use of biographical items: rational scaling, factorial scaling and sub grouping.

Rational scaling (or rational keying approach) was introduced to support biographical items with a thorough analysis, resting on a theory to create a solid framework for linking theory, item content and criterion (Kavanagh & York, 1972; Mumford & Owens, 1987; Quaintance, 1981; Williams, 1961).

Factorial scaling resembles factor analysis. Items, expected to have high predictive validity, are a priori selected. By means of factor analysis, these items are grouped, or factorized. So far this method is exploratory, requiring cross validation of the primary factor solution in an independent sample. With a corroborative cross validation, a foundation for theory development and further application is laid (Childs & Klimoski, 1986; Baehr & Williams 1968; Klimoski, 1973; Mitchel & Klimoski, 1982; Eberhardt & Muchinsky, 1982).

Subgrouping, as developed by Owens (1982) is a statistical classification procedure aimed at identifying (groups of) individuals with the same characteristics. Subgrouping is especially used in case of predicting multiple criteria.

The predictive value of biodata

Predictive validities of biodata in the IO Psychology context have been found to be quite impressive, about .30 to .40 (Asher, 1972; Ghiselli, 1973; Hunter & Hunter, 1984; Mumford & Owens, 1987; Owens, 1976; Reilly & Chao, 1982). Reilly and Chao (1982), in an overview, studied the predictive validity of biodata items for administrative, managerial, scientific, and military jobs. For items such as previous education they found predictive validities ranging from .25 to .50 for various criterion measurements such as tenure, training and productivity.

Generalizing the predictive validity of biodata items beyond the IO context however, has not been without problems. Academic success seems merely predicted by high school GPA (Dyer, 1987), previous academic success (Federici & Schuerger, 1974) and personality (Daehnert & Carter, 1987). In all these three studies the predictive power of biographical aspects (e.g. family life, attitudes, values, school related activities, life experience, sex, age, marital status and type of high school) was of minimal or no importance, marking a sharp contrast with the predictive value of biodata in the IO context.

Melamed (1992) conducted a study on the predictive validity of biodata items for academic success, using items from the Standard Admission Form. Results showed biodata items only to have predictive validity for the criterion 'graduating as psychologist' and not for grade point average (GPA). Biodata failed to predict future academic success.

We conclude that biodata show impressive (generalizable) predictive validity in the IO or occupational context. Predictive validity of biodata for academic success has scarcely been studied and the few results on this subject clearly show less impressive validities. A probable cause could be restriction of range, due to the much more homogeneous educational context.

Personality

Any history of personality research will ultimately have to deal with the 'Big Five' as the rock bottom of personality (Barrick & Mount, 1991). From the first systematical inventory of personality descriptors by Allport and Odbert (1936) to the fine and narrow tuning by Cattell (1943) and Eysenck (1947, 1970), the Big Five have been regarded as the more or less definitive sediment of human personality (Digman, 1990). Basically the Big Five is rooted in what is called the lexical hypothesis (Goldberg, 1990; Hendriks, Hofstee & De Raad, 1998). This hypothesis states that all important aspects of human personality are represented in natural language. Just as in research on intelligence, factor analysis provided the way for a more or less definitive structure, resulting in five broad independent 'traits' or dimensions.

The Big Five or 'The Five Factor Model' (FFM) (Digman, 1990; Goldberg, 1990, 1992; De Raad, 1992) consists of the following five dimensions: I. Extraversion, II. Agreeableness, III. Conscientiousness, IV. Emotional stability and V. Openness to experience. If we look at the dimensions of the Big Five, Extraversion and Emotional Stability can be considered personality classics, residing in almost any theory of personality (Matthews & Deary, 1998). Agreeableness and Conscientiousness are relatively new, emanating from the lexical approach. The fifth factor, known as Openness to experience (Costa & McCrae, 1985), Intellect (Digman & Takemoto-Chock, 1981), or (Intellectual) Autonomy (Hendriks, 1997) is still subject of discussion with regard to its definitive interpretation. Regarding the small but significant correlation with intelligence (Goldberg, 1990) it is clear that some aspect of intelligence should be captured in the final denotation of this dimension. Mainly for measurement reasons we chose to interpret this factor as *intellectual autonomy*, based on the recent work of Hendriks (1997; Perugini & Ercolani, 1998). These authors argue that the fifth dimension is best interpreted as Intellectual Autonomy, denoting people with a strong tendency to do things their own (creative) way, as opposed to people only following the thoughts and commands of others. More narrow facets within this dimension are 'critical', 'having an opionion of oneself', 'sharp minded' versus 'obedient' and 'submissive'.

Notwithstanding the broad scientific consensus on the Big Five, critical reviews (Pervin, 1994; Block, 2001) point to two lines of criticism. One line focuses on the Big

Five as a whole, concentrating on the number of factors. The other line concentrates on the exact interpretation of the individual factors, concentrating for the most part on the conscientiousness and intellectual autonomy factors, on whose meaning there is not yet a full consensus.

Recent suggestions for modifications within the Big Five framework have come from Paunonen and Jackson (2000) and Hofstee, De Raad and Goldberg (1991, 1992). Paunonen and Jackson criticize the interpretation of the conscientiousness dimension, arguing that it lacks coherence and that it should be split in three separate dimensions: (1) a dimension emphasizing methodical and orderly behavior; (2) a dimension emphasizing trustworthy and reliable behavior and (3) a dimension emphasizing ambitiousness and drivenness. Loevinger (1994) also criticizes the conscientiousness factor in that it lacks behavioral aspects of moral nature, especially integrity.

Hofstee, De Raad and Goldberg (1991), using the lexical approach, suggest a more rigorous restructuring of the Big Five. They showed that everyday personality descriptives usually load significantly on two of the Big Five dimensions. Their approach resulted in the Abridged Big Five Dimensional Circumplex Model (AB5C-model; Hofstee & De Raad, 1991). This model, with 90 facets of behavior, leads to a 'periodic system' of personality, with every facet based on the combination of two Big Five factors (for a detailed description of this model, see De Raad, 1997; De Raad & Schouwenburg, 1996).

Criticism on the number of factors has come from Paunonen and Jackson (2000), Tellegen (1993) and Carroll (2001). Paunonen and Jackson argue that the conventional Big Five dimensions still leave variance unexplained. Building on a study by Goldberg and Saucier (1998) they offer an extra nine bipolar dimensions of personality, not, or not adequately implied within the Big Five. Among their dimensions are religiousness, manupulativeness, honesty and humor. Tellegen (1993) offers some empirical arguments for a Big Seven, pointing to the importance of two extra dimensions: negative and positive valence. Digman (1997) and Carroll (2001) both using higher order factoranalytic methods present two higher order factors, respectively socialization and personal growth, and control and adaptability.

The two lines of criticism in the end lead to the same destination: estimating the exact number of relevant dimensions for describing and explaining human personality. Higher order factoranalytic studies can shed some light on the overall structure of human personality, concentrating on basic drives. Studies aimed at refinement and extension of the Big Five factors can point to important behavioral aspects not (fully) covered by the basic five dimensions. These studies are useful for practical purposes, were there is a need for more (situational) specific behavioral predictions such as in work and educational settings. In the next paragraph, we explicitly focus on the role of the Big Five for achieving academic study results.

The predictive validity of personality dimensions for academic success
Research shows only limited success for predicting academic achievement from personality dimensions (De Raad & Schouwenburg, 1996). The most likely personality dimension for predicting study success seems conscientiousness, although dependent on the specific content of the criterion measure(s) other dimensions could equally well be predictive of 'success'. Okun and Finch (1998) found a relatively strong (negative) correlation between conscientiousness and study drop-out. Busato (1998) also points to the importance of conscientiousness for (long term) study success in psychology. Results on other dimensions are mixed, probably due to sampling error (Van Dam, 1995; Dyer, 1987; Blicke, 1996; Busato, 1998).

According to Van Dam (1995) the Big Five dimensions are 'too general' to predict specific behavior in a specific situation. She advises the construction of narrower trait based questionnaires, specifically designed to meet the needs of a specific situation. Although Van Dam seems to give a concrete and useful advice, her argument on situational specificity is rebutted by research on the predictive value of personality dimensions in the IO context (Barrick & Mount, 1991). Research in the IO context has generated quite stable predictive validity coefficients for personality dimensions, setting conscientiousness on top and emotional stability second. Predictive validities of other dimensions show some task or job related specificity, for example agreeableness for commercial jobs. The question on situation specificity regarding effects of personality variables on academic achievement cannot fully be answered until large scale meta-

analytic studies will shed more light on the *SD* of the operational validities of studies included in the meta-analyses.

The relation between biographical aspects and personality

As stated earlier, biodata and personality descriptors can be seen as two opposing poles of the continuum covering personal characteristics. Biodata represent the 'hard' objective facts; personality descriptives represent the 'soft' interpretative side. The aim of both poles however is the description and future prediction of human behavior by means of reliable and valid characteristics and behavioral traits. Attempts like those of Thenophyr (1994), trying to integrate the soft and hard side of behavior into a hierarchical model can be useful from the viewpoint of theory construction but from our viewpoint this integrative approach is more mystifying than clarifying.

Mount and others (2000) studied the incremental predictive validity of biodata together with General mental ability and the Big Five. They too presupposed an overlap between personality and biodata, which was confirmed by their study. But they conclude that it is not the case that the use of both biodata and personality questionnaires leads to redundancy. For a more precise prediction of future behavior, Mount, Witt and Barrick (2000) suggest the construction of situation tailored instruments.

From the discussion on biodata and personality variables it can be concluded that it is important to keep the two spheres well apart, thus marking a clear distinction between 'objective' and 'interpretative' facts of personality. Questionnaires should try to capture as much criterion related behavior of respondents as possible.

Hypotheses

For our study, the central question is twofold:
- What is the relation between the first and second exam result and biographic, respectively, Big Five factors?
- What is the relation between biographic and Big Five factors?

Based on empirical results so far, we expect conscientiousness to be an important predictor of exam results. Emotional stability is also expected to have a positive effect.

With regard to biographical factors we expect previous education and living conditions to have a significant influence on the first and second exam result.

Method

Instruments

The biodata questionnaire (BQ)[6]

We based the construction of our BQ on the standard enrollment form for first year psychology students. Nearly all the items on this form met the restrictions for 'hard' biographical items: factual, specific and verifiable. To obtain an adequate coverage of all the possible important criterion related characteristics of an individual we created eight specific fields of biodata characteristics: personal descriptors (hard), life conditions (hard), social status (hard/soft), financial situation (hard), reasons for studying psychology at the OUNL (soft), friction regarding study activities (soft), social support (hard/soft) and study expectancies (soft). With various degrees of hard/softness we hoped to gain more insight in the predictive nature of both concepts for study success.

In line with Asher's 1972 advice on bio-item constructing, a primary version of the BQ was constructed, containing 65 items. This draft version was pretested with 100 psychology students to find out if items differentiated among students, if they were clearly formulated and unambiguous. Analysis of this pilot study resulted in removing 12 items because of lack of responsiveness and score variance. Eleven questions were reformulated. Six questions appeared to be redundant and were also removed. Following suggestions of respondents, eight items were added to the field 'Reasons for studying psychology at the OUNL'. The response options for one question were extended to eight alternatives. After correcting the primary version of the BQ, the definitive version contained eight scales with 55 items: personal descriptors (7 items), life conditions (6 items), social status (2 items), financial situation (2 items), reasons for studying psychology at the OUNL (28 items), friction regarding study activities (5 items), social

---

[6] All materials concerning the construction and administration of the BQ as well as the BQ itself can be obtained from the author: wim.bloemers@ou.nl

support (2 items) and study expectancies (3 items). Completion of the BQ took about 25 minutes.

Personality Questionnaire FFPI

For the inventory of personality aspects we used the Five Factor Personality Inventory (FFPI), (Hendriks, 1997; Hendriks, Hofstee & De Raad, 1998). The FFPI measures in an efficient way the position of a subject on all five Big Five dimensions. It consists of 100 items, 20 per dimension. The FFPI can both be used for self- and other assessment. In our study we used the self assessment version. Completing the FFPI takes about 15 minutes. The FFPI is a unique personality questionnaire because of its scoring principles (Hofstee, Ten Berge & Hendriks, 1998). Normally, scale scores on a personality questionnaire are calculated by summing the respective item scores. But this method leads to high intercorrelations among scales. Moreover, the classical method is based on an arbitrary 'neutral' point, where the population mean of a scale is set at zero, whereas in reality this point can be well above or below zero. For example the mean population score on agreeableness might be 3 on a scale from 0 to 4, indicating that the mean population score represents more friendly than unfriendly behavior. By setting this mean scale score of 3 to zero (as it is the population mean), a subject scoring zero on this questionnaire gets a 'neutral' score on agreeableness. However, in reality this score means 'showing more friendly than unfriendly behavior', although compared to the population mean this is average behavior. With the FFPI, scale scores are computed using anchored factor scores (Hofstee & Hendriks, 1998), based on a representative sample of the Dutch population (N=2494). This allows giving participants information on their position on a scale compared to the population mean, as well as on this mean itself. All the data on the FFPI were transformed to anchored factor scores with the computer program FFPI.exe (Hendriks, 1998). The total variance explained by the five FFPI factors is about 40%, with all factors showing acceptable reliability independence and stability (Hendriks, Hofstee & De Raad, ibid.)

Criterion scores

For the criterion scores, the result on the exams Psychology 1A and 1B, data were collected respectively 5 and 18[7] months after the start of each course. Five months seemed long enough for our participants to prepare for and complete the first exam. At the OUNL there is no fixed exam date; students are free to determine their own exam moment. The exams consist of 40 multiple choice questions on a broad introduction to Psychology. Although no reliability estimates of the exam scores are available, the scores are assumed to have good reliability, as the items are continually upgraded, based on item analysis. Apart from the 1A and 1B exam results, we decided to create a second dichotomous criterion, labeled 'taken first (and second) exam versus no exam'[8]. This was done because distinguishing between these two groups could contain valuable information about which predictor variables could differentiate between students taking and not taking exams. As seen below, only about half of our participants had taken the first exam at the time of our analyses, thereby also implying the possibility of a severe restriction of range.

At the moment of data collection for each respective course (January 2001/January 2003), 108 respondents had completed the first exam, psychology 1A (55%) and 86 (44%) the 1B exam. Success percentages were 82.4 % (n=89) for the first exam and 92% (n=79) for the second exam.

Participants

A random sample was taken from the first year students of psychology at the OUNL. A total of 370 students were sent both the FFPI and the BQ. A total of 195 participants returned both the FFPI and the BQ. Females accounted for 81% (n=158), males for 19% (n=37) of the respondent group. The mean age was 35 years (SD= 10.4, with a range of 52. The mean age of the males was 40 years (SD= 11.7), of the females 34 years (sd=9.8), with a range of 46. About 37% of the respondents (n=72) had a previous

---

[7] Due to practical problems, the initial analysis included only the 1A exam data. Data concerning the 1B exam were collected at a much later moment and may therefore differ from other study results.
[8] Participants who did not take exams were excluded from analyses concerning the first and second exam result.

education at or above high school level (HBO/WO). More than two third (70%) were part of the working population.

Procedure

Both the FFPI and the BQ were sent by mail to 370 students at the beginning of the semester. A letter, explaining the procedure and the aim of the study as well as a stamped envelope for return were included. Both questionnaires were returned by 195 participants. Five returned FFPI questionnaires turned out to be unusable for further analysis, having too many items unanswered; so 190 participants remained for further analysis. After completing the data analysis, all participants were sent a summary of the study results.

Results

Overall results showed that the 'harder' biodata showed greater predictive validity. 'Previous education', a 'hard' variable was the strongest predictor. The 'amount of spent study time' (hard) also had predictive validity, together with 'the need for flexible studying' (soft). Students who especially valued possible economic and societal advantages from studying psychology scored slightly lower on the 1A exam. Those who were highly attracted by the 'free entrance' system of the OUNL scored slightly lower on both exams. Students who studied more out of recreational interests scored slightly higher on both exams, and those who valued the flexibility of the study scored relatively higher on the 1A exam.

Apart from one significant ß value for Agreeableness on the criterion 'taken exam versus no exam', Big Five dimensions showed no predictive validity for exam results in our study. Relating biodata and Big Five variables, women scored higher on Agreeableness and on Extraversion. A negative correlation was found between age and extraversion. In the next section we will give a more detailed analysis of our BQ and FFPI results.

Biodata and study success

Our BQ results were disappointing in a sense that only few of the scales showed acceptable reliability. First, we computed reliabilities (Cronbach's α) for the interval BQ scales. Reliabilities varied between .73 and .18 with a mean of .42. Based on the reliability figures, we decided to focus on the 'reasons for studying at the OUNL' scale. We factoranalyzed the items of this scale. Two meaningful factors emerged from the varimax analysis. The first factor (4 items; α=.74) was labeled 'Economic/societal advantage of studying Psychology' pointing to economic and status like profits from studying psychology. The second factor was labeled 'Easy access of study at the OUNL' (2 items, α=.75).

Table 3.1 lists the correlations of these two scales, as well as other relevant isolated BQ items (only significant results shown) with the criterion score first exam result.

Table 3.1:

Significant correlations of bio scales and items with exam results Psychology 1A and 1B.

Items/scales	M	SD	r_{1A}	r_{1B}	$N_{1A/1B}$
Scale 1: Economic/societal advantage	11.4	3.41	-.20*	-.12	104/86
Scale 2: Easy access (no entrance level required)	5.98	2.39	-.28**	-.27*	106/86
Highest form of previous education (item 10)	5.43	1.75	.48**	.39**	105/86
Being able to read and understand a difficult text in English (item 53)[1]	3.76	1.04	.24*	.17	108/86
Study as a form of recreational activity (item 17)	3.78	.87	.22*	.31**	108/86
Distance learning facilitates flexible studying (contents, tempo and examinations) (item 30)	4.49	.56	.33**	.18	108/86

*p <.05 **p<.01; r_{1A}= correlation with result exam 1A; r_{1B}= correlation with result exam 1B;

$N_{1A/1B}$= number of participants 1A and 1B exams

[1]Although this item had a significant correlation with the 1A exam result we did not explore this item because from September 2003 the introductory text for the 1A and 1B exam will be in Dutch.

From table 3.1, 'easy access', 'previous education' and 'Study as a form of recreational activity' seem the most consistent predictors for passing the 1A and 1B exams. It is noteworthy that both constructed scales have negative correlations with both criterion scores. That means that students who value status and economic profit as important study goals, as well as students who are attracted by the OUNL because of easy access, score relatively lower on the first and second exam. Students however who seem to value the study of psychology more or less as a recreational activity and who are attracted by the high flexibility of studying at the OUNL scored relatively higher on the 1A and 1B exam.

Table 3.2:

Significant correlations of bio items with the dichotomous criteria 'exam 1A' and '1A/1B passed or not'.

Items	M	SD	r_{1A}	$r_{1A/1B}$	N
How many hours do you spend weekly on your psychology study? (item 44)	12.35	5.50	.18*	.24**	178
I would like to know more about a course subject (item 21)	3.91	.85	.18*	.09	193
How many courses do you think on *average* to complete on a yearly basis? (1 course = 100 hours of study, the average is six courses on a yearly basis)? (item 38)	5.50	1.56	.16*	.17*	179

$*p < .05$ $**p < .01$; r_{1A} = correlation with 1A passed; $r_{1A/1B}$ = correlation with 1A and 1B exam passed

From table 3.2 it can be read that the BQ items 'Hours spent weekly on studying', and 'number of courses intended to complete' show low to moderate but significant correlations with the criterion variables. The BQ item 'I would like to know more about a subject' correlated only significant with the criterion '1A exam passed'.

We performed two stepwise multiple regressions (Table 3.3 and Table 3.4), to get a clearer picture of the proportion of variance in the (continuous) exam scores explained by relevant BQ predictors. Table 3.3 shows that 'previous education' generates significant ß values for both exams, whereas 'distance learning facilitates flexible studying' generates a significant ß value only for the first exam. Table 3.4 shows that only the biovariable 'amount of spent study time' explained some significant variance in our two exam measures. The explained variance in the two exam measures was however only a few percent.

Table 3.3:

Regression analysis of criteria 'exam result Psychology 1A and 1AB' on BQ variables

Items	M	SD	B_{1A}	β_{1A}	B_{1AB}	β_{1AB}
Highest level of previous education (item 10)	5.51	1.72	.29	.37**	.28	.37**
Distance learning facilitates flexible studying (item 30)	4.53	.57	.84	.30*	.39	.17
Constant			3.3		3.9	
$\underline{R}_{1A} = .46$, $\underline{R}^2_{1A} = .21$, adjusted $\underline{R}^2 = .20$, $\underline{SE} = 1.19$, $\underline{n} = 111$, $\underline{F} = 14.7$, $\underline{p} = .00$						
$\underline{R}_{1B} = .43$, $\underline{R}^2_{1B} = .18$, adjusted $\underline{R}^2 = .16$, $\underline{SE} = 1.21$, $\underline{n} = 83$, $\underline{F} = 8.8$, $\underline{p} = .00$						

Only significant ß values are listed. *p<.05; **p<.01

Table 3.4:

Regression analyses of criteria 'Passed exam 1A and 1A/1B versus not passed' on BQ variables

Items	M	SD	B_{1A}	β_{1A}	B_{1AB}	β_{1AB}
How many hours do you spend weekly on your psychology study? (item 44)	11.7	5.03	.02	.19*	.02	.25**
Constant			.50		-.36	
$\underline{R}_{1A} = .22$, $\underline{R}_{1A}^2 = .05$, adjusted $\underline{R}_{1A}^2 = .02$, $\underline{SE}_{1A} = .49$, $\underline{n} = 190$, $\underline{F} = 1.80$						
$\underline{R}_{1AB} = .26$, $\underline{R}_{1AB}^2 = .07$, adjusted $\underline{R}_{1AB}^2 = .04$, $\underline{SE}_{1AB} = .48$, $\underline{n} = 190$, $\underline{F} = 2.7$, p<.02						

Only significant ß values are listed. *p<.05; **p<.01

It is noteworthy that the figures from the regression analyses contradict some of the correlational outcomes, in particular those of the two constructed scales. While both scales show a moderate negative correlation with the 1A and/or 1B exam, the multiple regression analysis shows that the two scales do not explain significant variance in the exam outcomes. The two exam results correlated .61 (p<.01); having passed the first exam correlated .74 with having passed the second exam (r point biserial p<.01), while the result on the 1A exam explained 37% of the variance in the second exam result (R=.61, se=1.08, F=48.6, p=.00, B=.78, se=.11, ß=.61, p<.01, C=1.3, se=.85).

Personality dimensions and study success

Table 3.5 lists the correlations between the Big Five personality dimensions and the exam results. As Table 3.5 shows, no significant predictor-criterion correlations were found. The slightly negative correlations between Emotional stability and exam result were unexpected, but they seem worth further investigating. They could fit a pattern were insecure students try to maximize on their chances of success on the two exams. We found however no significant correlation between Emotional Stability and spent study time. Among the Big Five dimensions we found a few significant intercorrelations. Autonomy correlated negative with Conscientiousness ($\underline{r}$= -.18, $\underline{p}$ <.05). Emotional stability correlated with Extraversion (r = .16, $\underline{p}$ <.05). Whereas FFPI factor scores should show zero intercorrelations (Hendriks et al, 1998) our intercorrelations must be attributed to sample characteristics.

Regression analysis of the exam 1A and 1B results on the FFPI dimensions showed no significant predictive validity for any of the Big Five factors, with $\underline{R}^2$ of .03 for the 1A exam result and 0.4 for the 1B exam result.

Table 3.5:

Correlations between personality dimensions and exam results (n=186)

	1	2	3	4	5
1 Extraversion	-				
2 Agreeableness	-.01	-			
3 Conscientiousness	-.03	-.03	-		
4 Emotional Stability	.16*	-.02	.04	-	
5 Autonomy	-.06	.11	-.18*	-.14	-
Exam 1A result (n=111)	-.06	-.03	-.01	-.17	-.03
Exam 1B result (n=84)	.00	.00	-.07	-.13	-.01

*p≤.05 (two tailed)

With respect to our second criterion 'succeeded on first (and second) exam' versus 'not succeeded' one FFPI dimension generated one significant correlation: Agreeableness ($\underline{r}$=-.14, $\underline{p}$< .05) with the 1A exam passed. This small negative correlation was not substantiated in a regression analysis (ß=-.13, p=.07). The total explained variance of the multiple regressions of 'succeeded on exam versus not succeeded' on the Big Five dimensions was negligible for both exams.

A closer look at the FFPI dimensions

Realizing our relatively small sample size, it was easy to assume that the lack of predictive value of the Big Five dimensions was attributable to sampling c.q. restriction of range. Because the number of respondents in our study was much smaller than that of the norm group (184 versus 2494) it was not surprising that ranges and means of some of the Big Five factors in our studies differed from those of the norm group. Our study sample scored significantly higher on Extraversion and Autonomy, but lower on Conscientiousness than the norm group. This finding, however, is no hard proof for a significant difference in distribution.

To find out if our sample Big Five dimensions factor scores distributions were actually different from the population Big Five dimensions distributions we tested for each Big Five dimension the difference between the standard deviation found in our sample and that of the norm group. This difference was calculated as follows: the standard error of measurement of a standard deviation is by definition $SD/\sqrt{2N}$ (For the norm group this value is a constant: $1/\sqrt{2N}$). The standard error of the difference between two standard deviations, in our study the standard deviation of the norm group Big Five dimensions and the study sample Big Five dimensions, $Se_{(SDng-SDx)}$, equals $\sqrt{(SD^2_{ng}/2N + SD^2_{x}/2N)}$. The last step consisted of dividing the difference of the norm group and study sample standard deviations of the Big Five dimensions by the standard error of difference: $(SD_{ng}-SD_{x})/ Se_{(SDng-SDx)}$. The outcomes for each Big Five dimension were tested against the corresponding z-values of the normal distribution (two-tailed). The results of this last step are given in Table 3.6. As can be seen in Table 3.6, the differences in standard deviation are not significant (two-tailed), implying that the Big Five profile of our study sample did not differ from the norm group Big Five profile.

Table 3.6:

Significance of differences between sample and norm group Big Five dimension standard deviations

Big Five dimension	$(SD_{ng}-SD_x)/Se_{(SDng-SDx)}$	p-value
Extraversion	.77	.22
Agreeableness	-.15	.44
Conscientiousness	.38	.35
Emotional Stability	.54	.29
Autonomy	-.15	.44

Big Five personality dimensions and biodata

Correlations between Big Five dimensions and main (hard) biodata are listed in Table 3.7.

Table 3.7:

Correlations between main biodata variables and Big Five personality dimensions

		E	A	C	S	At
Sex	(n=184)	.18*	.15*	.01	-.13	-.14*
Age	(n=182)	-.15*	-.07	.06	-.06	.13
Previous educ.	(n=180)	-.04	-.12	.04	-.01	.01

* $p < .05$; E= Extraversion; A=Agreeableness; C=Conscientiousness; S= Emotional Stability; At= Autonomy

In our study, the women scored more extraverted and more agreeable than men. A negative correlation was found between age and extraversion. Our study failed to replicate well known correlations between Conscientiousness and age (Hendriks, 1998), Autonomy and previous education, and sex and Emotional Stability.

Correlational analysis of the other BQ variables with Big Five personality dimensions showed only scattered significant results, e.g. Agreeableness and 'helping motives' ($r=.36$, $p<.01$). Agreeableness also showed a significant correlation with the wish to advance in a scientific domain ($r=.26$, $p<.01$).

Discussion

In this study we tried to establish the predictive validities of biodata and Big Five personality dimensions for the first year exam results of psychology students at the OUNL. Our aim was to find out if these two groups of variables could and should be used for inclusion in a web based instrument (HEIDI) for giving potential students advice on their chances for study success. We will now reflect on the study outcomes by instrument, starting with biodata. We will conclude with the implications of this study for possible inclusion of our instruments in a web based diagnostic inventory.

Biographical variables and first year study success

The BQ used for this study delivered some strong and some surprising predictors for the score on the first and second exam result in Psychology. Important biodata for the first and second exam result were level of previous education, as we hypothesized, and giving high priority to flexibility of study and studying as a form of recreational activity. The two scales 'Easy access' and 'Economic societal advantage' also showed small but significant correlations with exam results. Our hypothesis regarding living conditions was not confirmed. Most biodata variables of our constructed BQ showed no relation with exam scores or with either or not taking (and passing) the exams.

Upon comparing the students who passed the first and the second exam with students who did not, the number of hours spent on studying, and intended number of courses on a yearly basis show small but significant relations with succeeding on exams. The extremely high success ratio on both exams explains why previous education does not differentiate between passing and failing an exam. Previous education is, however, strongly related with the result on exams. Turning from correlational data to regression analysis, the role of the two constructed scales 'economical societal' and 'easy access' decreased, showing that these motives are of minor importance for explaining criterion variance.

The overall predictive result of biographical variables for first year study success is somewhat disappointing, taking into account the number of variables used. It shows

52

that first year success of an academic career is hard to predict, even from a minute biographical perspective.

Having no or only little experience with previous education appears to be a serious handicap for our students. This is the more dissatisfying because the OUNL has a strong mission for attracting people with 'missed educational opportunities'. Our study shows that chances for graduating for these people are small, even considering the fact that they have unlimited time of study. However, the percentage of this group is small; in our sample 15% of the participants had an education below high school level. In general, results on the BQ show that a high level of previous education, a strong need for flexibility in study conditions and a serious investment in study time are indicative of successful first year study results and probably for graduating, as Busato, (1998) has shown.

Big Five personality dimensions and first year study success

Based on the results of other studies (De Raad & Schouwenburg, 1996; Okun & Finch, 1998; Barrick & Mount, 1991; Van Dam, 1995; Busato, 1998) we expected Conscientiousness and Autonomy, and to a lesser degree Emotional Stability and Extraversion to have some predictive validity for the first year exam results. But personality dimensions, as assessed with the FFPI did not manifest predictive validity for first year exam results in this study. Autonomy, as defined by Hendriks (1998), and Extraversion do not seem plausible predictors for multiple choice exam results. In the case of Autonomy the small but significant correlation with intelligence could lead to a positive correlation using more aggregated criterion behavior (Ackerman & Heggestad, 1997).

The lack of predictive value of the Big Five dimensions could only partially be attributed to restriction of range. Our predictor scores did not show any significant restriction of range, but it is likely that some restriction of range occurred in the criterion scores. Only about 50% of our subject completed the first exam, and about 80% scored over a 6 on a ten-point scale. For the second exam, this figure is even higher, thus limiting possible predictor-criterion correlations.

Splitting our sample, based on having or not having done exams, surprisingly showed Agreeableness and Emotional Stability to have a potential *negative* relationship with succeeding on the first exam. Although explanations for this strange phenomenon can be made up easily (e.g. people being too agreeable have difficulty in asserting study behavior; people low on emotional stability lack self assuredness and therefore they study harder and longer) these results only get significance when cross-validated in a larger sample.

Another possible explanation for the Big Five dimensions failing to discriminate between exam results and students succeeding on exams vs. students not succeeding could be the fact that personality measures need more time and more aggregated criterion behavior to be able to make a difference. We used only two simple multiple choice criterion measures of 'maximal' nature, whereas our predictor was an instrument assessing typical behavior. The fact that the Big Five dimensions do not predict first year study success is probably due to the phenomenon that these general measures of typical behavior should be tested against more longitudinal, varied measures of study success (Spence & Helmreich, 1983; Helmreich, Sawin & Carsrud (1986).

A final note on the lack of predictive validity of the FFPI comes from authors who assert that the Big Five dimensions are too broad to predict specific behavior, leaving much of the relevant criterion variance unexplained (Mershon & Gorsuch, 1988; Rothstein, Paunonen, Rush & King, 1994; Van Dam, 1995; Paunonen, 1998; Paunonen & Ashton, 2001). These authors argue that for predicting specific behavior the use of specific and narrow predictors is needed. The use of a special 'achievement goal' scale for academic settings could generate more positive results than focusing on a general dimension of conscientiousness (Hough, 1992; Harackiewicz, Barron, Tauer & Elliot, 2002). The fact that biographical variables like 'economic and societal perspective' and 'intellectual development' showed some predictive power coincides with the use of a specific education directed 'achievement goal' scale.

Limitations of the present study

Sample

Women were strongly overrepresented among our subjects, making up about 80% of our sample. Apart from the fact that Psychology at the OUNL attracts women in large volumes, this female overrepresentation could have a significant influence on mean FFPI dimension scores and correlations. But as stated above, our Big Five score distributions did not differ significantly from the norm group distributions. Looking at the population scores from Hendriks' 1998 study, the females in our study showed a more or less 'male profile'. It is obvious that the study of psychology at the OUNL is highly favored by females, but the specific distribution of personality characteristics among this group, as well as the implications for the prediction of study results should be the object of further study.

A second limitation is the severely restricted group taking exams. The initial sample was reduced by 45% on the first exam. It is therefore the question whether our results are generalizable to the population of first year psychology students. The fact is that the OUNL has no fixed exam moment, as students are free to set their own exam moment. Considering the drop-out rate on the first exam, from a study proficiency point of view, this seems a bad policy. It is well known from research on goalsetting theory (Locke & Latham, 1990) that at least some pressure is required to get results. Introducing fixed exam dates and creating a more stronger commitment by means of close contact with a study counselor could facilitate study progress.

Restriction of range

Possible restriction of range on the FFPI scores and on the criterion exam scores has already been dealt with in the results section. Restriction of range occurred in the criterion scores. Only a very limited group of students took both exams and about 80 percent of these students passed. The main task seems to be getting more students to take exams.

Criterion measurement

We used scores on the first two multiple choice exams, measures of maximal performance. This might have seriously hampered the possible predictive power of measures of typical performance, like the FFPI and the more 'softer' items of the BQ. But as Busato (1998) has shown, the result on the first exam is probably decisive for long term study success. Given the extremely high correlations we found between the exam results 1A and 1B, and between passing both exams, Busato's conclusion about the importance of the first exam result seems justified. From our study it can be further concluded that initial exam results are best predicted by hard, 'maximal' predictors, like previous education and the amount of spent study time. The statistical power of more soft and 'typical' measures is probably only revealed in longitudinal, aggregated multiple criterion measures also including measures of more 'typical' behavior. As 'typical' behavior may be predictive of a successful and professional attitude, 'maximal' behavior is decisive in being able to cope in the first place with the complex and substantive nature of the academic study of psychology.

Implications for HEIDI

From our study it appears that for successfully passing the first year of psychology at the OUNL a previous education of high school level is preferable and this advice should be given from the HEIDI website. As no formal level of previous education is required for enrolling at the OUNL, students with no clear or no sufficient level of previous education should be able to obtain an indication of the level of their information processing capacities from the HEIDI intelligence test as discussed in chapter two.

Furthermore, BQ questions concerning attraction to the flexible study system of the OUNL and taking sufficient study time should be included as well.

Personality measures showed no direct influence on first year study results and should therefore not be included in an advice concerning first year study success. From

the viewpoint of a democratic university and professional diversity, this is obviously a positive point. If however longitudinal research with multiple criteria shows that a certain personality profile or profiles are beneficial from a professional functioning point of view, we would advise to include a personality measurement in the HEIDI website as well. This would enable our students to make a vocational choice within the field of psychology where their chances of study success, and thus of personal and societal wellbeing are maximized, being the ultimate goal of HEIDI.

CHAPTER 4

The effects of prior knowledge and study expectations

4. Effects of prior knowledge and study expectations

In the previous chapters, we looked at the predictive power of (practical) intelligence, personality and biodata for first year study results. Results so far show a distinct influence of general intelligence (g), previous education and spent study time on first year study success, with other predictors showing no or only minor predictive values.

One of the HEIDI factors still to be investigated and subject of this study is the role of prior knowledge, combined with study expectations. Because of the high drop-out rate, it might be assumed that enrolling students make their study choice based on a more or less 'romantic' image of academic psychology, thereby lacking the necessary (scientific) knowledge and attitude to succeed. This latter phenomenon could also be labeled 'negative knowledge', that is, assumptions and knowledge about psychology, which, when becoming operational within an academic environment, are handicapping rather than facilitating. Examples of this 'negative knowledge' are expectations about self and other knowledge, issuing in thoughts about being better able to solve one's own and others' life problems.

If first year students with less prior knowledge of their chosen field of academic study score significantly lower on examinations, and if their drop-out rate is significantly higher than that of students showing more relevant prior knowledge, this finding is of importance both for enrolling students and the faculty in question. Providing potential students with adequate feedback on this issue could prevent them from confronting a major disappointment in their study career. Moreover, the success ratio of the faculty could profit in a substantial way from such a realistic study advice for enrolling students, based of course on the assumption that students with a more or less negative advice do indeed have a lesser chance of study success, and that they do not enroll.

The central question of this study is: 'what (scientific) knowledge do enrolling psychology students possess of the academic field of psychology, what are their expectations of the study and how do these two aspects influence study success'?

This study should result in the construction of a test of prior knowledge of the academic field of psychology, showing at least some predictive validity for first year study success and suitable for being administered by means of an internet appliance.

Prior knowledge

What precisely is prior knowledge? With regard to the literature on this subject, a fairly diverse picture enrolls. A structural distinction often made is that between declarative and procedural prior knowledge (McCloy, Campbell & Cudeck, 1994; Posner, 1978; De Jong, 1986). Declarative knowledge is about 'knowing what is the case' for example 'what courses do I get in my first year of study', or 'who is the founding father of behaviorism?' Procedural knowledge is about knowing 'how' for example 'knowing how to be successful in the study of psychology' or 'knowing how to prepare an exam in an efficient and effective way'. The distinction made above is further articulated by introducing episodic and semantic prior knowledge (Cohen, 1983), relating to specific content aspects of prior knowledge.

Glaser (1984) emphasized a more quantitative distinction, pointing to the difference between domain specific and domain transcending prior knowledge. A domain is a relatively 'complete' part of some specific field of (prior) knowledge, like academic psychology, or mechanics. Dochy (1993), trying to integrate both the declarative and procedural aspects of prior knowledge into a more comprehensive dynamic definition, concluded that:

> 'Pre Knowledge makes up the total amount of knowledge, present within an individual for executing specific learning tasks, structured in (cognitive) schemata, declarative and procedural, part implicit part explicit and consisting of content specific and meta-cognitive knowledge of a dynamic character' (p.100).

This rather elliptical definition does in fact no more than putting together all the various aspects of prior knowledge into one grand complex amalgam. In our study, the concept of prior knowledge will be truncated to one simple aspect. More precisely we are only interested in the question 'What and how much knowledge do enrolling first year psychology students of the OUNL have of the academic field of psychology, and what is the relation of this knowledge with first year study success?' Thus, the question regarding the nature of prior knowledge is treated from a simple declarative point of view.

Effects of prior knowledge on study success

Ausubel, one of the founding fathers of modern educational psychology, writes:

> 'The most important single factor influencing learning is what the learner already knows. Ascertain this and teach him accordingly' (Ausubel, 1986, p. 267-272).

According to Anderson and Pichert:

> 'The knowledge a person possesses has a potential influence on what he or she will learn and remember…' (1978, p. 440-443).

The above authors seem to assume an a-priori positive influence of prior knowledge on study results. This influence seems most plausible for the initial phase of an academic study. Adequate prior knowledge may have a 'boosting effect', that is, it may help to facilitate the first encounter of a student with a given field of study. Students with more adequate prior knowledge will have to spend less energy on getting familiar with the concepts and modes of thinking they already (partially) possess. Prior knowledge may also boost confidence because students with adequate prior knowledge receive positive feedback on having relevant knowledge of a field of study. Considering long term effects, when the complexity of the study increases, it seems plausible that study motivation together with discipline, time devoted to study and general cognitive capacity for processing knowledge, also known as g, become more important. To our knowledge, long term effects of prior knowledge on study success have not been investigated.

Considering the role of prior knowledge, two basic assumptions seem to dominate cognitive psychology:

1) The acquisition of (new) knowledge and skills is dependent on what is already present at the level of entry (Neisser, 1976; Dochy & Van Luyk, 1987).

2) Academic performance is dependent on an existing domain specific knowledge (Chi, 1985; Glaser, 1984; Minneart & Jansen, 1996).

Many correlational studies support the facilitating role of domain specific prior knowledge on study success (Chi, Glaser & Rees, 1982; Chase & Simon, 1973; Lesgold, Feltovich, Glaser & Wang, 1981; Voss, Greene, Post & Penner, 1983). Dochy (1992) was able to attribute about 40% of the variance in study achievements to existing prior

knowledge, pointing to exceptionally high correlations. Bloom (1976) even found correlations as high as .80 between prior knowledge and study results with first graders. Furthermore, a recent meta-analytic study on Graduate Record Examination (GRE) results shows relatively high validities for specific subject knowledge tests, suggesting that academic knowledge of the field of study plays an important role in the quality of graduating (Kuncel, Hezlett, & Ones, 2001).

Data from correlational and experimental research can be combined and articulated into causal models. In causal modeling, or structural equation modeling, direct and indirect effects of different variables on study achievements can be traced. Parkerson, Lornax, Schiller and Walberg (1984) concluded from their causal model that prior achievement had the highest direct (.72) and indirect effects on study achievement. This result seems in support of the important role of prior knowledge for study results, although the prior achievement results were not controlled for *g,* thus making the exact interpretation of prior achievement somewhat unclear. Parkerson et al. also found an indirect effect of prior achievement on motivation (.25) and study time spent (.07), and a direct effect of home/living conditions on prior achievement (.14). Together, the results discussed above support the facilitating role of prior knowledge on study results.

Realistic study expectations

With respect to the large first year drop-out rate of OUNL psychology students, the question arises whether a proper amount of prior knowledge is the only factor responsible for this large drop-out rate. In the model of Parkerson and others (1984), the factor (lack of) motivation can very well be ruled out in the case of the OUNL students, because it can be assumed that they are optimistic and enthusiastic about studying Psychology. Quality of instruction and aspects of social integration show low effects on study results. The latter is almost absent at studying Psychology at the OUNL. Quality of instruction is another issue; the character of the study is heavily based on the 'distant learning' concept.

But a factor that comes up as a candidate for influencing (lack of) study success is what might be labeled as 'unrealistic expectations', more or less a form of 'negative prior knowledge'. The assumption is that many students have a grossly 'romanticized' picture of the academic study of psychology, thereby underestimating its scientific character.

This romantic picture entails associations with practical self knowledge and being able to solve one's own and others' life problems, completely ignoring the fact that methodology and statistics make up a large part of (the first year of) the study of psychology, thus requiring specific capacities and skills. Realizing that studying psychology at the OUNL, because of its individual character of distant learning, lacks for a great deal what Tinto (1993) called 'social integration', the focus of this study merely will be at what Tinto (1993) labeled 'academic integration'. Academic integration denotes the phenomenon of being able to meet the required academic standards. For the OUNL this would mean passing two introductory exams within the first half year of study. The question is 'do enrolling students have a realistic picture of the subjects covered by these two exams?'

Little or no research has been done on the concept of 'realistic study expectations'. For our study we defined 'realistic study expectations' as having true knowledge of what it takes (abilities, capacities, time) to succeed in the academic study of psychology and what the study leads to (qualifications, knowledge and skills). 'Realistic' in our study thus was split in 'effort expectations' and 'profit expectations'. Effort expectations point to the investment side of studying: do students have an adequate picture of the capacities, skills and time needed for succeeding in the study of psychology? Profit expectations concentrate on having a true knowledge and understanding of study benefits, such as acquired scientific knowledge and skills. The assumption is that students with expectations less matching the actual study situation will perform significantly less than students with expectations matching the actual study situation.

Hypotheses

Two hypotheses were formulated:

Hypothesis 1: A positive significant correlation will be found between the amount of academic psychological prior knowledge and the exam results of the two introductory exams.

Hypothesis 2: Students having more realistic study expectations will score significantly higher on the two introduction exams than students with less realistic study expectations.

<center>**Method**</center>

Participants

A pilot study was undertaken with 98 first year students to construct a prior knowledge test and a realistic study expectations questionnaire (RSEQ) (see *Instruments*). From these 98, a number of 55 responded (56%). For the main study, 212 students were randomly selected from the first year group, called 'kennismakingstraject'[9] (KMT). The KMT consists of two introductory courses, with separate exams (Psychology 1A and Psychology 1B). The two courses are being supported with eight non-compulsory lectures. From these 212, a number of 63 responded (30%). For the definitive analyses, both respondent groups were taken together (55+ 63=118 participants), using only the group of items of both instruments which showed acceptable reliability and discriminating power after analyzing the pilot study results. Table 4.1 lists the subject characteristics of the main study.

Table 4.1:

Respondent sex, education and age (percentages between brackets), n=118.

sex		education*			age		
male	28 (24)	LVC (lbo)	3	(3)	21-30 year	25 (21)	
female	81 (69)	MPGE (mulo/mavo)	7	(6)	31-40 year	53 (45)	
unknown	9 (8)	HPGE (havo)	14	(12)	41-50 year	25 (21)	
		High school (vwo)	10	(9)	51-60 year	4 (3)	
		MVE (mbo)	34	(29)	unknown	11 (9)	
		College (hbo)	27	(23)			
		University (wo)	12	(10)			
		unknown.	11	(9)			

* LVC= Lower Vocational Education; MPGE= Middle Preparing General Education; HPGE=Higher Preparing General Education; MVE=Middle Vocational Education

---

[9] Introductory course of Psychology

From table 4.1 it can be seen that women make up about 70 percent of the first year group. More than half of this group has an education at least at the middle vocational or college level. The mean age was 36 year. Of these students 46% had the intention to graduate in Psychology.

Instruments[10]

Prior knowledge test

Dochy (1992**)** discriminates three different ways of testing prior knowledge. With Subject Oriented Knowledge State Tests (SO KST) one tests only part of a certain subdomain; with Cross Domain Knowledge State Tests (CD KST) one tests a part of all existing subdomains, and with Domain Specific Knowledge Tests (DS KST) knowledge of a whole domain is being tested. We were interested in domain specific knowledge, or, to be precise: what knowledge do enrolling students have of the field of academic psychology, as found in introductory texts such as Hilgard, Atkinson and Smith (2000), Gleitman (1993) and Bernstein, Roy, Srull & Wickens (1992). We constructed a prior knowledge test having the same format as the official exams, that is, 40 multiple choice questions with four alternatives each. With 40 items, an acceptable content validity of the domain in question seemed guaranteed. The prior knowledge test consisted of general questions concerning the field of academic psychology as taught in the Introductory course (KMT). Sub-domains consisted of physiology, needs, learning and conditioning, perception, intelligence and individual differences. Coefficient α was .64. After skipping the 5 items with a very small or negative item-total correlation (items 3, 6, 19, 24 and 38) α increased to .70, which seemed acceptable. Deleting more items would lead to an insufficient coverage of the domain in question. So the definitive prior knowledge test consisted of 35 items.

---

[10] Instruments can be obtained from the author. Address: wim.bloemers@ou.nl

The realistic study expectations questionnaire (RSEQ)

The concept of 'realistic study expectations' was divided into four subdomains: study contents, study benefits (knowledge and skills), study skills and time spent. For each subdomain, 3 to 5 items were constructed, with the exception of the domain time spent, which consisted of only one question. All 14 items had to be scored on a five point Likert scale, running from 1: very little/very small, to 5: very large/very much. The minimum score was 14, the maximum score was 70.

The pilot study (n=55) revealed low alpha's, ranging from -.40 to .40 for the different subscales. Items with a negative item-total correlation were removed and several items were reformulated. For the main study, twelve items remained. An additional four extra exploratory questions, scaled on a five point Likert scale were included, concerning issues as the goal of enrollment, difficulty of examination, expected exam performance and one's expectations about graduating in psychology. The general idea was that the data on these questions might reveal important information about students' exam performances.

Criterion measures: Exams in Psychology 1A and 1B

The criterion consisted of the exam result on the two introductory courses in psychology (Psychology 1A and Psychology 1B). Both exams contain 40 multiple choice questions. Although no reliabilities of the exam scores are known, both exams are regularly controlled and updated by means of item analysis. So the reliability should be more than acceptable ($\alpha \geq .75$). At the time of analysis, the 1A exam had been taken by 62 students, and the 1B exam by 30 of the group of 118 respondents. While students are free to set their own exam date the majority of the exams are usually taken about four months after the start of the introductory course.

Procedure

For the pilot study, the prior knowledge test and study expectations questionnaire were sent to 98 randomly selected first year psychology students one month before the start of

the introductory course. A total of 55 tests and questionnaires were returned (56%). All respondents were given feedback on their score on the prior knowledge test.

After analysis and revision of the two instruments (see above) the definitive version of both instruments was sent to 212 randomly selected first year psychology students one month before the start of the second introductory course. The number of respondents was 63 (30%). All respondents received feedback on their results on the prior knowledge test. Because of the low number of respondents, both the respondents of the pilot and the main study were merged into a definitive file, using only those items of the instruments which showed sufficient item-total correlation in the pilot-study. This meant that final analyses were conducted on the 35 items of the PKT and 8 items of the RSEQ.

Results

In order to get a clear picture on the two hypotheses as stated, preliminary reliability analyses were undertaken on the prior knowledge test and the RSEQ. The 35 item prior knowledge test showed rather poor reliability with an α of .68 and an Se of 2.6. After eliminating 10 more items, the resulting 25 item test showed an α of .75, a mean score of 12.7 (sd 4.4) at a maximum of 25, with an average p value of .51. The standard error of measurement of this remaining 25 item test was 2.15.

Cluster analysis and successive PCA combined with varimax on the 14 study expectation items resulted in two scales showing acceptable reliability:

(1) A two item 'biological knowledge' scale, consisting of the items:

-How much neurological knowledge do you think the study of Psychology contains?
-How much biological knowledge do you think the study of Psychology contains?

(2) A four item 'self knowledge' scale, consisting of the items:

-In what way do you expect that a scientific study will increase your self knowledge?
-In what way do you expect that the scientific study of psychology will help you to solve your own (life) problems more adequate?
-In what way do you expect that the scientific study of psychology will help you in being able to solve other people's problems?
-In what way do you expect that after the study of psychology you will have a better understanding of the behavior of other people?

The two item biological scale had a mean total score of 7.10 (sd=1.21), α=.74. The four item self knowledge scale had a mean total score of 13.84 (sd=2.56), α=.76.

After establishing the 'definitive' predictors a multiple regression analysis (Table 4.2) was executed with the introductory exams 1A and 1B as criterion variables, and the prior knowledge test and the biological and self knowledge scales and general background variables as predictor variables. These latter are important for our study as the OUNL has a student population with a considerable higher age and lower level of previous education than a 'regular' university.

Table 4.2:

ß-coefficients of predictor and background variables in relation to the two exam results Psychology 1A (N = 62) en Psychology 1B (N = 30) and 1A + 1B taken together (N = 30).

	ß-coëfficients 1A[1]	ß-coëfficients 1B[2]	ß-coëfficients 1A + 1B	
Definitive Prior Knowledge Test	.03	.31	.38	
'Biological scale'	-.08	-.19	-.22	
'Self knowledge scale'	-.25*	.14	.11	
Age	-.42*	-.28	-.33*	
Sex	-.20	.27	.11	
Previous education[3]	.12	.16	.23	
Intended study time	.26	.20	.16	
R	.49	.61	.65	
R^2	.24	.37	.42	* = $p<.05$

[1]N=62, no. of items =40, mean grade=6.85, sd=1.51; [2]N=30, no of items=40, mean grade=7.27, sd=1.45
[3] Previous education coded as follows: LBO/Lower Vocational Education=1; Mavo/Mulo/MPGE= Middle Preparing General Education=2; HAVO/Higher Preparing General Education=3; MBO/Middle Vocational Education=4; VWO/High school=5; HBO/College=6; University=7

Both our two hypotheses had to be rejected, as can be seen from Table 4.2. Neither prior knowledge nor realistic study expectations as measured with the Biological and Self knowledge scales show consistent significant results with the criterion variables 1A and 1B. Age, intended study time and the self knowledge scale show the largest ß-weights for the first exam. The change in ß value for the self knowledge scale could be explained by the fact that most students scoring high on this scale dropped out after the first exam or

failed the second exam. The definitive prior knowledge test shows relative large ß-weights for the second exam and both exam results combined. Age is the only predictor with more or less consistent (significant) results across both exams. Age however correlated significantly with previous education (R= -.24, p< .01).

Tabel 4.3 sheds some more light on the intercorrelations of the main predictors of this study and the two criteria. This table adds one significant aspect to the findings

Table 4.3:

Matrix of intercorrelations of scores on definitive prior knowledge test (DPKT), biological and self knowledge scales and psychology exams 1A and 1B.

	DPKT	'Biological'	'Selfknowledge'	Psychology 1A
'Biological'	-0.15			
'Selfknowledge'	-0.27**	0.06		
Psychology 1A (N=62)	.03	-0.17	-0.19	
Psychology 1B (N=30)	.33*	-0.28	0.03	0.70**

*p<.05 **p<.01

already discussed, the negative correlation (-.27) between the 'self knowledge' scale and the score on the DPKT. In order to get some more understanding of the rather ambiguous influence of the DPKT results on exam scores an analysis of variance was conducted, creating three groups of students: one group having done no examination at all (n=56); one group having done one examination (1A; n=62) and a group having done both examinations (1A and 1B; n=30). Analysis of variance showed no significant differences between the three groups on the DPKT (F=.99, ns.). This meant that results on the DPKT do not differ between students having taken no, one ore two exams.

Discussion

The results showed that prior knowledge in this study is a limited predictor of study success for first year psychology students. Study advice based on prior knowledge, as it seems from this study, should be handled with care. Nobody profits from a situation were enrolling students are being discouraged with a low score on an instrument of which the predictive power is still unclear. Other predictors like the 'self knowledge' scale, age, sex, previous education and intended study time showed more promising results,

although the combined multiple R of these factors is not very impressive, being largest for the combination of the 1A and the 1B exams. For the first exam, the 'self knowledge' scale (in a negative way) and intended study time seem the most important predictors of success.

How is it possible that prior knowledge, a concept backed by clear empirical support (Chi, Glaser & Rees, 1982; Chase & Simon, 1973; Lesgold, Feltovich, Glaser & Wang, 1981; Voss, Greene, Post & Penner, 1983; Dochy, 1992) does not show any statistically significant predictive power in this study? A few possible explanations come to mind. The small sample seems the most obvious reason for the lack of predictive power of our study. Furthermore, it could be that the DPKT has insufficient content validity for the domain in question, especially for the first exam. The shift in ß values suggests that the DPKT had more content validity for the second than for the first exam. However, all the items of the prior knowledge test were directly based on item format and content of the actual two exams. Another explanation is that the prior knowledge test shows little differentiation among students. From the mean p-value of the DPKT it could be argued that the DPKT is too difficult for potential students. When actual examination norms are applied, only about 17% of the students would score a sufficient result on the DPKT, whereas about 80% of the students, who do the 1A and 1B exams, pass. In other words: students performing badly on the DPKT can perform (very) well on the exams. Analysis of variance supported this argument, as no difference was found between the DPKT scores of the groups of students having passed no, one or two exams. An explanation of this relative 'underperformance' on the DPKT could be that enrolling students have barely any knowledge of the domain of academic psychology. Comparing this with the field of economics, being a regular subject in secondary education, it should not be that surprising that Dochy (1992) found a far greater predictive validity for a DPKT for students of economy. Although the DPKT shows some predictive value for the second exam, the lacking predictive power of the DPKT for the first exam is especially something to be further investigated. Studies like Busato's (1998) show the utmost importance of the first exam result for study success. However, from our study, amount of effort, age and previous education seem more important than prior knowledge. This is

an important result that can be of great use for enrolling psychology students of the OUNL.

Recommendations for further research

Future studies should concentrate on a larger group of students, especially in view of the high drop-out rate of first year students. This high drop-out range leads to a very severe restriction of range in the criterion variables, leaving little room for significant correlational results.

Furthermore, keeping Busato's (1998) result in mind, it seems advisable to concentrate DPKT items fully on the first exam content. Curriculum developments at the OUNL facilitate this point of view, because in the near future there will be only one introductory exam for psychology students, containing both the domains of the 1A and 1B exam.

The concept of realistic study expectations can be further articulated, focusing on the content of study expectations. Naïve 'romantic' expectations show a promising (negative) predictive power, making a point for a more thorough and scientific oriented briefing of potential students of psychology at the OUNL. That fewer students will enroll with such a policy seems inevitable, facing the OUNL with the dilemma of sticking to the commercial success of selling first year courses, or increasing its percentage of graduates. From a utility point of view, a free entrance seems a self killing prophecy, thereby crushing the myth that just 'anybody' can graduate in an academic study. Moreover, this study points to the fact that younger people, with more and higher previous education have more chances of graduating. Having more and higher education is an advantage in two ways: one is familiar with what it takes to succeed in an academic environment, and there is a substantial relation with information processing capacities (g; Neisser et al., 1996). Future research therefore should concentrate on the role of general information processing capacity, personality variables (Okun & Finch, 1998), including self efficacy (Bandura, 1977) and biographical aspects (Reilly & Chao, 1982).

Implications for HEIDI

Prior knowledge from our study seems a construct too weak in predictive validity to be of practical use for potential students. Therefore we would advise at this moment not to include measurements of prior knowledge in the HEIDI with regard to exam result predictions. If future research with an improved prior knowledge test would show substantial predictive validity for the first (and second) exam result, inclusion of this instrument could be taken into consideration. There is however another point of view on the usefulness of prior knowledge as a feedback measure for potential students. Such a measure could give potential students useful information on what sort of knowledge is considered important, and what kind of questions they can expect. Thus, from a realistic study preview standpoint, it could be important to supply students in advance with information on what kind of psychological knowledge they are expected to master, and in what way this knowledge is being examined (multiple choice). This kind of information also pertains to the field of 'realistic study expectations'. Thus, utilizing a PKT in the HEIDI website from a 'realistic study preview' focus could synergize the concepts of prior knowledge and realistic study expectations.   The items of the PKT could give potential students relevant information about what (not) to expect during the study of psychology. The website furthermore seems an ideal medium to draw potential students' attention to possible bottlenecks in the study, bottlenecks which are obviously not taken to heart from the available study brochures.

CHAPTER 5

The effects of Self Efficacy and personality characteristics

(Submitted for publication: *The Journal of Adult Education*)

5. Effects of self efficacy and personality on study success

In 1999, the project Higher Education Interactive Diagnostic Inventory (HEIDI, Bloemers, 1999) started at The Open University of the Netherlands (OUNL). The aim of this project was to create a website for potential students of psychology, giving them feedback on relevant predictors for study success. The idea behind this is to enable students at to make a more realistic and fruitful choice of study. Results from HEIDI are not to be used as a selection decision; its only aim is to give an advice on which students themselves make a decision to enroll.

The choice of predictors for inclusion in the HEIDI website is based on Schouwenburg (1998). Schouwenburgs model (see chapter 1, *General introduction)* is built on four clusters:

(1) General intelligence and practical study intelligence

(2) Personality and biodata

(3) Prior knowledge and study know-how

(4) Biographical aspects

According to Schouwenburg, all clusters have a direct influence on study success. Apart from these direct effects, according to Schouwenburgs model, personality variables also have an indirect influence on study success. This indirect effect is moderated by Self Efficacy (SE): students' perceptions of their capabilities for succeeding in the study (Bandura, 1986). In chapters 2, 3 and 4 we studied the predictive value of intelligence and practical study intelligence, personality and biodata, prior knowledge and realistic study expectations. In this chapter the focus is on the role of personality variables and SE. The central question is: 'What is the relationship between personality dimensions, self efficacy and study results of first year Psychology students at the OUNL?'

For inventarising personality measures the Five Factor Personality Inventory (FFPI) (Hendriks, 1998) was used. A self constructed SE scale was used for SE measures. Criterion measures consist of the two introductory multiple choice exams, covering knowledge of the first year course 'Introduction to Psychology' (KMT).

We expect to find a significant relation between relevant personality variables (conscientiousness, emotional stability and openness/autonomy) and study success. Furthermore, the role of SE for study success was investigated as an independent construct and from the viewpoint of a moderating role of SE on the relevant personality variables.

In the next paragraph, we will focus on the importance of the Big Five for study success, followed by a concise treatise of SE, and SE and study success. Also highlighted is the relation between the Big Five and SE, before turning to our current study.

The Big Five and academic achievement

For a concise sketch of the Big Five personality dimensions, the reader is addressed to chapter 3 were we gave an overview of the Big Five, it's development and recent criticisms. We will now turn to the role of the Big Five for study success.

Compared to research on the Big Five in Industrial Organizational settings, research within the Big Five framework on study success is relatively rare. But from studies of Digman (1990), De Raad and Schouwenburg (1996) and Okun and Finch (1998) some conclusions can be drawn. According to Okun and Finch, emotional stability, agreeableness, intellectual autonomy and conscientiousness should show a positive correlation with study proficiency. They further claim that students high on Openness (intellectual autonomy), a dimension associated with eagerness and imagination, would score better exam results than students low on this dimension. However, only Openness showed a positive correlation with study results (Okun & Finch, 1998). Digman's study on freshmen showed that Conscientiousness is positively correlated with study results. Students high on Conscientiousness work harder, more organized and with greater determination.

According to De Raad and Schouwenburg (1996), all Big Five factors seem to influence learning and education in a variety of ways. Reasoning within the AB5C framework (Hofstee & De Raad, 1991), they claim that Conscientiousness fulfils a prominent role, combined or not with facets from extraversion, Emotional Stability and Intellectual Autonomy. Within the domain of Emotional Stability, facets concentrating on perseverance, stability and control should be of special importance for study results. The

positive side of Intellectual Autonomy also seems a promising candidate for study results, but empirical evidence is lacking on this issue. Facets of the Extraversion domain could go either way, depending on the demands of the study environment. The influence of Agreeableness on study results seems hard to predict, but quiet, cooperative students seem to be at an advantage.

Summing up it seems that Conscientiousness and to a lesser degree Intellectual Autonomy and Emotional Stability seem important predictors for study success (Schouwenburg, 1998; Okun & Finch, 1998; Wolfe & Johnson, 1995). In our current study it is of importance that the learning process at the OUNL is for the most part characterized by self-study. The OUNL utilizes distant learning as its main feature, thus capitalizing on student discipline and self control. From industrial psychology studies (Barrick & Mount, 1991, 1995; Hurtz & Donovan, 2000) Conscientiousness emerges as the (relatively) most important predictor for work performance. Compared to industrial settings, the study environment at the OUNL is characterized by low external control and feedback, putting even more importance on student self discipline and perseverance. So for our present study, we expect a significant influence of Conscientiousness on study results, and moderate effects of Emotional Stability and Autonomy. We will now turn to the role of Self Efficacy, the other predictor of importance in our present study.

Self efficacy

Self Efficacy (SE), a term first used by Bandura (1986), has its roots in social-cognitive theory. Starting point is the assumed mutual relationship between behavior, cognitions and environment. Bandura defines SE as the judgment of an individual about his or her own capability of successfully executing specific behavior in specific situations to attain a desired result (Bandura, 1986). SE expectations are thus situation-specific and they will vary with the activity in question (Bandura, 1997).

According to Bandura (1982, 1986) SE is the most important aspect of human action/behavior and, dependent on the acquired knowledge and skill level, it also determines individual performance. People positive on SE will develop behavior which contributes to the (further) mastering of a domain specific task, the development of task relevant skills, knowledge and the capacity to perform. People low on SE will not, or to a

lesser degree, develop such cognitive behavioral features. Bandura further reasons that people high on SE will develop a higher level of aspiration; they will have more perseverance and are more open to new stimuli and challenges than people low on SE. They will also show more perseverance and effectiveness in goal directed behavior. People low on SE will experience more tension and stress executing goal directed behavior (Bandura, 1997). For our present study, reasoning from Schouwenburgs model, the above characteristics of SE imply that our students should have a rudimentary impression of their intelligence, personality, study related behavior, study circumstances and biographical features, and that this impression in turn determines their level of self efficacy. SE in turn influences future study behavior and study results.

SE and study results

Multon, Brown and Lent (1991), in their meta-analysis, found a significant positive effect for SE on study results ($\underline{r}$ = .38). Taking a broader perspective, their results confirmed the overall implications of Bandura's theory on SE for study performance and study results. In general, students high on SE show better time management, use more effective problem solving strategies and show prolonged effort compared to students low on SE (Collins, 1982; Bouffard-Bouchard, Parent, & Larivee, 1991; Chemers, Hu, & Garcia, 2001).

Moreover, positive study results reinforce the positive image of one's own competency (Vrugt, Hoogstraten, & Oort, 1998; Silver, Mitchell, & Gist, 1995). Bouffard-Bouchard (1990) and Cervone en Peake (1986) manipulated SE perceptions of students by using imaginary performance information during feedback on performance. Students given positive feedback showed a better performance than students (with the same skills level) given less positive feedback.

Although performance boosting techniques will not lead to ever increasing effects, based on the results above and Bandura's SE theory in general it can be expected that students showing high confidence in their capacities to perform well on an academic level will show a better study performance than students low on academic SE. Our current study concentrates on Academic SE of psychology freshman at the OUNL. Based

on the literature above and student interviews we decided to split the general field of Academic SE into four SE competency domains:

(1) SE-Academic (SEA): The perception of one's capacity to understand and process a large amount of complex information at an academic level.

(2) SE-Networking (SEN): The perception of one's capacities to activate and utilize others for study support and study organization.

(3) SE-Selfmotivation (SES): The perception of one's abilities for self motivation during a long period of time with considerable setbacks and difficulties.

(4) SE-Learning (SEL): the perception of one's capabilities to plan and organize the study and execute apt study skills at the academic level.

In general, we expect a significant positive correlation of all four SE domains with first year study results.

Apart form a positive influence of SE on study results, the question on the construct validity of SE itself arises. In our study we take a closer look at possible links between SE and Big Five personality dimensions. At first sight, SE seems to show considerable overlap with some Big Five dimensions, especially Conscientiousness, Emotional stability and Autonomy. We do not investigate the possible (contaminating) role of intelligence or other factors, which could have a moderating role on SE. We will now take a closer look at possible Big Five-SE relations, before stating our hypotheses on expected Big Five-SE relations.

The Big Five and SE

Conscientiousness and Emotional Stability seem at first sight good candidates regarding prolonged effort and self confidence at executing task directed behavior. We will now consider each of the five personality dimensions and their possible relation with SE.

Extraversion

Extraversion is of interest for SE because of the relation with a high energy level. This caused Hendriks and others (1998) to talk about 'doers'. According to these authors a positive score on Extraversion refers to energetic behavior. Studies by Friedman and Rosenman (1974) and Taylor, Locke and Gist (1984) showed that extraverts or 'type A'

people have a higher energy level, which can result in a higher SE level. So for our current study we expect a positive relation between Extraversion and SEN.

Agreeableness

Agreeableness is mainly about behavior towards other people (Hendriks et al., 1998). Relevant aspects of behavior residing under this factor are mild, peaceful and lenient versus uncomplying and stubborn. We did not find any studies on the relation between agreeableness and SE. Moreover, since the (first year) study at the OUNL has a strong emphasis on self study we do not expect to find significant relationships between Agreeableness and aspects of SE. The only link that comes up could be of a negative nature, implying that people high on SE could show low Agreeableness because of their strict task directedness.

Conscientiousness

Conscientiousness refers to a precise and systematic way of working, to planned goal directedness (Hendriks et al., 1998). In the fulfillment of tasks, but also in general, the conscientious person manifests responsible and performance directed behavior. According to Bandura (1997) persons high on SE will show greater determination, they have a stronger commitment to often more difficult goals. They are more organized and more goal-directed. They also are better in planning their study. For our current study we expect a positive significant correlation between Conscientiousness and SEA and SEL.

Emotional Stability

People high on Emotional Stability will have more self-confidence and feel better equipped when confronted by difficult situations. Gist and Mitchell (1992) presuppose that people with a low self esteem also are low on SE for various tasks. These people are not sufficiently self-confident with regard to the necessary competencies for succeeding on the task in question. We therefore expect in our present study a positive correlation between Emotional Stability and SEA. Because of the relation between Emotional Stability and being able to cope with serious problems we also expect a positive relationship between Emotional Stability and SES.

Intellectual Autonomy

Intellectual Autonomy is associated with intelligent, imaginative, original and artistic behavior. Although we found no studies articulating possible relations between SE and this dimension, because of the small but significant relationship with intelligence we expect a positive relationship between this dimension and SEA. People high on Intellectual Autonomy should better be able to process and understand complex information on an academic level.

We will now state all of our hypotheses explicitly regarding our current study.

Hypotheses

Our current study is based on four groups of hypotheses:

1) Big Five: We expect Conscientiousness, Emotional Stability and Intellectual Autonomy to have a significant positive influence on first year study results.
2) SE: We expect the concept of SE and all four subscales as used in our study to show a significant positive influence on first year study results.
3) Big Five and SE: We expect significant positive correlations between Extraversion and SEN, Conscientiousness and SEL, Emotional Stability and SES and Intellectual Autonomy and SEA.
4) Moderating effect of SE: We explore the possible moderating influence of SE on the relationship between Big Five dimensions and study results.
    We also explore the influence of a number of biodata, such as age and sex, on first year study results.

Method

Participants

Participants in our study were students, starting September 2000 with the introductory course of Psychology at the OUNL. This first year course consists of the subjects Psychology 1A and 1B, based on an adaption of Gleitman's *Introduction to Psychology* (Gleitman & Tomic, 1993). All 496 students registered in the Netherlands were asked to participate. Of these, 161 students (27.9%) were willing to participate. Of these 161, 23 (14.3%) were males and 138 (85.7%) were females. Of all the participants, 45% had a previous education at or below the middle vocational level (MBO). About 70% of the participants indicated the attainment of the bachelors degree in Psychology as a serious study goal; 32.3% wanted to realize a masters degree. The mean age of our participants was 36.14 years (SD=10.25), with a range of 19-71. Males had a mean age of 42.04 years (SD=13.48); females 35.15 (SD=9.32).

Instruments[11]

  FFPI

As a measure of personality characteristics we used the Five-Factor Personality Inventory (FFPI), a reliable and valid personality questionnaire (Hendriks, 1997). With the FFPI someone's Big Five profile is established in an efficient way, taking about 15 minutes. The FFPI consists of 100 compact and comprehensible behavioral statements. All items are formulated in third person singular, scored on a five-point scale with the extremes 'not at all applicable' and 'very applicable'. Every Big Five dimension consists of 20 items. Examples of items are: 'Likes to talk' (Extraversion); 'Considers the feelings of others (Agreeableness); 'Likes working according to a schedule' (Conscientiousness); 'Is able to detach from problems' (Emotional Stability) and 'Conforms to the majority' (Intellectual Autonomy).

    For scoring the FFPI a special computer program is utilized. Scoring is based on anchored factor scores (Hofstee & Hendriks, 1998). This means that not, as usually is done, the mean score of a population on every dimension is set at zero and the standard deviation at 1. With the FFPI, information on the distance of a participants score to the

---

[11] Instruments can be obtained from the author. Address: wim.bloemers@ou.nl

81

mid scale point is kept available. With this technique, FFPI participants can be given extra information, that is, both on their score from scale midpoint and their score from the mean of the norm group. For individual research this creates the possibility to give participants feedback in the form of 'You scored on the positive side of Emotional Stability (above scale mid point), but lower than the mean norm group score on this dimension' (Hendriks, Hofstee & De Raad, 1998).

The Self Efficacy (SE) scale

The relevant questionnaires of SE we encountered all presupposed some familiarity with the task at hand. This is in accordance with the theory of SE, and with empirical results which show that more task experience leads to more precision of task performance predictions (Gist & Mitchell, 1992; Vrugt, Langereis & Hoogstraten, 1993).
Our participants however, psychology freshman, had no or very little experience with the task at hand and they had not been given any feedback whatsoever on study relevant competencies and performances. Therefore, and because any SE operationalization refers to domain and task specific capacities, we decided to construct our own version of an SE questionnaire, attuned specifically to our participants and the tasks at hand. By inventarising similar instruments (Wood &Locke (1987); Bandura (1997); Pintrich & De Groot (1990); Teeuw, Schwarzer & Jerusalem (1994) we obtained and adapted a first set of 37 useful items. Items were initially assigned to four SE scales:

   (1) Academic SE (SEA): Confidence in one's own capacity to process and understand large amounts of complex information at an academic level

   (2) SE for Networking (SEN): Confidence to be able to utilize others for organizing and supporting the study

   (3) Self regulating SE (SES): Confidence to be able to motivate oneself during longer periods with considerable problems and setbacks

   (4) Self regulated Learning SE (SEL): Confidence in being able to plan and organize the study and having sufficient study skills

Items were measured on a five point Likert scale, ranging from (1) *Totally disagree* to (5) *Totally agree.*

A group of 117 first year students took this draft SE questionnaire. Principal components analysis (PCA) generated 11 factors with eigenvalues > 1. Varimax rotation however failed to generate a clear and simple factor structure.

We used reactions of students to further improve our SE questionnaire. Ambiguous items were removed and we increased the level of competence in several items to get more differentiation between students. Furthermore, we rewrote items as much as possible in actual behavioral terms ('I am able to…..'). Following advice from Bandura (1995), to increase differentiation, we decided to use a seven point scale instead of the five point scale from the pilot study. Scale extremes varied from (1) 'not good at all' to (7) 'very good'. The first version of the SE questionnaire thus consisted of 33 items, spread over four scales. Examples of items, followed by the initial scale and item number between brackets are:

How good do you think you are at this moment at:

-Remembering relevant concepts and facts for an exam (SEA06)

-Being able to get help from your study counselor when you have study problems (SEN13)

-Staying calm when confronted with unexpected study problems (SES16)

-Creating time to organize the study (SEL24).

All four scales listed α's ≥ .84, with a mean α of 87.

Further validating of the SE questionnaire

We factoranalyzed the above version of the SE Questionnaire with our definitive sample (n=160). PCA, followed by varimax (Table 5.1) this time resulted in seven factors, of which the first four can be clearly identified as our original four factors. Together, these four factors explained 55.2% of the variance. The fifth factor could be interpreted as a concentration factor. The interpretability of factors six and seven was difficult.

To get further evidence on our four specific domains of SE resulting from the PCA, we performed a confirmatory factor analysis (CFA) using AMOS 3.6 (Arbuckle, 1997). This CFA resulted in two possible models for the factor structure of our SE questionnaire (Figure 5.1 and Figure 5.2). According to the Comparative Fit Indexes (CFI) (Bentler, 1990 and the Tucker-Lewis Indexes (TLI) (Tucker & Lewis, 1973; McDonald, & Marsh, 1990) both models have an acceptable fit. For both indexes a value of ≥ .90 refers to an

acceptable model fit (Bentler & Bonett, 1980). Although the fit of model 1 is slightly higher, model 2 has more degrees of freedom. Model 2 is more parsimonious, because it specifies only 4 relations, while model 1 specifies 6 relations. Furthermore, from table 5.3 it can be concluded that all the SE subscales show significant positive inter correlations. Evaluating both models, model two (Figure 5.2) seems preferable because of its more parsimonious structure, pointing to one central SE factor at the root of the model.

All further results in our study are based on the SE scales resulting from the above CFA. That means we used sum scores for the respective scales SEA based on 7 items ($\alpha = .86$), the scale SEN (3 items; $\alpha = .86$), the scale SES (4 items ($\alpha = .83$) the scale SEL (4 items; $\alpha = .83$) and the Total SE scale (18 items ($\alpha = .88$).

Criterion measures

As criterion measures for first year study results we used results on the exams psychology 1A and Psychology 1B. These exams both consist of 40 multiple choice questions, covering an introductory text on Psychology (Gleitman & Tomic, 1993). We shall also look at the criterion 'succeeded on two exams versus not succeeding'.

Procedure

In April 2000 all first year study counselors were asked to participate in the SE pilot study. The draft version of the SE questionnaire, together with an introduction letter was sent to them by e-mail, asking them to distribute the draft questionnaire among their students. Based on the results of this pilot study, a definitive version of the SE questionnaire was constructed. In September 2000, all first year students registered in the Netherlands were sent two questionnaires, the definitive SE questionnaire and the FFPI accompanied with a letter explaining the purpose of the study. Students not responding were sent a reminder after six weeks.

Table 5.1:

Results of the PCA on 33 SE Questionnaire items after varimax-rotation

		Factors						
Items		1	2	3	4	5	6	7

Scale: SEA: Self-efficacy expectations towards academic performance (n = 158, n items=9)

SEA1	Individually grasping the essence of a subject	.77	--	--	--	--	--	--
SEA2	Understanding facts, concepts and arguments	.73	--	--	--	--	--	--
SEA3	Grasping the essence of a subject in taking notes	.75	--	--	--	--	--	--
SEA4	Explaining facts concepts and arguments to others	.63	--	--	--	--	.48	--
SEA5	Determine important facts and concepts for an exam	.69	--	--	--	--	--	--
SEA6	Remembering important facts and concepts for an exam	.62	--	--	--	--	--	--
SEA7	To pass the KMT within one year	.52	--	--	--	--	--	.56
SEA8	To pass the Psychology 1A exam	.66	--	--	--	--	--	--
SEA9	To pass exams on the first time	.60	--	--	--	--	--	--

Scale: SEN: Self-efficacy expectations for utilizing others (n = 160, n items=5)

SEN10	Getting study support from family and friends	--	--	--	.84	--	--	--
SEN11	Getting study interest from significant others (SO)	--	--	--	.80	--	--	--
SEN12	Getting support from SO, for creating study time	--	--	--	.79	--	--	--
SEN13	Realizing support from teachers by study problems	--	--	--	.53	--	--	--
SEN14	Realizing support from students by study problems	--	.43	--	.61	--	--	--

Scale: SES: Self-efficacy expectations for self regulation (n = 160, n items=6)

SEZ15	To show perseverance with boring subjects	--	.42	.67	--	--	--	--
SEZ16	To stay calm at unexpected study problems	--	--	.70	--	--	--	--
SEZ17	To trust your problem solving capacities	--	--	.66	--	--	--	--
SEZ18	To show determination at reluctant study problems	--	--	.79	--	--	--	--
SEZ19	To show extra effort when failing an exam	--	--	.69	--	--	--	--
SEZ20	To motivate yourself without support from others	--	.56	.43	--	--	--	--

Scale: SEL: Self-efficacy expectations towards self-regulated learning (n = 161, n items=13)

SEL21	Solely concentrating only and only on your study	--	.78	--	--	--	--	--
SEL22	To realize study tasks within given deadlines	--	.58	--	--	--	--	--
SEL23	Actually realizing your study planning	--	.75	--	--	--	--	--
SEL24	Creating time for organizing your study	--	.81	--	--	--	--	--
SEL25	Concentrating on the subject during college	--	--	--	--	--	.75	--
SEL26	To concentrate yourself while studying	--	--	--	--	.68	--	--
SEL27	Concentrating while executing study tasks	--	--	--	--	.68	--	--
SEL28	Concentrating on the questions during an exam	--	--	--	--	.71	--	--
SEL29	Studying individually -without counselor contacts-	.56	--	--	--	--	--	--
SEL30	Doing individual papers and projects	.58	--	--	--	--	--	--
SEL31	To realize a quiet place for studying	--	--	--	.42	--	--	--
SEL32	Motivating yourself to study at home	--	.64	--	--	--	--	--
SEL33	Participating in discussions during college	--	--	--	--	--	.70	--

	1	2	3	4	5	6	7
Sum of squared loadings	**11.02**	**3.35**	**1.97**	**1.86**	**1.48**	**1.28**	**1.11**
% variance	**33.4**	**10.2**	**6.0**	**5.6**	**4.5**	**3.9**	**3.4**
Total explained % variance (cumulative)	**33.4**	**43.6**	**49.6**	**55.2**	**59.7**	**63.6**	**67.0**

NB: only loadings >.40 are given. All items adapted and abbreviated from the original Dutch questionnaire.
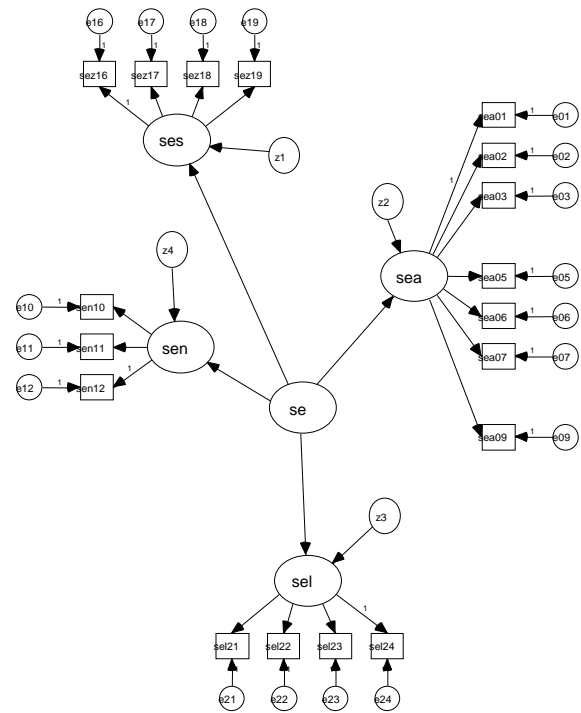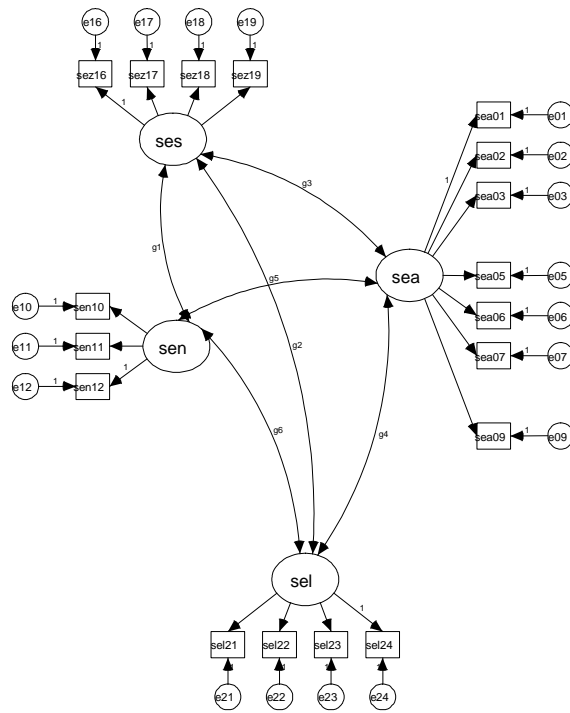
| Figure 5.1: | Figure 5.2: |

Model 1:Possible CFA structure in Amos 3.6.

CFI = .92; TLI = .91; df = 129

Model 2:Possible CFA structure in Amos 3.6.

CFI = .92; TLI = .90; df = 131

Marks for the criterion scores were taken from the Student Information System (SCIS). Both the 1A and 1B examination dates can be individually chosen by our students. Criterion scores were collected between January and August 2001. During that period, 114 students took the 1A exam (M= 7.26, SD=1.40) and 89 students took the 1B exam (M=7.13; SD=1.46).

<div align="center">**Results**</div>

Contrary to our expectations, none of the Big Five personality dimensions showed a significant relationship with study results. Of the Self efficacy scales, SEA showed a positive relationship with the first exam result. Contrary to our expectations, SEN showed a negative relationship with both the 1A and the 1B exam results. Correlations between Big Five dimensions and Se scales in general supported the construct validity of the SE scales. We will now reflect on the study outcomes in more detail, starting with hypothesis one.

Hypothesis 1: Influence of Big Five dimensions on study results

First we computed the correlation coefficients between the five personality dimensions and the results for the two exams (Table 5.3). Contrary to our expectations we found no significant relationships. A further multiple regression analysis of the two exam results on the big five personality dimensions supported the zero-order correlations: no significant ß values were found, with ß values ranging from -.14 to .07. The explained variance for the 1A exam was .05, for the 1B exam .03.

Hypothesis 2: Influence of Self-Efficacy on study results

Table 5.3 demonstrates that among our predictors, SEA and SEN are showing the best results, with correlations of respectively .28 (p< .01) and -.16, (p< .05) for the 1A exam, and with a correlation for SEN of -.18 (p< .05) with the 1B exam. The negative correlation of SEN with both exam scores is contradictory to our initial expectations. A further multiple regression analysis revealed that only SEA showed a significant ß value for the 1A exam (β = .39, p < .01) and SEN for the 1B exam (β = -.22, p < .05) (Table 5.2). As can be seen from Tables 5.2 and 5.3, the ß values of Table 5.2 do not correspond fully with the zero order correlations from Table 5.3. This is caused by the fact that the four SE scales are statistically highly dependent.

Table 5.2:

Results of the multiple regression analysis of Psychology 1A and 1B exam results on the four SE scales*

SE Predictor	Psychology 1A (n=111) Beta	Psychology 1B (n=86) Beta
1. SEA	.39***	.17
2. SEN	-.10	-.22*
3. SES	-.15	-.05
4. SEL	-.12	.11
R	.39	.28
R²	.15	.08

*SEA: Academic Self-Efficacy; SEN: Self-Efficacy for networking; SES: Self-Efficacy for self regulation; SEL: Self-Efficacy for self regulated learning. * $p < .05$; *** $p < .001$

Hypothesis 3: Relation of the Big Five dimensions and Self-Efficacy

To test our hypotheses we computed correlations between the Big Five personality variables and the four SE scales (Table 5.3). The outcomes show low but consistent (significant) correlations between the SE scales and relevant personality dimensions, supporting our hypotheses and thus the construct validity of the four SE scales. Self-Efficacy for Networking (SEN) correlated with Extraversion ($r = .23$, $p < .01$), Self-Efficacy for Self Regulated Learning (SEL) correlated with Conscientiousness ($r = .25$, $p < .01$), Self-Efficacy for Self Regulation (SES) correlated with Intellectual Autonomy ($r = .33$, $p < .01$), Self Efficacy for Self Regulated Learning (SEL) correlated with Emotional Stability ($r = .28$, $p < .01$). Furthermore, Academic Self-Efficacy (SEA) correlated with Intellectual Autonomy ($r = .22$, $p < .01$). Self efficacy for Self Regulation (SES) further correlated with Extraversion ($r = .17$, $p < .05$). The total SE scale, not surprisingly, correlated significantly with Extraversion ($r = .22$, $p < .01$) and Intellectual Autonomy ($r = .22$, $p < .01$), correlations with Conscientiousness and Emotional Stability being just below significance level. Agreeableness correlated nearly zero with the total SE scale.

With regard to further construct validation of the SE scale(s) we did a number of multiple regression analyses to find out how much of the variance of the SE scales could be explained by the five FFPI-dimensions (Table 5.4).

Table 5.4 shows that the construct of self efficacy, as operationalized in our study, cannot empirically be regressed to the five personality dimensions of the Big Five, as measured with the FFPI. There is however a significant relationship between the personality dimensions (Agreeableness excluded) and Total SE, which supports our hypothesis 3. Also, the regression results are consistent with expected relationships between personality dimensions and domain specific SE scales.

Hypothesis 4: Self Efficacy as a moderator for the relationship between Big Five dimensions and study results.

Based on a study by Gist and Mitchell (1992) we investigated the possible moderating role of the SE construct. If SE would have a separate and distinct status, it is conceivable that it would influence other behavioral relations, thus influencing the relationship between Big Five personality dimensions and study success. We computed partial correlations between the five FFPI factors and the 1A exam results, with the SE scales and SE-Total held constant. But, contrary to our expectations, we found no substantial differences between the zero order correlations and the partial correlation coefficients.

Biographical variables

As in every HEIDI study, we explored the influence of biographical variables on study success (Table 5.5). The correlation between previous education and both the results for the 1A and 1B exams is significant and of reasonable magnitude, respectively $r = .33$, $p < .01$ and $r = .26$, $p < .05$. Sex and age showed no significant relationship with study results.

Table 5.3:

Means, standard deviations and correlations of predictor and criterion variables

Variable	M	SD	n	1	2	3	4	5	6	7	8	9	10	11
1. Extraversion	.55	.88	155	_										
2. Agreeableness	2.30	1.06	155	-.09	_									
3. Conscientiousness	.47	.93	155	.02	-.08	_								
4. Emotional Stability	.89	.95	155	.03	.13	-.04	_							
5. Intellectual Autonomy	1.71	.94	155	.05	.08	-.18*	.03	_						
6. Academic SE (SEA)	37.20	4.25	158	.04	-.04	.08	.09	.22**	_					
7. Networking SE (SEN)	14.91	3.37	160	.23**	-.04	.13	.08	.05	.21**	_				
8. Self regulating SE (SES)	21.07	3.11	160	.17*	.10	-.04	.28**	.33**	.40**	.37**	_			
9. Self regul. Learning SE (SEL)	19.07	3.74	161	.15	.00	.25**	.10	.11	.38**	.41**	.38**	_		
10. SE Total	92.28	10.42	156	.22**	-.01	.16	.16	.22**	.73**	.68**	.72**	.75**	_	
11. Psychology 1A	7.26	1.39	114	-.14	-.11	-.04	-.12	-.01	.28**	-.16*	-.06	-.07	.03	_
12. Psychology 1B	7.13	1.46	89	-.10	-.06	.07	-.11	.03	.15	-.18*	-.05	.06	.02	.46**

N.B.: SE: self-efficacy expectation; * $p < .05$; ** $p < .01$

Table 5.4:

Results of the multiple regression analysis of the four SE scales on Big Five dimensions [1]

			SEA	SEN	SEZ	SEL	SETOT
Predictors					Beta's		
1.	Extraversion		.02	.22**	.14	.14	.19*
2.	Agreeableness		-.06	-.02	.06	.01	-.01
3.	Conscientiousness		.12	.14	.03	.28***	.20*
4.	Emotional Stability		.10	.08	.26***	.11	.15*
5.	Intellectual Autonomy		.25**	.06	.32***	.15	.25**
		R	.28	.28	.46	.35	.39
		R²	.08	.08	.21	.12	.15
		n	152	154	154	155	150

[1]SEA: Academic Self-Efficacy; SEN: Self-Efficacy for Networking; SES: Self-Efficacy for self regulation; SEL: Self-Efficacy for self regulated learning; SETOT= SEA + SEN + SES + SEL; * $p < .05$; ** $p < 0.1$; *** $p < .001$

Table 5.5:

Correlation matrix of biographical variables and exam results

		Psychology 1A (n=114)	Psychology 1B (n=89)
	Predictor		
1.	Sex	-.07	.14
2.	Previous education	.33**	.26*
3.	Age	-.11	.10
	** $p < .01$; * $p < .05$		

We also computed relations between study success and some additional questions. The results (Eta's) are given in Table 5.6.

Table 5.6:

Matrix of associations between background variables and exam results

	Predictor	Psychology 1A (n=114)	Psychology 1B (n=89)
		Eta	Eta
1.	Clear study goal	.25	.26
2.	Previous education: graduated?	.10	.06
3.	Dropped out of previous study?	.17	.04
4.	Positive evaluation of prev. study perf.	.46	.27

Table 5.6 shows that a positive perception of previous study performance seems to be positively associated with study success in Psychology.

The prediction of success on both the introductory courses (1A and 1B)

Instead of focusing on the (linear) relation between predictors and the specific mark for an introductory exam, we also interested in a measure on the importance of our predictors on succeeding on both the introductory exams. To accomplish this, all predictor variables were dichotomized and contrasted with the criterion 'having passed/not having passed the two exams'. We computed Chi^2 (df=1) for every dichotomized predictor and this criterion. Results of these Chi^2 analyses showed a significant relationship between 'Evaluation of previous study performance' and the criterion (χ^2 (1, $\underline{n}$ = 161) = 5.57, $\underline{p}$ = .02). We also found a significant relationship between SEL (χ^2 (1, $\underline{n}$ = 161) = 3.98, $\underline{p}$ = .05) and SE-Total (χ^2 (1, $\underline{n}$ = 156) = 4.77, $\underline{p}$ = .03) and the criterion. These results imply that SEL and the SE Total score together with 'Positive evaluation of previous study performance' are useful predictors for success on the exams of the introductory course (KMT) of Psychology.

Discussion

The aim of his study was to find out if, and in what way, Big Five personality dimensions and Self Efficacy can predict first year study results of psychology students at the OUNL. Our results showed that none of the Big Five personality dimensions predicted first year study results. With respect to SE outcomes, the results were mixed. We found some predictive power regarding exam results for Academic SE and SE for Networking (negative). SE for Self Regulated Learning and the total SE score differentiated between students who passed the two exams versus students who did not. We will now reflect in more detail on flaws and possible refinements in the two basic predictor outcomes of our study, starting with Self Efficacy.

Self efficacy and study success

Although some aspects of SE in our study turned out to predict study results in a moderate way, and the SE total scale differentiated between successful and unsuccessful first year psychology students, Bandura's claim (1982, 1986) that SE directly influences the performance level of tasks was not supported in a convincing way.

The first explanation that comes to mind is the fact that the SE estimates, made by our participants, where inaccurate. Two phenomena could be responsible for this inaccuracy: lack of information on the actual level of performance and unrealistic optimism due to low previous education.

As for the first phenomenon, various authors have pointed to the fact that the accuracy of SE as a predictor of study success increases when students have more information about their study competencies (Bandura, 1986; Wood & Locke, 1987; Vrugt, Langereis & Hoogstraten, 1997). A further refinement was made by Gilovich, Kerr and Medvec (1993) who showed that SE tends to be lowered when the moment of task execution approaches. At that moment, the task suddenly seems more difficult and students begin to realize their personal restrictions. Our participants had no feedback on their academic competencies at the moment their SE measures were taken. This could

mean that their estimates of SE were not accurate, resulting in low SE- criterion correlations.

As for the second phenomenon, about 45% of our participants had a previous educational level at or below Middle Vocational Education (MBO). A study by Agnew and Jones (1998) shows that students with low previous education tend to overestimate their academic capacities. Both phenomena mentioned above make it at least plausible that a substantial part of our participants could have shown a form of 'unrealistic optimism' at the time of scoring the SE-questionnaire on expected study performance, biasing their SE scores in the upward direction.

To overcome problems of a lack of feedback and unrealistic optimism, future SE measures should be of a repeated character. Initial SE measures at the beginning of the study should be followed by measures close to the first examination. In between these two measures, accurate feedback on study proficiency should be given. In this way, a more accurate and realistic differential SE measure for students can be composed showing possible SE change and its influence on study results.

The Big Five and study success

As noted earlier, and corresponding with previous study results and our hypothesis, conscientiousness was expected to emerge as a moderate but significant predictor of study success. To answer the question if the non-occurrence of conscientiousness as a predictor of study success in our study could be attributed to specific sample characteristics, we compared the standard deviations of the Big Five factor scores of our participants with the standard deviations of the norm group scores[12]. Comparing the standard deviations would give a more reliable picture if the factor score distributions were actually different, instead of comparing ranges and mean scores which can easily differ with a small sample used as in our study (mean scores of our participants on Extraversion and Intellectual Autonomy were significantly higher, and on Conscientiousness significantly lower than mean norm group scores). In Table 5.7, the

---

[12] For the exact calculations, see p. 50.

Big Five factor scores standard deviations of the study sample and norm group are compared.

Table 5.7:

Significance of differences between sample and norm group Big Five dimensions standard deviations

Big Five dimension	$(SD_{ng}-SD_x)/ Se_{(SDng-SDx)}$	P-value
Extraversion	.92	.18
Agreeableness	-.32	.37
Conscientiousness	.54	.29
Emotional Stability	.38	.35
Autonomy	.46	.32

The results in Table 5.7 show that no significant differences in distribution occurred between our study sample Big Five profile and the norm group Big Five profile.

Another possible explanation of relevant Big Five factors not showing any influence on study results could be the fact that the Big Five are too general in character to bring to light individual differences in highly specific multiple choice measures. Some authors point to just this need for more specific situationally relevant behavioral facets in explaining academic task differences (Mershon & Gorsuch, 1988; Rothstein, Paunonen, Rush & King, 1994; Paunonen, 1998; Paunonen & Ashton, 2001; Harackiewicz et al., 2002), but others (Ones & Viswesvaran, 1996) showed the superiority of general and broad personality dimensions. According to Paunonen and Jackson (2000) specific behavioral facets could explain substantial criterion related variance, variance that is otherwise 'hidden' in the amalgam of the original broad Big Five dimensions. Future studies of personality and study success should therefore look at the general factor level and the more specific facet level as well.

A third possible explanation for not finding any relevant correlations between Big Five dimensions and study success lies in the specific nature of the predictor and criterion measures. The criterion consisted of two single measures of maximal behavior, whereas the predictor (Big Five) measures typical behavior (see chapter 2). This predictor-

criterion dissimilarity could be of influence in not detecting any common variance. It would be of interest if a more complex criterion (more and different study measures over a longer period of time) would show a more robust correlation with personality dimensions. This would mean that if personality makes a difference for study results, these differences will occur only over a longer period of time, including multiple and different criterion measures.

The specific nature of our criterion measures (MC) makes it furthermore plausible why no correlation with Intellectual Autonomy was found, because our criterion measures leave little or no room for any autonomous behavior. This too suggests the need for more and different criterion measures over a longer period of time.

SE and the Big Five

There are relationships between our operationalization of the different SE constructs and the Big Five dimensions. Correlations between Big Five dimensions and SE scales seem to support the construct validity of the four SE scales (Table 5.3), an image further refined by the multiple regression analysis (Table 5.4). But these relationships are not strong enough to draw a definitive conclusion on the status of SE as a distinct construct of human behavior. The SE construct seems positively related to all Big Five dimensions, with the exception of Agreeableness. The relation with Agreeableness seems to point in a negative direction, which is unsurprising if one realizes the 'preoccupation' of the SE construct with individual performance. More studies are needed on this issue, specifically for measures correlating specific facets of Extraversion, Conscientiousness and Intellectual Autonomy with SE.

General limitations of our study

General limitations of our study reside in a more profound influence of restriction of range. Of all the first year students of 1999, only 28% participated in our study. Of this group, 71% did the psychology 1A exam and 55% the Psychology 1B exam. The means of these criterion measures are relatively high and the SD's are relatively low. Both exam

scores were measured on a ten point scale, with a mean score on the 1A exam of 7.26 (SD=1.4) and on the 1B exam of 7.13 (SD=1.46). So there seems to be restriction of range in our criterion scores, probably caused by the fact that our students only do an exam when they feel certain to succeed. Together with the signaled possible restriction of range in the SE and FFPI scores, this makes it hard to generate substantial correlations.

Furthermore, our sample was not fully representative, showing an overrepresentation of females. Females made up 86% of our sample, whereas for the total population this was 77%. This could be an additional factor for the meager results of SE as an academic predictor, because SE influences on study performance for men are more profound than for women (Bandura, 1997).

A final comment on the SE questionnaire is that both the PCA and the CFA were done on the same sample. To find out if the assumed factor structure of the CFA is valid, cross validation on a new sample is necessary.

Consequences for HEIDI

Because personality does not seem to matter for first year study results, the simple conclusion would be not to include a personality questionnaire in a web based diagnostic tool like HEIDI. But taking into account the limitations of our study, just such an inclusion is necessitated for doing extended research on facet scores and longitudinal outcomes. It is possible that personality does make a difference in the long run, especially with an eye to the required professional profile of practicing psychologists. Being able to give students advice on important personality characteristics, such as conscientiousness, openness and agreeableness could improve the professional level of graduates. So we do advice to include a personality measure in the diagnostic tool, not so much for advising on study success during the first year, but with a long term perspective on graduating and working as a professional psychologist.

As for Self efficacy, the CFA resulted in an acceptable model fit for our instrument, showing the model with a central SE factor at the root to be the more parsimonious. Furthermore, SE subscales and the SE total scale showed predictive value

for first year study results. Because of the high scale intercorrelations our advice would be to include the total SE scale in the web based instrument. As many of our students lack a 'normal' level of previous education needed for an academic study, any aspect that could facilitate their academic study proficiency should be put to use. Self efficacy, as a personality characteristic to persevere on difficult and long lasting tasks, could compensate for a relatively low information processing capacity. This outcome however, should be determined by empirical research, with an emphasis on a possible overlap with Big Five dimensions. It could be very well the case that, as our study results suggests, the construct of SE is a hybrid of 'being on the positive side of the Big Five', Agreeableness excepted. From a viewpoint of parsimony and compactness of our web tool this would make an independent SE measure superfluous.

CHAPTER 6

General discussion

6. General discussion

The aim of this project was to find relevant basic predictors for first year study success in psychology at the OUNL. With these predictors, a website (HEIDI) could be created where potential OUNL psychology students could get feedback on their chances to pass the first two examinations. In this way they would be able to make a more thorough decision to start the study of psychology at the OUNL or not. At this moment, more than 50 percent of the first year students drop out of the psychology program within three months, suggesting a choice of study based on erroneous information or expectations.

Taking the model of Schouwenburg (1998) as a starting point, four separate studies were conducted, all with one central question: '*What basic factors do differentiate between first year psychology students' results on two introductory exams Psychology*?' Based on the model of Schouwenburg and relevant previous study outcomes, we chose to investigate five distinct fields of possible relevant predictors for study success:

- Verbal, numerical and practical intelligence (chapter 2)

- Big Five Personality dimensions (chapters 3 and 5)

- Biodata (chapter 3)

- Prior knowledge and realistic expectations (chapter 4)

- Self efficacy (chapter 5)

We now summarize the most important findings of our studies in Table 6.1. In this table we distinguish between three separate stages in the process of passing the two first year psychology exams:

-Phase 1: factors relating to interest in studying academic psychology

-Phase 2: factors relating to (not) doing the first (and second) exam and succeeding or not[13]

-Phase 3: factors relating to the actual exam results

---

[13] Although there is a difference between doing exams and succeeding on exams, from our studies the mean success ratio of students doing exams was .85 for the 1A exam (range 79.1-93.1) and .88 for the 1B exam (range 80.5-93.3). If we combine both exams, taking an exam results in 87% of the cases in passing that exam. Therefore these two different aspects (taking and succeeding) were taken as one category. Students only do an exam when they are certain to succeed; a strategy that works out well, but which also has serious consequences for the drop out percentage.

This tripartite outlook above is made to generate useful information on the process of advancement versus dropping out and the role played by various factors during this process. If we would look only at the performance of students on the two introductory exams, we would already have lost sight of about 50% of the students who dropped out, taking no exam at all. To enlarge the success ratio of the first year it is important to find out at what moment students drop out, what factors are responsible, and what eventually can be done about this phase-bound drop out phenomenon. As we shall see, different factors are responsible for enrolling and dropping out at various phases.

For each phase, relevant positive, neutral and negative factors, resulting from our studies, are identified. We will end with a comment on the results, with suggestions for further research and implications for the construction of the HEIDI website.

Phase 1

As Table 6.1 shows, the process of starting a study in psychology, followed by exam preparation, doing the exam and the exam outcome shows the influence of different factors or predictors. In phase 1, the decision to enroll for the study of psychology, we see that sex factors play a major role. Women make up for about 75% of the (new) students of psychology at the OUNL. It is tempting to attribute special personality characteristics to this enrollment group, with relatively high extraversion, high autonomy and lower conscientiousness scores. But based upon the comparison of the standard deviations, our (two) samples Big Five profile did not differ from a norm group profile. A preliminary study on the Big Five profile of psychology students (Hofstee & Hendriks, 1998), based on self and other ratings, supports (only) the relatively high score on Autonomy. The fact that our respondents score relatively high on Emotional stability, compared to the Hofstee and Hendriks outcomes is probably due to differences in age. Normally, women tend to score lower on Emotional Stability than men. The fact that (female) students at the OUNL are much older than students at a regular university could explain this relatively high score on Emotional Stability. Further research must show if the choice for the study of psychology is more a matter of sex, or that there exists something like a distinct 'psychology personality profile'.

101

Table 6.1:

Possible relevant positive, neutral and negative predictors for three phases of study advancement for first year psychology students at the OUNL, based on significance of ß coefficients.

	Phases of study advancement		
	Phase 1: interest in Psychology: enrollment*	**Phase 2: doing exams: succeeding or not?**	**Phase 3: Exam result**
Predictors			
Sex (female)	+		
Previous education			+
Verbal intelligence		+	+
Practical study intelligence		?	?
Autonomy	+		
Spent study time		+	
Positive evaluation previous study		+	
Need for flexible studying			+
Prior knowledge		?	?
Need for self knowledge			-
Total Self Efficacy scale		+	
Academic Self Efficacy			+

*Results for this phase are inferred from general results of studies in chapters 2-5.

Phase 2

In phase 2, defined by the decision whether or not to take the first exam and thus to succeed or not, personality dimensions and sex run out of sight. The most important variables at stake in phase 2 seem 'Spent study time' and 'Verbal intelligence', denoting that verbal intelligence, together with having enough study time and working hard form the main route to success. Also, a positive self evaluation based on previous study results and a sufficient amount of self efficacy play a role in this phase. In a way, personality could still play a role in phase 2, as 'hours spent on study' seems to refer to the conscientiousness domain, although this measure was inferred from the question 'how many hours do you spend on your study weekly?' and not from the personality

questionnaire. But a situational explanation (having enough study time) can also be given. The already noted differences in predictor and criterion measurements could explain this 'tacit' influence of personality dimensions, pointing once more to the need of longitudinal and multiple criterion measurements to unravel the role of personality on study results. Other studies (Okun & Finch, 1998) showed a moderate but significant influence of personality variables on study success. It is most likely that our two single multiple choice criterion measurements do leave little room for variance to be explained by 'typical' and general behavior predictors like the FFPI we used. The general picture of our two studies on the influence of personality dimensions suggests that personality does not play any significant role in the two first year exam outcomes.

Phase 3

Phase 3, the actual exam results, shows the emergence of straight and robust predictors like previous education and verbal intelligence. In fact, these two predictors, probably correlated around .40, were the most solid across all our results. Other predictors did not add any significant explained variance in exam results. The role of academic self efficacy and the need for flexible studying in this phase are thus very small. The actual performance on the two first year exams seems to be determined by the capacity to deal with verbal complexity, with numerical intelligence playing a role on the first exam outcome. This means that a latent g score is probably the best predictor for straight exam success, in practice best represented by the (manifest) level of previous education. The fact that intelligence, or previous education, is the most robust predictor is not very surprising. There is a broad empirical consensus that correlation of general intelligence and study success lies around .50 (Neisser et al., 1996). Our study showed that being able to handle a large amount of complex verbal information is probably the decisive factor for passing the two first year exams. The implication of this finding is that an indication of one's level of cognitive capacity is a useful and even necessary piece of information in deciding to start the study of psychology at the OUNL.

Comparison with other studies

The main outcomes of our four studies are in line with findings by Harackiewicz et al. (2002). When commenting on predicting success in college they state that '…..two variables stand out as key predictors, students ability, typically measured in terms of SAT or ACT scores, and prior academic performance, typically assessed using high school GPA or high school graduation rank.' (p.562). Overlooking our three phase model of first year study success in psychology, we can conclude that the more typical side of human personality can lead people to start a study in psychology and bring them to the point where maximal behavior (taking exams) is required. From that moment on, intelligence is the most important factor. One of the main reasons why intelligence emerged from our studies as an important factor explaining exam results is probably the fact that students of the OUNL are not preselected for intelligence, as is the case with students from conventional universities. Kuncel, Hezlett and Ones (2001) found true validities[14] for the Verbal and Quantitative scales of the GRE for first year graduate grade point average in the social sciences of the same magnitude as our verbal and numerical intelligence correlations with first year study success. An exception in our study was the correlation between NAT scores and the Psychology 1B exam result, which was substantially lower than the true validity of the Quantitative GRE scale. But this could be caused by restriction of range.

The conclusion therefore seems justified that people low on information processing capacities will have severe problems in passing the first year exams. Because of these (perceived) problems, they do not even attempt to take the first (and second) exam. The most important question on this issue is in what way a persistent study effort can compensate and overcome possible liabilities in information processing. Our studies showed that the amount of study time spent is an important factor in students taking exams.

Big Five personality variables however had no effect on exam outcomes. This could be caused by our small samples. Busato (1997) found small but significant correlations for conscientiousness and extraversion with the first exam result in a much larger sample. Okun and Finch (1998) also found a substantial (negative) effect for

---

[14] True validities corrected for restriction of range and criterion unreliability.

conscientiousness on institutional departure with first year psychology students. A large scale study is needed to determine the 'true' effects of personality variables like conscientiousness on first year study results. Such a study could also reveal if the amount of spent study time can compensate for weaknesses in information processing capacities and if and how 'spent study time' relates to the conscientiousness domain. Estimating the 'true' validity of typical behavior variables is especially important because it could attend students to the fact that their own efforts could make a substantial contribution to study success.

Apart from student factors (ability and motivation), success on examinations depends on the quality and quantity of instruction, institutional factors and the influence of the broader social context (Walberg, 1981). Of the instructional and institutional factors influencing drop out listed by Prins (1997), student coaching, contacts with staff and contact time emerge as specific liabilities for OUNL students. As noted before, Studying at the OUNL is based on a distance learning concept, implying only a very limited amount of contact and direct support. The possibilities for academic and social integration, important for study success (Tinto, 1987) are minimal in a distant learning environment. This means that 'success' is largely dependent on individual factors, making students vulnerable because of the limited institutional support. It should therefore be investigated if a more supportive learning environment could facilitate first year study success.

When we contrast our findings with our 'starting' factors from the adapted model from Schouwenburg (Figure 1.1), we see that for explaining first year study success in psychology at the OUNL, this model can be reduced to two basic factors: previous education and verbal intelligence. Probably sharing a lot of common variance with the latter, a high level of previous education adds experience in dealing with complexities and the demands of an educational environment. In terms of our study, this experience is likely to generate a more realistic and higher level of academic self efficacy, thus stimulating sustained study effort, taking exams and working towards a specific goal. The central role of self efficacy from the original Schouwenburg model is thus de-emphasized, making it more a peripheral phenomenon in explaining first year study success. To get a more definitive view on the role of self efficacy, a dynamic measure on

105

this subject is needed. This means a self efficacy measure at entry, regularly providing feedback on study behavior outcomes, and a self efficacy measure at the point of taking the first exam. In this way an accurate dynamic self efficacy score, fine-tuned to the exam performance can be obtained.

Age was in our studies of limited influence on study results. The population of the OUNL differs highly in age from the conventional student population. We found only one significant (negative) effect for age on the first exam result in study 4. The other studies showed mostly small, insignificant correlations between age and exam results. We take this as a positive finding, meaning that students of highly diverse age have more or less equal chances. Furthermore, if age plays a significant role, this will be probably be moderated by previous education, because older students have a lower level of previous education.

Our study showed further that personal characteristics residing under the heading of the 'understanding and helping kind', if playing any role whatsoever, emitted a negative influence on first year study success. This may seem somewhat disillusioning, but in our opinion it is a positive indicator. It confirms that a 'true' psychologist is in the first place characterized by adequate scientific knowledge and skills. 'Social' motives may play a background role, but they are rather a handicap than a facilitator in becoming a professional psychologist at the OUNL. Table 6.1 gives an overview of the model for first year study success emerging from our studies.
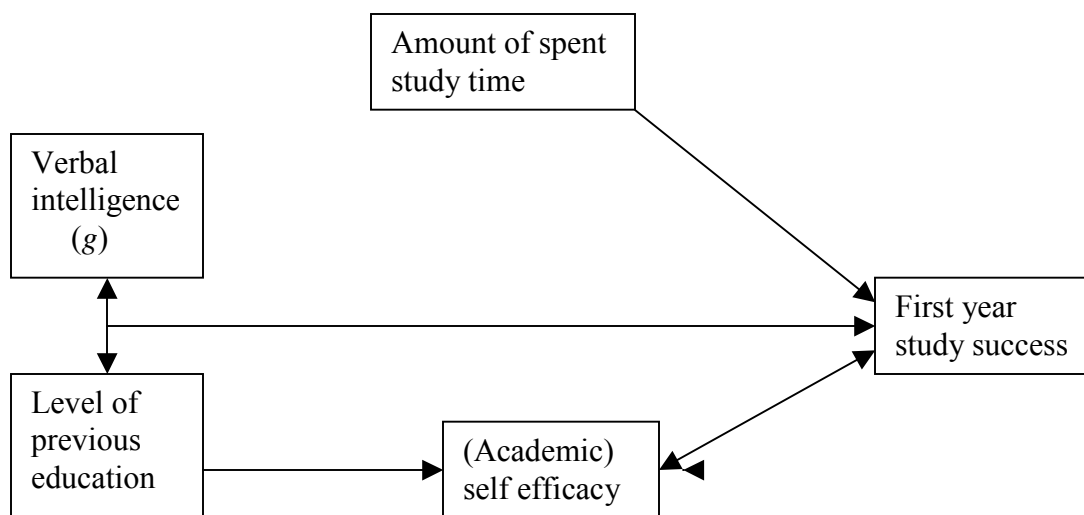

Figure 6.1: *Relevant factors for first year study success in Psychology*

Implications for the HEIDI website

If we would make a profile of a potential successful first year OUNL psychology student, this profile, based on our study outcomes, would be as follows:

Has information processing capacities at least at the level of higher general preparatory education (HAVO) *and has discipline in study behavior (invested study time).*

Thus, from the point of view of predicting success in the first year of the study of psychology, the HEIDI website should contain questions about previous education, a general intelligence test, and questions aimed at clarifying the minimal amount of needed study time. This however is a typical 'maximal' viewpoint, aimed at maximizing first year exam results. From a broader perspective, variables not directly related to study outcomes, like personality dimensions and prior knowledge, can reveal important information about what potential students will encounter during the course of their study. Feedback from these instruments can function as a Realistic Study Preview (RSP), confronting students with the merits and implications of their intended vocational choice.

So from a vocational and longitudinal perspective, additional instruments such as a PKT and a personality questionnaire could be included in the HEIDI website. By including these instruments students get two sorts of information: information on their capacities for passing the first year exams, and information about their personal study fit. Also, students should be enabled to monitor their study advancement on a systematic and regular basis, thus building towards a maximum level of confidence towards the two first year exams. The OUNL lacks the normal 'social' components (Tinto, 1987) of an academic environment. It would be wise to create some sort of a systematic feedback/feedforward mechanism, by which students at various moments of exam preparation can check if the necessary level of competence is acquired. Supplying questions to acquire this is one thing, but an extensive and clear explanation of the right

and wrong answers should be provided as well. Otherwise, an instrumental exam-driven method would prevail, while only the latter method leads to a true understanding of the (complexity of the) subject in question.

Limitations of our studies

Small number of subjects

Our studies on the role of intelligence and prior knowledge were handicapped by a small number of subjects. This had serious consequences for statistical analyses and generalizability of the results.

Sample characteristics

Although some mean Big Five scores differed from those of a representative population norm group, no significant differences in distribution were detected. A large scale study is needed to find out if psychology students match a distinct personality profile.

Dropout

A large dropout manifested itself during our studies. Criterion measurements were severely restricted by this dropout. We partly solved this problem by focusing not only on available exam results, but on creating a second criterion, labeled 'whether or not taking and succeeding on exams'.

Single criterion measurements

Our criterion measurements consisted of two multiple choice exams. Being measures of maximal performance, these measurements do not easily correlate with measurements of 'typical performance', such as the FFPI and the PSIQ. Typical performance is best measured by complex criterion measures which include some form of typical behavior. But this asks for longitudinal measures, where our focus was on the first two exam results of the study of psychology. Busato (1997) has showed the importance of the first examination result. A compensatory grade system, in which insufficient results on the first exam to a certain degree can be compensated by other exam results, could prevent students from immediately dropping out. To be more precise, in all our studies a total of

518 students participated. Of these 518, a number of 355 (69%) took the first exam within six months, of which 85 percent passed (n=303). A total of 251 students (48%) also took the second exam within one year, of whom 88 percent succeeded (n=221). This brings the success ratio of the first year students of our four studies to 43 percent. So instead of presenting more than half of the first year students with an invincible hurdle, a compensatory system could stimulate these very students to continue their study because they maintain a perspective on graduation. As from the summer curriculum of 2003 there will be only one introductory exam, this case becomes more urgent, while the workload and the complexity of the first year exam will increase.

Instruments

Not all of the instruments used in our studies showed sufficient construct validity. This is especially the case with the self constructed PSIQ. More research is needed to investigate if practical intelligence can be operationalized in a reliable and valid way, and if it plays any role in passing the first year exams. With respect to the FFPI, it is the question if such a global measure of personality covers the situation of taking the first year exams adequately. A more specific personality questionnaire, specially aimed at the educational setting, could possibly reveal variance undetected by the FFPI.

Procedure

Although we tried to keep our research procedures in line with methodological demands, we already signaled that in the study on self efficacy a different procedure probably would have led to more convincing results. A meaningful measure of self efficacy demands at least two measures of this construct, thus enabling an accurate prediction of the effective amount of self efficacy and its influence on exam results. Furthermore, in study three on biodata and personality variables, data on the second exam result were collected at a later moment than in the other studies. This may have led to a slight overestimation of exam results.

Recommendations for further research

In chapter 2, the role of verbal intelligence as a valid predictor for first year study success was demonstrated. Numerical intelligence showed only a limited contribution, although the two measures correlated at about .40. This high overlap was confirmed by the construction of a 'compact' intelligence measure of 20 numerical and verbal items. Selected on the item-total correlations, this compact intelligence test showed a much higher predictive validity for exam results than the original tests. This result however needs cross validation. Further, it is likely that the influence of numerical intelligence should increase while proceeding through the curriculum, because the student has to master methodological and statistical subjects. Thus from a longitudinal point of view a study is needed on the long-term influence of numerical and verbal intelligence on study success. Information on information processing capacities is especially useful for potential students who lack the necessary level of previous education, or who suspect a 'hidden potential'.

Chapter 3 also points to the need for further research on the role of personality dimensions and biodata in combination with longitudinal, aggregated study outcomes. The role of a tailor made study personality inventory could be investigated, focusing on specific facets of the conscientiousness domain. A possible link is an approach concentrating on mastery goals, performance approach goals and performance avoidance goals (Harackiewicz et al., 2002). Biodata research should concentrate on this role of motives and goals, possibly in combination with prior knowledge.

Chapter 4, on the role of prior knowledge, points to the need for further research on prior knowledge and longitudinal study success. Although we did not find a strong and convincing effect of prior knowledge, the regression results of the prior knowledge test for the second exam and the combined exam score showed significant ß values. As noted before, the combination of more relevant academic psychological knowledge and having a clear study goal should be the subject of further investigation. Another line of study could be the short term and/or longer term influence of relevant prior knowledge.

Chapter 5, on the role of self efficacy, points to the need for a study on the role of a dynamic self efficacy measure, fine-tuned to the task of passing the first year exams. Especially the role of feedback should be scrutinized, because in our study the lack of feedback may have distorted the self efficacy judgments in a serious way.

Overlooking the *core* problem of drop out, a simple approach could be just to ask students *why* they dropped out. This simple person oriented approach could reveal useful information not discovered by a more 'distant' psychometric approach.

As a final word, the Open University tries to attract potential students by pointing to its flexible study possibilities. It is questionable whether this advertising slogan, apart from its inherent appeal, serves educational and individual goals. Looking back on our relevant study outcomes it would be wise to emphasize the minimum amount of hours of study needed and being able to deal with verbal complexity. Also having a clear study goal, leading to an occupational activity seems more important for passing the first year's exams than being able to wander about at one's own pace. Furthermore, students with 'inside hunting' and 'social' motives' could benefit from a realistic study-demands preview because these motives only play a (very) minor role in the psychology curriculum and they seem negatively related to study advancement.

Apart from the HEIDI website we think that it would be fruitful to take a more realistic approach in student recruiting in general by the OUNL. Only in this way, a high quality educational institution will succeed in bringing a higher proportion of students to pass the first hurdles of becoming a qualified professional psychologist.

References

Ackerman, P.L. & Heggestad, E.D. (1997). Intelligence, personality and interests: evidence for overlapping traits. *Psychological Bulletin*, *101*, 213-232.

Agnew, R., & Jones, D.H. (1988). Adapting to deprivation: An examination of inflated educational expectations. *The Sociological Quarterly*, *29*, 315-337.

Anderson, R.C., & Pichert, J.W. (1978). Recall of previously unrecallable information following a shift in perspective. *Journal of Verbal Learning and Verbal Behavior*, *17*, 12.

Aiken, L.S., West, S.G., Sechrest, L. & Reno, R.R. (1990). Graduate training in statistics, methodology and measurement in psychology: a survey of PhD programs in North America. *American Psychologist*, *45*, 721-734.

Arbuckle, J.L. (1997). *Amos users guide version 3.6*. Chicago: SPSS Inc.

Asher, J.J. (1972). The biographical item: Can it be improved? *Personnel Psychology*, *25*, 251-269.

Ausubel, D.P. (1968). *Educational Psychology: A Cognitive View*. New York: Rinehart and Winston.

Baehr, M., & Williams, G.B. (1968). Predictors of sale success from factorially determined dimensions of personal background data. *Journal of Applied Psychology*, *52*, 98-103.

Baird, L.L. (1976). Biographical and educational correlates of graduate and professional school admissions test scores. *Educational and Psychological Measurement*, *36*, 415-420.

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Psychological Review*, *84*, 191-215.

Bandura, A. (1982). Self-efficacy mechanism in human agency. *American Psychologist*, *37*, 122-147.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

Bandura, A. (1995). *Manual for the construction of self-efficacy scales.* Stanford University. Stanford.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: Freeman.

Barrick, M.R., & Mount, M.K. (1991). The Big Five personality dimensions and job performance: a meta-analysis. *Personnel Psychology, 44*, 1-26.

Bentler, P.M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness of fit in the analysis of covariance structures. *Psychological Bulletin*, *88*, 588-606.

Bernstein, D.A., Roy, E.J., Srull, T.K. & Wickens, C.D. (1991). *Psychology*. Boston: Houghton Miflin Company.

Blicke, G. (1996). Personality traits, Learning strategies, and Performance. *European Journal of Personality, Vol. 10*, 337-352.

Block, J. (2001). Millenial Contrarianism: The Five Factor Approach to Personality Description 5 Years Later. *Journal of Research in Personality*, *35*, 98-107.

Bloemers, W. (1999). *Adviezen aan de poort: Higher Educational Interactive Diagnostic Inventory*. Dissertatievoorstel. Heerlen: OUNL.

Bloom, B.S. (1976). *Human characteristics and school learning*. New York: McGraw-Hill.

Boring, E.G. (1923). Intelligence as the tests test it. *New Republic*, *35*, 35-37.

Bouffard-Bouchard, T. (1990). Influence of self-efficacy on performance in a cognitive task. *Journal of Social Psychology*, *130*, 353-363.

Bouffard-Bouchard, T., Parent, S., & Larivee, S. (1991). Influence of self-efficacy on self-regulation and performance among junior and senior high-school age students. *Instructional Journal of Behavioral Development*, *14*, 153-164.

Briel, J.B., O'Neill, K., & Scheunemann, J.D. (Eds.) (1993). *GRE technical manual*. Princeton, NJ: Educational testing service

Bryson, J.B. (1969, April). The dimensions of early human experience. In W.A. Owens (Chair), *Developmental implications of biodata*. Symposium presented at the meeting of the Southern Society for Philosophy and Psychology, Miami.

Buchanan, T. & Smith, J.L. (1999). Using the Internet for psychological research. *British Journal of Psychologyy*, *90*, 125-144.

Busato, V. (1998). *Leerstijlen nader geanalyseerd*. Academisch proefschrift, Universiteit van Amsterdam.

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

Carroll, J.B. (2001). The five-factor personality model: How complete and satisfactory is it? In: Braun, H., Wiley, D & Jackson, D (eds.). *Under Construction: The role of constructs in psychological and educational measurement.* Mahwah, NJ: Erlbaum.

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

Ceci, S.J., & Liker, J. (1986). Academic and nonacademic intelligence: An experimental separation. In R.J. Sternberg & R.K. Wagner (Eds.), *Practical intelligence: Nature and origins of competence in the everyday world*. New York: Cambridge University Press.

Cervone D., & Peake, P.K. (1986). Anchoring, efficacy, and action: The influence of judgmental heuristics on self-efficacy judgments and behavior. *Journal of Personality and Social Psychology*, *50*, 492-501.

Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive Psychology*, *4*, 55-81.

Chemers, M.M., Hu, L., & Garcia, B.F. (2001). Academic self-efficacy and first-year college student performance and adjustment. *Journal of Educational Psychology, 93*, 55-64.

Chi, M.T. (1985). Interactive roles of knowledge and strategies in the development of organized sorting and recall. In S. Chipman, J. Segal & R. Glaser (Eds), *Thinking and learning skills: current research and open questions.* Hillsdale, N.J.: Erlbaum.

Chi, M.T., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. Sternberg (Ed.), *Advances in the psychology of human intelligence*.Hillsdale, N.J.: Erlbaum.

Childs, A, & Klimoski, R.J. (1986). Succesfully predicting Carreer succes: An application of the biographical inventory. *Journal of Applied Psychology*, *1*, 3-8.

Cohen, G. (1983). *The Psychology of Cognition.* London, New York: Academic Press.

Collins, J.L. (1982). *Self-efficacy and ability in achievement behavior*. Paper presented at the annual meeting of the American Educational Research Association, New York.

Cornelius, S.W. & Caspi, A. (1987). Everyday problem solving in adulthood and old age. *Psychology and Aging*, *2*, 144-153.

Costa, P.T. & McCrae, R.R. (1985). *The NEO Personality inventory manual*. Odessa, FL: Psychological Assessment resources.

Daehnert, C., & Carter, J.D. (1987). The prediction of succes in a clinical psychological graduate program. *Educational and Psychological Measurement*, *47*, 1113-1125.

Dam, K. van (1995). Het Vijffactormodel en Personeelsselectie. *Nederlands Tijdschrift voor de Psychologie*, *50*, 149-161.

Dawes, R.M. (1971). A case study of graduate admissions: Application of three principles of human decision making. *American Psychologist*, *26*, 180-188.

Denney, N.W. & Palmer, A.M. (1981). Adult age differences on traditional and practical problem-solving measures. *Journal of Gerontology*, *36*, 323-328.

Digman, J.M. (1990). Personality structure: Emergence of the five-factor model. *Annual review of psychology*, *41*, 417-440. Palo Alto, CA: Annual Reviews.

Digman, J.M. (1997). Higher-order factors of the Big Five. *Journal of Personality and Social Psychology*. *73*, 1246-1256.

Digman, J.M. & Takemoto-Chock, N.K. (1981). Factors in the natural language of personality. Re-analysis, comparison and interpretation of six major studies. *Multivariate behavioral research*, *16*, 19-170.

Dochy, F.J.R.C. (1992). *Assessment of prior knowledge as a determinant for future learning*. Heerlen: Open Universiteit Nederland.

Dochy, F.J.R.C. (1993). De invloed van voorkennis op het leerresultaat en het leerproces. In: W.Tomic en P. Span, *Onderwijspsychologie. Beïnvloeding, verloop en*

resultaten van leerprocessen. Open Universiteit Nederland. Heerlen/Utrecht: LEMMA.

Dochy, F.J.R.C. & Luyk, S.J., van (1987). *Handboek Vaardigheidsonderwijs.* Lisse: Swets & Zeitlinger.

Drenth, P.J.D & van Wieringen, P.C.W (1969). *Verbale Aanleg Testserie,* Amsterdam: Swets & Zeitlinger.

Drenth, P.J.D. & Hoolwerf, G. (1970). *Numerieke Aanleg Test.* Amsterdam: Swets & Zeitlinger.

Dyer, E.D. (1987). Can university success and first-year job performance be predicted from academic achievement, vocational interest, personality and biographical measures*? Psychological Reports*, *61*, 655-671.

Eberhardt, Bruce J., & Munchinsky, Paul M. (1982). An empirical investigation of the factor stability of Owens' Biographical Questionaire. *Journal of Applied Psychology*, *67*, 138-145.

Elshout, J.J. (1999). De Vijf Persoonlijkheidsfactoren Test (5PFT) 1973-1999. *Nederlands Tijdschrift voor de psychologie*, *54*, 195-207.

Elshout, J.J. & Veenman, M.V.J. (1990). Wat doet intelligentie voor studenten? In: Ippel, M.J. & Elshout, J.J. (red). *Training van hogere-orde denkprocessen*. Amsterdam: Swets & Zeitlinger B.V.

England, G.W. (1960). *Development and use of weighted application blanks*. Dubuque IA: Williams C. Brown.

Epstein, S. (1979). The stability of behavior: I. On predicting Most of the People Much of the Time. *Journal of Personality and Social Psychology*, *37*, 1097-1126.

Federici, L., & Schuerger, J. (1974). Prediction of success in an applied M.A. psychological program. *Educational and Psychological Measurement*, *34*, 945-952.

Feldman, K.A. (1994). Review Essay: 'What matters in College? Four Critical Years Revisited, by Alexander W. Astin. *Journal of Higher Education*, *65*, 615-622.

Friedman, M., & Rosenman, R. (1974). *Type A Behavior and Your Heart*. New York: Knopf.

Gardner, H. (1983). *Frames of mind: the theory of multiple intelligences*. New York: Basic Books.

Ghiselli, E.E. (1973). The validity of aptitude tests in personnel selection. *Personnel Psycholog*y, 26, 61-477.

Gilovich, T., Kerr, M., & Medvec, V.H. (1993). Effect of temporal perspective on subjective confidence. *Journal of Personality and Social Psychology, 64*, 552-560.

Gist, M.E., & Mitchell, T.R. (1992). Self-efficacy: A theoretical analysis of its determinants and malleability. *Academy of Management Review*, *17*, 183-211.

Glaser, R. (1984). Education and thinking. The role of knowledge. *American Psychologist*, *39*, 93-104.

Gleitman, H. (1993). *Psychology.* New York: Norton.

Gleitman, H., & Tomic, W. (1993). *Inleiding in de Psychologie.* Heerlen: Open Universiteit.

Goff, M., & Ackerman, P.L. (1992). Personality-intelligence Relations: Assessment of Typical Intellectual Engagement. *Journal of Educational Psychology*, *84*, 537-552.

Goldberg, L.R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed.), *Review of personality and social psychology*, *2,* 141-165. Beverly Hills, CA: Sage.

Goldberg, L.R. (1990). An alternative "Description of Personality": The Big-Five factor structure. *Journal of Personality and Social Psychology*, *59*, 1216-1229.

Goldberg, L.R. (1992). The development of markers of the Big-Five factor structure. *Psychological Assessment*, *4*, 26-42.

Goldsmith, D.B. (1922). The use of a Personal History Blank as a Salesmanship Test. *Journal of Applied Psychology*, *6*, 149-155.

Gordon, R.A. (1997). 'Everyday Life as an intelligence Test: Effects of Intelligence and Intelligence Context'. *Intelligence*, *24*, 203-320.

Gottfredson, L.S. (1997). Why *g* Matters: The Complexity of Everyday Life. *Intelligence*, *24*, 79-132.

Gottfredson, L.S. (1998). The General Intelligence Factor. *Scientific American Presents*, *9*, 24-29, 51.

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

Guilford, J.P. (1988). Some changes in the structure-of-intellect model. *Educational and Psychological Measurement*, *48*, 1-4.

Guion, R.M. (1966). *Personnel testing*. New York: McGraw-Hill.

Harackiewicz, J.M., Barron, K.E., Tauer, J.M. & Elliot, A.J. (2002). Predicting success in College: A longitudinal Study of Achievement Goals and Ability measures as predictors of Interest and Performance From Freshman Year Through Graduation. *Journal of Educational psychology*, *3*, 562-575.

Helmreich, R.L.,Sawin, L.L. & Carsrud, A.L. (1986). The honeymoon effect in Job performance: temporal increases in the predictive power of achievement motivation. *Journal of applied Psychology*, *71*, 185-188.

Hendriks, A.A.J. (1997). *The construction of the Five-factor Personality Inventory.* Academisch proefschrift, Rijksuniversiteit Groningen.

Hendriks, A.A.J., Hofstee, W.K.B., & Raad, B. de (1998). *Handleiding bij de Five-Factor Personality Inventory (FFPI).* Lisse: Swets Test Publishers.

Hendriks, A.A.J. & Hofstee, W.K.B. (1994). Het karakteriseren van Persoonlijkheidsschalen in termen van de Vijf-dimensionele Eigenschappenruimte. *Nederlands Tijdschrift voor de Psychologie*, *49*, 131-136.

Hilgard, E.R., Atkinson, R.T. & Smith, C.D. (2000). *Hilgard's introduction to psychology.* Fort Worth, Texas: Harcourt College Publishers.

Hofstee, W.K.B. (2001). Intelligence and personality: Do they Mix? In: Collins, J.M & Messick, S. *Intelligence and personality. Bridging the Gap in Theory and Measurement*. NJ: Lawrence Erlbaum Associates.

Hofstee, W.K.B., & Raad, B. de (1991). Persoonlijkheidsstructuur: De AB5C-taxonomie van Nederlandse eigenschapstermen. *Nederlands Tijdschrift voor de Psychologie*, *46*, 262-274.

Hofstee, W.K.B., Raad, B. de, & Goldberg, L.R. (1992). Integration of the Big Five and circumplex approaches to trait structure. *Journal of Personality and Social Psychology*, *63*, 146-163.

Hofstee, W.K.B., Ten Berge, J.M.F & Hendriks, A.A.J (1998). How to score questionnaires. *Personality and individual Differences, 25*, 897-909.

Hofstee, W.K.B. & Hendriks, A.A.J. (1998). The use of scores anchored at the scale midpoint in reporting individuals' traits. *European Journal of Personality*, *12*, 219-228.

Hogan, J. B. (1994). Empirical Keying of Background Data Measures. In: *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

Holder, R. & Wankowski, J. (1980). *Personality and Academic Performance of students at University.* Research into Higher Education Monographs. Surrey, Society for research into Higher education.

HOOP'96: *Hoger Onderwijs en Onderzoeksplan 1996*. Den Haag, SDU.

Horn, J.L. (1982). The theory of fluid and crystallized intelligence in relation to concepts of cognitive psychology and aging in adulthood. In: F.I.M. Craik & S.E. Threhub (eds.), *Aging and cognitive processes*. New York: Plenum.

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

Hunter, J.E. (1980). *Validity generalization for 12000 jobs: An application of synthetic validity and validity generalization to the General Aptitude Test Battery (GATB).* Washington DC: U.S. Department of Labor.

Hunter, J.E., & Hunter, R.F. (1984). Validity and utility of alternative predictors of job performance. *Psychological Bulletin*, *96*, 72-98.

Hurtz, G.M. & Donovan, J.J. (2001). Personality and Job Performance: The Big Five revisited. *Journal of Applied Psychology*, *85*, 869-879.

Intelligence and its measurement: A symposium. (1921). *Journal of Educational Psychology, 12*, 123-147, 195-216, 271-275.

Jensen, A.R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger/Geenwood.

Jong, T., de (1986). *Kennis en het oplossen van vakinhoudelijke problemen.* Proefschrift Technische Hogeschool Eindhoven. Helmond: Wibro.

Jong, U. de, Roeleveld, J., & Webbink, H.D. (1997). *Verder studeren in de jaren negentig.Verder studeren, een panelstudie onder scholieren en studenten: eindrapport.* Zoetermeer: Ministerie van Onderwijs, Cultuur en Wetenschappen. Den Haag: SDU.

Kavanagh, M.J., & York, D.R. (1972). Biographical Correlates of the Middle Managers' Performance. *Personnel Psychology, 25*, 319-332.

Kenagy, H.G., & Yoakum, C.S. (1925). *The selection and training of salesmen*. New York: McGraw-Hill.

Klerk, L.F.W. de (1983). *Onderwijspsychologie*. Deventer: Van Loghum Slaterus b.v.

Klimoski, R.J. (1973). A biodata analyses of career patterns in engineering. *Journal of Vocational Behavior, 3*, 103-113.

Kuncel, N.R., Hezlett, S.A. & Ones, D.S. (2001). A comprehensive Meta-Analysis of the Predictive Validity of the Graduate Record Examinations: Implications for Graduate Student Selection and Performance. *Psychological Bulletin, 127*, 162-181.

Laak, J.J.F. Ter (1995). *Psychologische Diagnostiek. Inhoudelijke en Methodologische Grondslagen.* Lisse: Swets & Zeitlinger.

Lange, A., Van de Ven, Schrieken, B., Bredeweg, B., & Emmelkamp, P.M.G. (2000). Internet-mediated, protocol driven treatment of psychological dysfunctions. *Journal of Telemedicine and Telecare, 6*, 15-21.

Lesgold, A.M., Feltovich, P.J., Glaser, R. & Wang, Y. (1981). *The acquisition of perceptual diagnostic skill in radiology*. Pittsburg, P.A.: University of Pittsburg, Learning Research and Development Center.

Levine, A.S. & Zachert, V. (1951). Use of biographical inventory in the Air Force classification program. *Journal of Applied Psychology. 35*, 241-244.

Lidz, C.S. (1998). Dynamic Assessment Approaches. In: Flanagan, D.P., Genshaft, J.L. & Harrison, P.L. *Contemporary Intellectual Assessment*. New York: The Guilford Press.

Locke, E.A., & Latham, G.P. (1990). *A Theory of Goalsetting and Task Performance.* Englewood Cliffs, N.J.: Prentice Hall.

Lodewijks, J.G.L.C. (1981). *Leerstofsequenties van conceptueel netwerk naar cognitieve structuur.* Dissertatie, Katholieke Hogeschool Tilburg. Lisse: Swets & Zeitlinger.

Loevinger, J. (1994). Has psychology lost its conscience? *Journal of personality Assessment*, *62*, 2-8.

Mael, F.A. (1991) A Conceptual Rationale for the Domain and Attributes of Biodata Items. *Personnel Psychology*, *44*, 763-792.

Manson, G.E. (1925). What can the application blank tell? *Journal of Personnel Research*, *4*, 73-99.

Matthews, G. & Deary, I.J. (1998). *Personality Traits.* Cambridge: CUP.

McCloy, R., Campbell, J. & Cudeck, R. (1994). A Confirmatory Test of a Model of Performance Determinants. *Journal of Applied Psychology, 79*, 493-505.

McDaniel, .A., Morgeson, F.P., Finnegan, E.B., Campion, M.A. & Braverman, E.P. (2001). Predicting job performance using situational judgment tests: a clarification of the literature. *Journal of Applied Psychology*, *86*, 730-740.

McDonald, R.P., & Marsh, H.W. (1990). Choosing a multivariate model: Noncentrality and goodness of fit. *Psychological Bulletin*, *107*, 247-255.

Mcnemar, Q. (1964). Lost: Our intelligence? Why? *American Psychologist*, *19*, 871-882.

Meehl, P. (1996). *Clinical Versus Statistical Prediction: A Theoretical Analysis and a Review of the Evidence.* Jason Aaronson Publishers

Melamed, T. (1992). Use of Biodata for predicting academic success over thirty years. *Psychological Reports*, *71*, 31-38.

Meijer, J. (1999). Leerpotentieel en intelligentie: voortzetting van een debat over een complexe problematiek. In Hamel, R. (red.). *Meesterschap*. Amsterdam: Vossius pers AUP.

Melamed, Tuvia (1992). Use of Biodata for predicting academic success over thirty years. *Psychological Reports*, *71*, 31-38.

Mershon, B., & Gorsuch, R.L. (1988). Number of factors in the personality sphere: does increase in factors increase predictability of real-life criteria. *Journal of Personality and Social Psychology*, *55*, 675-680.

Messick, S.M. (1967). Personality Measurement and College Performance. In: Jackson, D.N. & Messick, S.M. *Problems in Human Assessment.* New York: McGraw-Hill Book Company.

Mikesell, R.H., & Tesser, A. (1971). Life History Antecedents of Autoritarianism: A Quasi-longitudinal Approach. *Proceedings of the 74th Annual Convention of the American Psychological Association, 6,* 136-137.

Minneart, A. & Janssen, P. (1996*).* How general are the effects of domainspecific prior knowledge on study expertise as compared to general thinking skills ? In: Birenbaum, M. & Dochy, F. *Alternatives in assessment of achievements,learning processes and prior knowledge. Evaluation in education and human services.* Boston: Kluwer Academic Publishers.

Mitchell, T.W., & Klimoski, R.J. (1982). Is it rational to be empirical? A test of methods for scoring biodata. *Journal of Applied Psychology, 67,* 411-418.

Morris, P., Cheng, D., & Smith H., (1990). Preliminary report: how and why applicants choose to study psychology at university. (Unpublished manuscript, Lancaster University).

Morrison, R.F. (1994). Biodata Applications in Career Development Research and Practice. *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

Mount, K.M., Witt, L.A. & Barrick, M.R., (2000). Incremental Validity of Empirically Keyed Biodata Scales over GMA and the Five Factor Personality Constructs. *Personnel Psychology, 53,* 307-315.

Multon, K.D. Brown, S.D., & Lent, R.W., (1991). Relation of self-efficacy beliefs to academic outcomes: a meta-analytic investigation. *Journal of Counseling Psychology, 38,* 30-38.

Mumford, M.D., & Owens, W.A. (1987). Methodology review: Principles, Procedures, and Findings in the Application of Background Data Measures. *Applied Psychological Measurement, 11,* 1-31.

Mumford, M.D., Snell, A.F. & Reiter-Palmon, R. (1994). Background Data and Development: Structural Issues in the Application of Life History Measures. *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

Neider, Andrew G., Owens, Willam A. (1985). Using Biodata to predict job choice among college graduates. *Journal of Applied Psychology*, *70*, 127-136.

Neisser, U. (1976). *Cognition and reality*. San Francisco: W.H. Freeman.

Neisser, U., Boodoo, G., Bouchard, T., Wade Boykin, A., Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J. & Urbina, S. (1996). Intelligence: Knowns and Unknowns. *American Psychologist*, *51*, 77-101.

Nickels, B.J. (1994). The Nature of Biodata. *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

Nijenhuis, J. te, Evers, A., & Mur, P. (2000). Validity of the Differential Aptitude Test for the Assessment of Immigrant Children. *Educational Psychology, 20,* 99-115.

Okun, M.A, & Finch, J.F (1998). The Big Five Personality Dimensions and the Process of Institutional Departure. *Contemporary Educational Psychology*, *23*, 233-256.

Ones, D.S., & Viswesvaran, C. (1996). Bandwith-fidelity dilemma in personality measurement for personnel selection. *Journal of Organizational Behavior*, *17*, 609-626.

Owens, W.A. (1976). Background data. In M.D. Dunnette (Ed.), *Handbook of industrial psychology*. New York: Rand-McNally.

Parish, J.A., & Drucker, A.J. (1957). *Personnel research or officer Candidate School* (USA TACO Personnel Research Branch Tech. Rep. No. 117, 1-22)

Parkerson, J.A. , Lornax, R.G., Schiller, D.P. & Walberg, H.J. (1984). Exploring causal models of educational achievement. *Journal of Educational Psychology*, *76*, 638-646.

Paunonen, S.V. (1998). Hierarchical organization of personality and prediction of behavior. *Journal of Personality and Social Psychology*, *74*, 538-556.

Paunonen, S.V., & Jackson, D.N. (2000). What is beyond the Big Five? Plenty! *Journal of Personality*, *68*, 821-835.

Paunonen, S.V., & Ashton, M.C. (2001). Big Five Predictors of Academic Achievement. *Journal of Research in Personality*, *35*, 78-90.

Perugini, M. & Ercolani, A.P. (1998). Validity of the Five Factor Personality Inventory (FFPI): An Investigation in Italy. *European Journal of Psychological Assessment*, *14*, 234-248.

Pervin, L. A. (1994). A critical Analysis of Current Trait Theory. *Psychological Inquiry*, *5*, 103-113.

Pintrich, P.R., & De Groot, E.V. (1990). Motivational and Self-Regulated Learning Components of Classroom Academic Performance. *Journal of Educational Psychology*, *82*, 33-40.

Plomin, R. & Petrill, S.A. (1997). Genetics and Intelligence: What's New? *Intelligence*, *24*, 53-77.

Polanyi, M. (1969). *Personal Knowledge.* London: Routledge & Kegan Paul.

Posner, G.M. (1978). *Cognitive science: Implications for curriculum research and development.* Paper presented at the AERA conference, Toronto.

Prins, J. (1997). *Studieuitval in het wetenschappelijk onderwijs. Studentkenmerken en opleidingskenmerken als verklaring voor studieuitval.* Proefschrift Universiteit van Nijmegen. Nijmegen: Nijmegen University Press.

Quaintance, M.K. (1981). *Development of a Weighted Application Blank to Predict Managerial Assessment Center Performance.* Unpublished doctoral dissertation, George Washington University: Washington, DC.

Raad, B. de (1992). The replicability of the Big Five dimensions in three word-classes of the Dutch language. *European Journal of Personality*, *6*, 15-29.

Raad, B. de (1997). De vijf kanten van uw persoonlijkheid. *Psychologie, 16 (12)*, 14-19.

Raad, B. de, & Schouwenburg, H.C. (1996). Personality in Learning and Education: a review. *European Journal of Personality*, *10*, 303-336.

Reilly, R.R., & Chao, G.R. (1982). Validity and fairness of some alternative employee selection procedures. *Personnel Psychology*, 35, 1-62.

Rothstein, M.G., Paunonen, S.V., Rush, J.C., & King, G.A. (1994). Personality and cognitive ability predictors of performance in graduate business school. *Journal of Educational Psychology*, *86*, 516-530.

Saucier, G. & Goldberg, L.R. (1998). What is Beyond the Big Five? *Journal of Personality, 66*, 495-524.

Schmidt, F.L. & Hunter,J.E. (1993). Tacit knowledge, practical intelligence, general mental ability and job knowledge. *Current Directions in Psychological Science, 2*, 8-9.

Schmidt, F.L. & Hunter, J.E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological Bulletin, 124*, 262-274.

Schouwenburg, H.C. (1998). *Studievaardigheid: Doet de persoonlijkheid van de student er ook nog toe?* Paper gepresenteerd op de landelijke dag studievaardigheden 29 april 1998, Fontys Hogescholen: Tilburg.

Scribner, S. (1984). Studying working intelligence. In B. Rogoff & J. Lave (Eds.). *Everyday cognition: Its development in social context*. Cambridge, MA: Harvard University Press.

Silver, W.S., Mitchell, T.R., & Gist, M.E. (1995). Responses to successful and unsuccessful performance: the moderating effect of self-efficacy on the relationship between performance and attributions. *Organizational Behavior and Human Decision Processes, 62*, 286-299.

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

Spence, J.T. & Helmreich, R.I. (1983). Achievement related motives and behaviors. In; J.T. Spence (Ed.). *Achievement and achievement motives: Psychological and sociological approaches*. San Francisco: Freeman.

Sternberg, R.J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York: Cambridge University Press.

Sternberg, R. (1989). *The Triarchic Mind. A new theory of human intelligence.* Penguin Books: USA.

Sternberg, R.J. (1993). *Sternberg Triarchic Abilities Test*. Unpublished test.

Sternberg, R.J. (1995). Theory and measurement of tacit knowledge as a part of practical intelligence. *Zeitschrift für Psychologie, 203,* 319-333.

Sternberg, R.J. (1996). *Successfull Intelligence*. New York: Simon & Schuster.

Sternberg, R.J. (1999). Personal e-mail.

Sternberg, R.J., Conway, R., Ketron, J. & Bernstein, M. (1981). People's conceptions of intelligence. *Journal of Personality and Social Psychology, 41*, 37-55.

Sternberg, R.J. & Detterman, D.K. (eds.). (1986). *What is intelligence? Contemporary viewpoints on its nature and definition*. Norwood, NJ: Ablex.

Sternberg, R.J. & Wagner, R.K. (1993). The *g*-ocentric view of intelligence and job performance is wrong. *Current Directions in Psychological Science*, *2*, 1-5.

Sternberg, R.J., Wagner, R.K. & Okagaki, L. (1993). Practical Intelligence: The Nature and Role of Tacit Knowledge in Work and at School. In Reese, H and Pucket, J. (Eds.). *Advances in Lifespan Development* (205-227). Hillsdale, NJ, Earlbaum.

Sternberg, R.J., Wagner, R.K., Williams, W.M. & Horvath, J.A. (1995). Testing common sense. *American Psychologist, 50,* 912-927.

Sternberg, R.J., Forsythe, G.B., Hedlund, J. et al. (2000). *Practical Intelligence in Everyday Life*. New York: Cambridge University Press.

Stokes, G.S., Mumford, M.D. & Owens, W.A. (ed.) (1994). *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

Taylor, S.E., Locke, E.A., Lee, C., & Gist, M.E. (1984). Type A behavior and faculty research productivity: What are the mechanisms? *Organizational Behavior and Human Decision Processes, 34*, 402-418.

Teeuw, B., Schwarzer, R., & Jerusalem, M. (1994*). Dutch adaption of the general Self-Efficacy Scale.* Berlin.

Tellegen, A. (1993). Folk concepts and psychological concepts of personality and personality disorder. *Psychological Inquiry*, 122-130.

Tenopyr, M.L. Big Five, Structural Modeling, and Item Response Theory (1994). *Biodata Handbook: Theory, Research, and Use of Biographical Information in Selection and Performance Prediction.* Palo Alto, California: CPP Books.

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

Tinto, V. (1987). *Leaving College*. Chicago, Il.: University of Chicago Press.

Tobias, S. (1994). Interest, prior knowledge, and learning. *Review of Educational Research, 64,* 37-54.

Tomic, W. (1989). Teaching behavior in Dutch Mathematics Classrooms. *Journal of classroom Interaction, 24,* 5-13.

Tomic, W. & Van der Molen, H.T. (1997). *Intelligentie en Sociale Competentie*. Lisse: Swets & Zeitlinger/Open Universiteit Nederland.

Tucker, L.R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood analysis. *Psychometrika, 38*, 1-10.

Vermunt, J.H.D.M. & van Rijswijk, F.A.W.M. (1990). *Inventaris Leerstijlen.* Heerlen/Tilburg: Open Universiteit/Katholieke Universiteit Brabant.

Vitelis, M.S. (1932). *Industrial Psychology*. New York: Norton.

Voss, J.F., Greene, T.R., Post, T.A., & Penner, B.C. (1983). Problem-solving skill in the social sciences. In: G.H. Bower (Ed.). *The psychology of learning and motivation: Advances in research and theory, 17,* 165-213. New York: Academic Press.

Vrugt, A., Langereis, M.P., & Hoogstraten, J. (1993). Perceived self-efficacy, perceived learnability of capabilities and academic performance. In Koppen, J.K., & Webler, W. (Eds.). *Strategies for increasing access and performance in higher education*. Amsterdam: Thesus publishers.

Vrugt, A., Langereis, M.P., & Hoogstraten, J. (1997). Academic self-efficacy and malleability of relevant capabilities as predictors of exam performance. *Journal of Experimental Education*, *66*, 61-74.

Vrugt, A., Hoogstraten, J., & Oort F.J. (1998). Doeloriëntaties, Waargenomen Eigen Competentie en Studieresultaten. *Tijdschrift voor Onderwijsresearch, 23 (3)*, 210-223.

Wagner, K.W. & Sternberg, R.J. (1991). *Tacit Knowledge Inventory for Managers*. San Antonio: Psychological Corporation.

Wickens, C. D. (1992). *Engineering psychology and human performance.* New York: Harper Collins.

Wilbrink, B., De Jong, U. & Voorthuis, M.E. (1993). *No-Show en Low-Show in het wetenschappelijk onderwijs.* Amsterdam, SCO-Kohnstamm Instituut.

Wilkinson, L. (1999). Task Force on Statistical Inference. Statistical methods in psychology journals: guidelines and explanations. *American Psychologist*, *54*, 594-604.

Williams, W.E. (1961, April). *Life History Antecedents of Volunteers versus Non-volunteers for an AFROTC Program.* Paper presented at the meeting of the Midwestern Psychological Association: Chicago.

Williams, W.M., Blythe, T., White, N., Li, J., Sternberg, R.J. & Gardner, H.I. (1996). *Practical intelligence for school*. New York: Harper Collins.

Wolfe, R.N., & Johnson, S.D. (1995). Personality as a predictor of college performance. *Educational and Psychological Measurement, 55,* 177-185.

Wood, R.E., & Locke, E.A. (1987). The relation of self-efficacy and grade goals to academic performance. *Educational and Psychological Measurement, 47,* 1013-1024.

Wood, J.L. & Wong, A.C. (1992). GRE scores and graduate school success. *International News and Notes*.

Zelfstudie Psychologie OUNL (2000). Heerlen, OUNL.

Summary

Summary

The study of psychology at the Open University of the Netherlands (OUNL) is characterized by a high drop out rate. About 50% of the first year students drop out within three months. To answer the question 'What causes this large drop out?' the project Higher Education Interactive Diagnostic Inventory (HEIDI) was started. The aim of this project was to identify factors responsible for the first year drop out, and to create a website were potential psychology students could get advice on their chances of passing the two first year exams by receiving feedback on their scores on relevant predictors. Furthermore, taking the instruments included in the website will give potential students a realistic picture of what the study of Psychology at the OUNL actually consists of. The website will have in fact the function of a Realistic Study Preview (RSP). With this feedback, potential students can make a more thorough and realistic choice whether the study of psychology is an undertaking that fits their capacities, skills and expectancies. As a starting point for our studies, we took the model of Schouwenburg (1998). Schouwenburg distinguishes four global factors responsible for study success: Prior knowledge, study know how, self regulation and competency feelings. The factor 'competency feelings' is at the centre of his model. We chose to expand and specify the model of Schouwenburg into (possible) independent and specific predictors for study success, so that each predictor could be adequately operationalized and implemented by means of a website. Four studies were undertaken to identify relevant predictors for the two first year exams of the study of Psychology:

(1) A study concerning verbal, numerical and practical study intelligence.

(2) A study concerning biodata and personality characteristics

(3) A study concerning prior knowledge and study expectations

(4) A study concerning self efficacy and personality characteristics.

In all studies criterion measures consisted of the two first year exam results. These two exams are 40 item multiple choice tests covering an introduction in psychology. A specific characteristic of these exams is that they are computer generated from a databank, and that students are free to set their exam dates to their own choice.

Study 1

      In the first study (chapter 2) we investigated the relevance of three forms of intelligence for predicting first year academic study success in psychology: numerical, verbal and practical study intelligence. For measuring numerical and verbal intelligence, existing subtests from popular test batteries were used (NAT 70' number series and VAT '69 analogies; Drenth & Hoolwerf, 1969,1970). For measuring practical study intelligence a new instrument was created, a 41 item Practical Study Intelligence Questionnaire (PSIQ), based on guidelines from Sternberg (1970). Criterion measures consisted of the two introductory exams in psychology, both 40 item multiple choice exams covering an introduction in Psychology. Participants were 94 freshmen from the OUNL Psychology study, registered in September 2001.

      Results showed that verbal intelligence has the highest predictive validity for the two exam results, followed by numerical intelligence which had predictive validity for the first exam result. PSIQ scores were difficult to interpret, due to the small number of subjects who completed this questionnaire. Verbal intelligence also differentiated between students who actually took and succeeded on the exams, versus students who did not take or failed exams. Results are being discussed from the perspective of a composite *g* measure as a solid indicator for first year study results, as are the implications of our findings for the construction of the HEIDI website.

Study 2

In the second study, we investigated the importance of biographic variables (biodata) and Big Five personality dimensions for first year study success in Psychology at the OUNL.

      Biodata were measured with a 45 item biodata questionnaire (BQ). Personality dimensions were measured with the FFPI, a 100 item Big Five personality Questionnaire (Hendriks, 1997). As criterion measures we used the results for the first and second exam, both 40 item multiple choice exams covering an introduction in psychology. Participants consisted of 190 psychology freshmen, registered in December 2001.

      Results showed that the items 'Previous education' and 'Distance learning facilitates flexible studying' had significant positive correlations and ß values for the

exam results. Two scales from the BQ, extracted by means of factor analysis, had a negative predictive value for the first exam result: a 4 item scale named 'Economic/societal advantage of studying psychology' and a 2 item scale labeled 'Easy access'. These two scales showed however no significant ß values for exam results. When a second criterion 'succeeded versus not succeeded on exams' was used, the bio items 'Hours spent on study' and 'Number of intended courses to complete on a yearly basis' had significant positive ß values.

Of the five broad personality dimensions measured by the FFPI, no dimension had any predictive value for the exam results. There was a tendency for Agreeableness and Emotional Stability to correlate negatively with exam results. Further analysis showed that Agreeableness had a significant (negative) correlation with the criterion 'Succeeded versus not succeeded on the first exam', but this effect did not hold up in a regression analysis. The fact that none of the Big Five dimensions correlated with the exam results was probably caused by the fact that the two criterion measurements contain too little 'typical' behavior. Another question is if psychology students are characterized by a specific personality profile.

Implications of our findings are discussed within the light of what kind of predictor should be included in the website to serve the interests of potential psychology students.

Study 3
In the third study, on the importance of realistic study expectations and prior knowledge for first year study success in psychology at the OUNL, two instruments were constructed: (1) A realistic study expectations questionnaire (RSEQ), containing 14 items and consisting of the scales study contents, study benefits (knowledge and skills), study skills and time spent. (2) A prior knowledge test (PKT) containing 40 multiple choice items. Both instruments were pretested. The final versions, containing respectively 12 and 35 items showed acceptable reliability. Participants were 118 first year psychology students, registered in September 2000. As criterion measurements were taken the two introductory exams on an introduction in Psychology, both consisting of 40 multiple choice items.

Result showed a small effect for the PKT on the second exam result, and a small (negative) effect of the RSEQ scale 'self knowledge' on the first exam result. Age had a significant (negative) effect on the first exam result and on the combined result of both exams. Further analysis showed that both instruments did not differentiate between the groups of students who had taken none, one or two exams.

Possible causes why the PKT and the RSEQ showed only limited predictive power for the two introductory exams are being discussed, as well as the possible use of both instruments for inclusion in a website aimed at giving potential psychology students advice on their chances of graduating in psychology. It is concluded that a PKT could serve two functions: (1) A realistic study preview function, giving students insight in what to expect during the first year of study, and (2), with sufficient predictive validity, informing students on their chances to pass the introductory exams.

Study 4

In the fourth study we investigated the effects of Big Five personality variables and Self Efficacy (SE) on the two first year exam results of Psychology students of the OUNL. Personality dimensions were measured with the FFPI, a 100 item Big Five personality Questionnaire (Hendriks, 1997). For measuring SE we initially constructed a 33 item self SE scale, consisting of the sub scales 'SE expectations towards academic performance', 'SE expectations for utilizing others', SE expectations for self regulation' and SE expectations towards self regulated learning'. A Confirmatory Factor Analysis (CFA) showed an acceptable fit for two models, of which we choose the most parsimonious one with one central SE factor at the root. The number of items of our definitive SE scale resuming from the CFA analysis was 18. As criterion measures we used the first and the second Psychology exam, both 40 item multiple choice exams covering an introduction in psychology. Participants were 161 psychology freshmen, registered in September 2000.

Results showed no effect of the Big Five personality variables, neither for the two exam results nor for the criterion 'Passed versus not passed first and second exam'. The SE subscale Academic SE had a significant positive correlation with the first exam result, substantiated in a multiple regression analysis. The subscale networking SE had a

133

significant negative correlation with both the first and the second exam results, of which only the latter was substantiated in a multiple regression analysis. On the criterion 'Passed versus not passed first and second exam' the SE total scale, the SE subscale 'Self efficacy for Self regulated learning' and the background variable 'Evaluation of previous study performance' showed a significant effect. Intercorrelations between Big Five dimensions and SE sub scales in general supported the construct validity of the SE subscales. Besides 'agreeableness' all the Big Five dimensions correlated positively with the respective SE subscales. Self efficacy however could not be regressed to the broader Big Five subscales, with R^2's ranging from .08 for the sub scales to .15 for the SE Total scale. Neither was SE a moderator for the relationship between Big Five dimensions and exam results. Implications of our findings for the construction of a website for potential students of Psychology are discussed.

Conclusions

Putting all the results of the four studies together, the conclusion is that the first year of the study of psychology at the OUNL can be represented by a three phase model:
-Phase 1, characterized by interest in the study of psychology
-Phase 2, characterized by doing and succeeding on exams or not
-Phase 3, characterized by actual results on exams

A positive predictor for Phase 1, being attracted to the study of psychology, was sex (about 75% of the students are females). In general, these (mainly female) students manifested a personality profile not distinct from the mean population profile, with a possible exception for Autonomy. In phase 2, doing and succeeding on exams or not, positive predictors were spent study time, verbal intelligence, self efficacy and a positive evaluation of previous studies. A negative predictor in this phase was 'self knowledge'. In phase 3, the actual exam results, positive predictors were verbal intelligence, previous education, and academic self efficacy.

Contrary to our expectations, personality dimensions, apart from being related to interest in psychology, showed no predictive power for doing exams and exam results. Only 'Agreeableness' showed a tendency to be negatively related to study advancement,

134

but this result should await further empirical evidence. On the role of practical study intelligence, no definitive conclusions can be drawn from our study. Self knowledge, in the meaning of improved self understanding, proved to be a negative predictor. Students high on motives of wanting to understand other people showed decreased chances of passing the first exam.

Consequences for HEIDI

The overall conclusion is that when giving potential students feedback on their chances of succeeding in the first year of psychology at the OUNL, this feedback should be based on information regarding previous education, verbal intelligence, self efficacy and study time required. Furthermore, instruments used to acquire this information should be compact and easy to administer. With respect to longitudinal results, the inclusion of a compact personality questionnaire on the website could be considered as well. Apart from providing information on exam results, the HEIDI website can fulfill another important function: that of a Realistic Study Preview (RSP). If a large part of the new student population has a false picture of what the study of psychology consists of, taking the instruments on the HEIDI website can give them valuable information on what to expect during the study of psychology. From this RSP perspective, it would be advisable to include the prior knowledge test and the FFPI as well on the website. These instruments can support students, not only as exam result indicators, but through giving valuable information about what to expect during the first year of study and about the way their personality profile fits that of the psychology students in general.

Samenvatting

Samenvatting

De studie psychologie aan de Open Universiteit Nederland (OUNL) kent een hoge uitval. Van de ongeveer 1200 studenten die per jaar instromen, valt de helft binnen drie maanden af. Om een antwoord te kunnen formuleren op de vraag 'Welke factoren zijn verantwoordelijk voor deze hoge uitval?' werd het onderzoeksproject Higher Education Diagnostic Inventory (HEIDI) gestart (Bloemers, 1999). Het doel van dit onderzoeksproject is (1) om in kaart te brengen welke factoren de grote uitval in het eerste studiejaar verklaren, en (2) het realiseren van een HEIDI website, waar potentiële studenten feedback kunnen krijgen met betrekking tot hun kans het eerste studiejaar van de studie psychologie aan de OUNL met succes te doorlopen. Deze feedback zal bestaan uit de terugkoppeling van de scores van potentiële studenten op instrumenten welke valide voorspellers blijken te zijn voor eerstejaars studieresultaten. Naast informatie over de slaag-zakkans zal de HEIDI website potentiële studenten ook informatie en inzicht geven wat de studie psychologie aan de OUNL inhoudt. In die zin functioneert de website als een Realistic Study Preview (RSP), waardoor potentiële studenten een meer afgewogen keuze kunnen maken of de studie psychologie aansluit bij hun capaciteiten, vaardigheden en verwachtingen.

Als uitgangspunt voor de te onderzoeken factoren binnen het HEIDI project werd het model van Schouwenburg (1998) gekozen. Schouwenburg onderscheidt vier globale factoren voor studiesucces: voorkennis, studie know-how, zelfregulatie en competentie, waarbij de factor 'competentie' een centrale rol vervult. Er werd voor gekozen het model van Schouwenburg te verfijnen met als doel het creëren van onafhankelijke en specifieke voorspellers voor eerstejaars studiesucces. De nadruk lag op eenvoudig te operationaliseren en af te nemen meetinstrumenten, geschikt voor toepassing in een webomgeving.

Vier aparte studies werden verricht om de voorspellende waarde van relevante predictoren voor eerstejaars studiesucces in kaart te brengen:

(1) Een studie gericht op het onderzoeken van de relevantie van verbale, numerieke en praktische studie-intelligentie.

(2) Een studie gericht op het onderzoeken van de relevantie van biografische gegevens (biodata) en persoonlijkheidskenmerken.

(3) Een studie gericht op het onderzoeken van de relevantie van voorkennis en studieverwachtingen.

(4) Een studie gericht op het in kaart brengen van de relevantie van waargenomen eigen competentie (Self Efficacy) en persoonlijkheidskenmerken.

Bij alle studies bestond het criterium uit de resultaten op de twee eerstejaars tentamens *Inleiding in de Psychologie*, tentamens bestaande uit 40 multiple-choicevragen over een brede, standaard academische inleiding in de psychologie. Een specifiek kenmerk van deze tentamens is dat ze automatisch gegenereerd worden via een databank, en dat studenten zelf hun tentamenmoment kunnen kiezen.

Studie 1

In het eerste onderzoek (hoofdstuk 2) werd onderzocht wat de voorspellende waarde was van drie vormen van intelligentie voor eerstejaars studiesucces: numerieke, verbale en praktische studie-intelligentie. Numerieke en verbale intelligentie werden gemeten met bestaande (sub) tests uit een veelgebruikte batterij voor het bepalen van academische geschiktheid, respectievelijk de NAT'70 getallenreeksen (Drenth & Hoolwerf, 1970) en de VAT'69 analogieën (Drenth & Hoolwerf, 1969). Voor het meten van praktische studie-intelligentie werd een nieuwe vragenlijst geconstrueerd, de uit 41 vragen bestaande Practical Study Intelligence Questionnaire (PSIQ). De constructie van de PSIQ werd gebaseerd op richtlijnen van Sternberg (1970). Als criteriummaat werden de resultaten op de twee eerstejaarstentamens genomen, beide bestaande uit 40 multiple-choicevragen over een brede inleiding in de psychologie. Participanten waren 94 eerstejaarsstudenten psychologie aan de OUNL, geregistreerd per september 2001.

Uit de resultaten bleek dat verbale intelligentie de beste en meest consistente voorspeller vormde voor eerstejaarsstudiesucces. De PSIQ leverde ambigue resultaten, waarschijnlijk mede door het geringe aantal proefpersonen op dit instrument. Scores op de PSIQ vertoonden een lichte correlatie met verbale intelligentie. Numerieke intelligentie had voorspellende waarde voornamelijk voor het eerste tentamenresultaat. Zowel verbale als numerieke intelligentie bleken ook van waarde voor het kunnen

voorspellen of er tentamen gedaan zou worden en met welk resultaat (slagen-zakken). In de discussie worden de resultaten van het onderzoek toegelicht, met name vanuit het perspectief om tot een compacte *g* maat te komen als indicator voor eerstejaars studiesucces. Ook worden de consequenties van het onderzoek besproken voor de constructie en inhoud van de HEIDI website.

Studie 2

In de tweede studie werd het belang onderzocht van biografische kenmerken (biodata) en persoonlijkheidseigenschappen (Big Five persoonlijkheidsdimensies) voor eerstejaars studiesucces van psychologiestudenten aan de OUNL.

Biodata werden in kaart gebracht met een uit 55 vragen bestaande biodata questionnaire (BQ). Persoonlijkheidsdimensies werden gemeten met de Five Factor Personality Inventory (FFPI; Hendriks 1997), een uit 100 vragen bestaande persoonlijkheidsvragenlijst. Participanten waren 190 eerstejaars studenten, ingeschreven per december 2001.

Uit de resultaten van de BQ bleek dat de items *vooropleiding* en *afstandsonderwijs maakt flexibel studeren mogelijk* significant samenhingen met examenresultaten. De variabele *aantal uren bestede studietijd* had voorspellende waarde ten aanzien van het al dan niet tentamen doen en het resultaat (slagen-zakken). Nader onderzoek van de BQ leidde tot de constructie van twee schalen, een uit vier items bestaande schaal *economisch- maatschappelijke voordelen van psychologie studeren* en een uit twee items bestaande schaal *studeren aan de OUNL is laagdrempelig*. De voorspellende waarde van deze schalen ten aanzien van tentamenresultaten en ten aanzien van het al of niet doen van tentamens met een (on)voldoende resultaat bleek echter gering.

Van de vijf persoonlijkheidsdimensies, gemeten met de FFPI, vertoonde geen enkele dimensie een significante samenhang met tentamenresultaten. De dimensie *vriendelijkheid* vertoonde een licht negatieve samenhang met het (succesvol) afleggen van het eerste tentamen. Het feit dat geen van de Big Five dimensies samenhing met tentamenresultaten kan gedeeltelijk verklaard worden door het feit dat er een grote

discrepantie bestaat tussen deze dimensies en de twee criteriummaten. De FFPI meet *typical behavior*, terwijl de twee criteriummaten betrekking hebben op *maximal behavior*.

De consequenties van de resultaten worden besproken met in het achterhoofd de vraag welke informatie voor potentiële psychologiestudenten nuttig is om tot een redelijke inschatting te komen wat betreft hun geschiktheid voor het volbrengen van de studie psychologie aan de OUNL.

Studie 3

In de derde deelstudie stond onderzoek naar de voorspellende waarde van voorkennis en realistische studieverwachtingen van eerstejaars psychologiestudenten aan de OUNL centraal. Voor deze studie werden twee instrumenten geconstrueerd: (1) Een vragenlijst betreffende realistische studieverwachtingen (RSEQ), bestaande uit veertien vagen, verdeeld over de schalen *studie-inhoud*, *studie-opbrengsten (kennis en vaardigheden)*, *studievaardigheden* en *bestede studietijd*. (2) Een voorkennistoets (Prior Knowledge Test, PKT) bestaande uit 40 multiple-choicevragen, qua opzet analoog aan de twee eerstejaars tentamens. Ten aanzien van beide instrumenten werd een pilot studie gedaan. Dit leidde tot iets compactere eindversies, respectievelijk bestaande uit 12 en 35 vragen met acceptabele betrouwbaarheden (Cronbach's α). Als criteriummaten werden de resultaten genomen op de twee eerstejaarstentamens psychologie, beide bestaande uit 40 multiple choice vragen over een brede inleiding in de psychologie. Participanten waren 118 eerstejaars psychologiestudenten, ingeschreven per 1 september 2000.

Uit de resultaten bleek dat de PKT een kleine positieve samenhang vertoonde met het tweede tentamenresultaat. De RSEQ schaal *zelfkennis* had een klein (negatief) verband met het eerste tentamenresultaat. *Leeftijd* correleerde significant negatief met het eerste tentamenresultaat, en met een maat van de gecombineerde tentamenresultaten. Verdere analyse wees uit dat geen van beide instrumenten differentieerde tussen studenten die geen, een of twee tentamens hadden gedaan.

In de discussie wordt ingegaan op mogelijke oorzaken voor de geringe voorspellende waarde van de PKT en de RSEQ voor eerstejaars tentamenresultaten. Bij de PKT kunnen een onvoldoende dekking en een te grote moeilijkheidsgraad een rol

gespeeld hebben, met name ten aanzien van het eerste tentamen. Ten aanzien van de RSEQ speelt waarschijnlijk predictor-criterium discrepantie een rol.

Ook wordt ingegaan op het nut van de implementatie van beide instrumenten in de HEIDI website voor potentiële psychologiestudenten. Geadviseerd wordt om de PKT op te nemen in de website vanuit het oogpunt van een *Realistic Study Preview* (RSP), zodat potentiële studenten een realistisch beeld krijgen wat hen te wachten staat gedurende het eerste studiejaar. Bij voldoende predictieve validiteit zou de PKT tevens aankomend studenten kunnen informeren over hun zak-slaagkansen ten aanzien van het eerste tentamen.

Studie 4

In de vierde studie werd de voorspellende waarde onderzocht van Big Five persoonlijkheidsdimensies en Self Efficacy (SE) voor eerstejaars tentamenresultaten van psychologiestudenten aan de OUNL. Persoonlijkheidsdimensies werden gemeten met de FFPI (Hendriks, 1997), een uit 100 vragen bestaande Big Five persoonlijkheidsvragenlijst. Voor het meten van Self Efficacy werd een uit 33 items bestaande SE schaal geconstrueerd, bestaande uit de subschalen *SE verwachtingen ten aanzien van academisch presteren*, *SE verwachtingen ten aanzien van het kunnen inschakelen van anderen*, *SE verwachtingen ten aanzien van zelfregulatie* en *SE verwachtingen ten aanzien van zelfregulerend leren*. Confirmatieve factoranalyse liet een acceptabele fit zien voor twee modellen, een model met vier SE factoren en zes relaties en een model met één centrale SE factor en vier relaties. Uit spaarzaamheidoverwegingen en vanwege het feit dat de vier SE subschalen hoge intercorrelaties vertoonden werd gekozen voor het laatste model. Dit impliceerde het gebruik van een SE vragenlijst bestaande uit 18 vragen. Als criteriummaten werden de resultaten genomen op de twee eerstejaars tentamens, beide bestaande uit 40 multiple-choicevragen over een brede inleiding in de psychologie. Participanten waren 161 eerstejaarsstudenten psychologie aan de OUNL, geregistreerd in september 2000.

Uit de resultaten bleek dat de Big Five dimensies analoog aan de resultaten van studie 2 geen voorspellende waarde vertoonden; niet ten aanzien van de twee tentamenresultaten en ook niet ten aanzien van het zakken of slagen voor de twee

tentamens. De SE subschaal *Academische Self Efficacy* correleerde significant positief met het eerste tentamenresultaat. Dit resultaat hield ook stand in een multiple-regressieanalyse. De SE subschaal *kunnen inschakelen van anderen* correleerde significant negatief met beide tentamenresultaten; alleen de laatste correlatie hield stand in een multiple-regressieanalyse. Ten aanzien van het criterium *geslaagd of niet geslaagd op beide tentamens* hadden de score op de totale SE schaal, de score op de subschaal *SE verwachtingen ten aanzien van zelfregulerend leren* en de achtergrondvariabele *Evaluatie van vroegere studieprestaties* een significante invloed. Intercorrelaties tussen de Big Five dimensies en de SE subschalen ondersteunden de constructvaliditeit van de SE subschalen. Afgezien van *Mildheid* correleerden alle Big Five dimensies positief met de SE subschalen. Uit verschillende multiple-regressieanalyses bleek echter dat de variantie in de Self Efficacy (sub)schalen niet herleid kon worden tot Big Five persoonlijkheidsdimensies. De verklaarde variantie varieerde van .08 tot .15, dit laatste voor de Totale SE schaal. Ook bleek uit deze studie dat SE geen moderatorfunctie had ten aanzien van de relatie tussen Big Five persoonlijkheidsdimensies en tentamenresultaten.

Ten aanzien van de HEIDI website wordt geconcludeerd dat de SE vragenlijst hiervan deel uit zou moeten maken met het oog op het predictieve aspect. Een Big Five persoonlijkheidsvragenlijst zou studenten inzicht kunnen geven in hoeverre hun persoonlijkheidsprofiel al of niet strookt met het profiel van afgestudeerde studenten.

Conclusies

Worden alle resultaten van onze vier studies op een rij gezet, dan is de conclusie dat het eerste jaar van de studie psychologie aan de OUNL weergegeven kan worden middels drie fasen:
- fase 1: interesse in de studie psychologie aan de OUNL
- fase 2: al of niet doen van tentamens (zakken versus slagen)[15]
- fase 3: behaalde tentamenresultaten

---

[15] Aangezien bijna 90% van de studenten die tentamen doen slaagt, komt het doen van een tentamen in de praktijk bijna altijd neer op slagen (zie voetnoot 13, p. 100).

Een positieve voorspeller voor fase 1, interesse in de studie psychologie aan de OUNL, was sekse (ongeveer 75% van onze participanten was vrouw). Hoewel respondenten gemiddeld hoger scoorden op extraversie en autonomie en lager op consciëntieusheid, week hun persoonlijkheidsprofiel niet significant af van dat van een representatieve populatiesteekproef. In fase 2, het al dan niet succesvol afleggen van tentamens, waren de belangrijkste voorspellers *de bestede studietijd*, *verbale intelligentie*, Self Efficacy en *een positieve evaluatie van eerdere studies*. Een negatieve voorspeller voor fase 2 was de schaal *zelfkennis*. Wat betreft fase 3 bleken verbale intelligentie, *de hoogte van genoten opleiding* en academische Self Efficacy positieve voorspellers voor behaalde tentamenresultaten.

In tegenstelling tot onze verwachtingen bleken Big Five persoonlijkheidsdimensies niet samen te hangen met zakken of slagen op tentamens of met behaalde tentamenresultaten. De persoonlijkheidsdimensie *mildheid* vertoonde een lichte negatieve tendens tot samenhang met tentamenresultaten, maar dit verband zal verder onderzocht moeten worden. *Praktische studie-intelligentie* liet geen significante samenhang zien met de drie fases. Een laatste enigszins verrassende uitkomst was het feit dat studenten met een vergrote hang naar zelfkennis en het beter kunnen begrijpen van de motieven van anderen juist minder succesvol bleken op tentamens.

Consequenties voor HEIDI

Feedback ten aanzien van de kans om het eerste jaar van de studie psychologie aan de OUNL met succes te kunnen doorlopen zou gebaseerd moeten worden op de hoogte van de genoten vooropleiding, verbale intelligentie, Self Efficacy en benodigde studietijd. Het is verder van belang dat de instrumenten die gebruikt worden om deze informatie te verkrijgen, compact van aard en gemakkelijk af te nemen zijn. Met het oog op lange-termijnresultaten zou een compacte, op academisch onderwijs gerichte persoonlijkheidsvragenlijst eveneens deel kunnen uitmaken van de website. Naast het verschaffen van informatie over slaagkansen kan de HEIDI website ook de functie vervullen van een Realistic Study Preview (RSP). Als, zoals uit onze studies blijkt, een groot gedeelte van de populatie potentiële psychologiestudenten een verkeerd beeld heeft van wat de studie psychologie inhoudt, dan zijn deze studenten gebaat met informatie die

hen een realistisch beeld geeft wat hen te wachten staat gedurende de studie. Vanuit dit perspectief zouden een voorkennistoets en een persoonlijkheidsvragenlijst deel uit moeten maken van de HEIDI website. Niet zozeer als voorspellers voor tentamenresultaten, maar als indicatoren van wat hen te wachten staat gedurende het eerste studiejaar en hoe hun persoonlijkheidsprofiel zich verhoudt tot dat van de 'gemiddelde' succesvolle psychologiestudent.

Dankwoord

Dit proefschrift is uitdrukkelijk een groepsproduct. Ik wil de Open Universiteit Nederland bedanken voor de tijd en de middelen die zij verschafte voor het vervaardigen van dit proefschrift. Verder de studenten van het KMT die belangeloos hun tijd investeerden in het invullen van de vragenlijsten en tests.

De verschillende hoofdstukken zijn tot stand gekomen dankzij de inzet en toewijding van een aantal studenten en medewerkers van de OUNL. Voor hoofdstuk 2, de studie over Intelligentie wil ik Ab van der Neut en Peter Hegi bedanken. Voor hoofdstuk 3, over biodata en de Big Five, Marja van der Blom en Wilco Kop. Voor hoofdstuk 4 over voorkennis en studieverwachtingen, Ron van Kampen. Voor hoofdstuk 5, over Self efficacy en de Big Five, Harold Menning. Fred de Roode van het OTEC was betrokken bij alle data analyses en André Brouwers deed de confirmatieve factoranalyse in hoofdstuk 5. Frank Meurs converteerde het manuscript naar PDF-format.

Last but not least mijn twee promotoren, Henk van der Molen en Wim Hofstee voor hun niet aflatende inspiratie, nauwgezetheid en toewijding. Henk was de eeuwig draaiende motor achter het geheel met een precisie die mij soms tot wanhoop dreef. Wim Hofstee was de geïnvolveerde kaartlezer, die ons soms deed inzien dat wij een afslag hadden gemist, in een doodlopende straat dreigden te verzanden, of ons er op attent maakte dat er buiten de geplande route nog veel schoons te genieten viel. Ik verkeer nu een kleine 30 jaar in de periferie van jullie bestaan. Een psycholoog kan zich geen betere leeromgeving wensen.

Curriculum Vitae

Wim Bloemers (1955) deed in 1975 Atheneum ß examen, waarna hij in Groningen Psychologie ging studeren. Na zijn doctoraaldiploma Persoonlijkheidspsychologie (1983) studeerde hij Wijsbegeerte aan de Rijksuniversiteit Groningen en de Universiteit van Amsterdam. Hij werkte onder andere bij een psychologisch adviesbureau en bij de HES-Rotterdam als docent psychologie, docent management en trainer communicatieve vaardigheden. Sinds 1998 is hij verbonden aan de Open Universiteit Nederland als studiebegeleider. Daarnaast publiceert hij regelmatig over personeelsselectie en aanverwante zaken en is hij zelfstandig gevestigd als psycholoog. Website: www.assessmentservice.nl.

Publicaties

Bloemers, W. (1988) Psychologische tests: If you can't beat them, buy them. Hoe de testpsychologie haar eigen ruiten in laat gooien. *Psychologie en Maatschappij*, *4*, 361-375.

Bloemers, W. (1989). De markt van angst en geluk. *Intermediair*, *25*, *48*, 57-61.

Bloemers, W., (1990). *Het psychologisch onderzoek: Een oefenboek*. Baarn: AMBO.

Bloemers, W. (1996). Het selectieproces: Een overzicht voor sollicitanten en selecteurs. In: Breed, A. & Kooreman, A. (red.)*: Psychologie in Arbeid en Loopbaan*. Lisse: Swets & Co.

Bloemers, W. (1997). Nederlandse Intelligentietests. In: Tomic, W & Molen, H.T. van der. *Intelligentie en Sociale competentie*. Heerlen/Lisse: Open Universiteit / Swets & Co.

Bloemers, W., Voorlichting over intelligentietests. In: Tomic, W., & Molen, H.T. van der (1997). *Intelligentie en Sociale competentie*. Heerlen/Lisse: Open Universiteit/ Swets & Co.

Dam, K. van & Bloemers, W. (1997) (red.): *Individu, Arbeid & Organisatie* Utrecht: De Tijdstroom.

Bloemers, W. (1997). Het selectieproces: Een voorstel voor een optimaal rendement. In: Dam, K. van, & Bloemers, W., (red.): *Individu, Arbeid & Organisatie*. Utrecht: De Tijdstroom.

Bloemers, W. & Hagedoorn, E. (1998). *Management, Gedrag & Organisatie*. Maarssen: Elsevier.

Bloemers, W. (2000). Persoonlijke ontwikkeling: Het eeuwigdurende minderwaardigheids-complex van de manager. In: *Opleiding & Ontwikkeling*, 13, 12.

Bloemers, W. & Van der Maesen de Sombreff, P.M.A.W. (2001). Persoonlijke Ontwikkeling: een kwestie van intelligentie, persoonlijkheid en/of sociale vaardigheden? In: *Human Resource Development*, 4, 27-36.

Bloemers, W. (2001). *De kleine assessmentgids.* Baarn: Ambo

Boekrecensies

Bloemers, W. & Van Dellen, T.: Professionele docenten maken het Hoger Onderwijs: een uitnodiging tot zelfreflectie. Recensie van *Onderwijskunde Hoger Onderwijs* (Ten Dam et al., 1997). *De Psycholoog*, *32*, 12, 1997.

Van Dellen, T. & Bloemers, W.: Integreren van Leren en werken met EPSS, vooralsnog een illusie? Recensie van *Leren en werken met EPSS* (T.H.J.Bastiaens, 1997). *Opleiding en Ontwikkeling, tijdschrift voor Human Resource Development,* mei 1998.

Bloemers, W. Recensie *Handboek Solliciteren* (T. Barning & S.Weusten). *Intermediair*, 1998.

Bloemers, W. & Van Dellen, T.: Het testen van allochtonen is (*g*)een probleem? Recensie van J. te Nijenhuis: Comparability of test scores for immigrants and majority group members in the Netherlands. *De Psycholoog,* (1998).

Bloemers, W: Recensie van: Individual Differences and Behavior in Organizations (K.R. Murphy). In: *European Journal of Selection & Assessment* (1998).

Bloemers, W: Recensie van: Current Perspectives in Industrial/Organizational Psychology (G. Lowenberg & K. Conrad). In: *European Journal of Selection & Assessment* (2000).