

**Determinants of success:  
a longitudinal study in higher professional education**

**Rutger Kappe**

ISBN 978-90-8891-346-4

*'academic failures are easier to predict  
than successes,  
the explanation being that a lack  
is hard to remedy,  
while an ability is easy to neglect'.*

Harris, D. (1940)  
*Factors affecting college grades* (p.150)



VRIJE UNIVERSITEIT

Determinants of success: a longitudinal study in higher professional education

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan  
de Vrije Universiteit Amsterdam,  
op gezag van de rector magnificus  
prof.dr. L.M. Bouter,  
in het openbaar te verdedigen  
ten overstaan van de promotiecommissie  
van de faculteit der Psychologie en Pedagogiek  
op dinsdag 29 november 2011 om 15.45 uur  
in de aula van de universiteit,  
De Boelelaan 1105

door  
Floris Rutger Kappe  
geboren te Dordrecht

promotor: prof.dr. H. van der Flier

copromotor: dr. J.A.H.M. Onstenk

# Contents

<i>Chapter 1</i>	
Introduction	9
<i>Chapter 2</i>	
The reliability of overall, and specific, GPA	23
<i>Chapter 3</i>	
Using multiple and specific criteria to assess the predictive validity of the Big Five on academic performance	35
<i>Chapter 4</i>	
Predicting academic achievement in higher education: What's more important than being smart?	45
<i>Chapter 5</i>	
Learning styles and academic achievement	63
<i>Chapter 6</i>	
Using multiple samples to replicate findings	71
<i>Chapter 7</i>	
Academic assessment centers and student competencies	83
<i>Chapter 8</i>	
Competencies, grades, intelligence, personality, and motivation as predictors of early career success	101
<i>Chapter 9</i>	
Summary of findings and recommendations	121
Samenvatting (summary in Dutch)	135
Appendices	147
List of abbreviations	151
References	153
Dankwoord (acknowledgements in Dutch)	167
Curriculum Vitae	169
List of publications	171





# Chapter 1

## Introduction

### 1.1 Problem statement and aim of the current study

There are many reasons why a country might provide its citizens with an education. Businesses need a continuous supply of talent and highly skilled workers to compete in a global market. Individuals want to be all that they can be and to self-actualize intellectually, morally, and in their career ambitions. Additional benefits of an education include improved health (Pascarella & Terenzini, 1991). And studies demonstrate unequivocally that the economic return associated with having a good education is substantial (e.g., Thomas, 2000).

On average, countries around the world spend 5.7% of their GDP on education (OECD, 2010). Education costs in the Netherlands constitute 5.6% of GDP, of which 26% is spent on higher education (OECD, 2010). The Dutch government, along with other European Union countries in 2009, helped develop a goal to turn the European Union into a competitive knowledge based economy. This means that The Netherlands will need to improve the educational level of its population. The chapters in this dissertation examine how intelligence, personality, motivation, anxiety and learning styles are related to academic achievement and early career success. Findings from the research serve the stakeholders of science, business, institutes of higher education, and students.

The number of students attending higher education has increased 2% annually in The Netherlands (CBS, 2010). More students have been pursuing higher education degrees, however time-to-graduation has been increasing as well. New targets have therefore been set. Whereas only 77% of the students who started in 2001, and 76% of those who started in 2002, finished within six years, the Dutch government has set an on-time target of 90% for the 2013 cohort (HBO-raad, 2009a). This target will be challenging because records show that time-to-graduation increased by three months on average between 2004 and 2008 (HBO-raad, 2008).

#### *Serving scientific interests*

This study examined the effects of intelligence, personality, motivation, anxiety, and learning style on academic achievement and early career success. An integrated approach was used to establish the joint effects of the predictors because they are interrelated (De Raad & Schouwenburg, 1996; O'Conner & Paunonen, 2007). In addition to overall GPA, this study used multiple measures of academic achievement: lectures, team projects, skills acquisition, internships, and thesis. This permitted a more detailed analysis of how the

various predictors influenced academic achievement. For example, it was possible to investigate whether different personality traits such as extroversion correlated more highly with grades on team projects compared to grades on exams in a traditional lecture environment, or how well different GPA's predicted early career success, individually as well as in conjunction with other variables such as intelligence and personality.

### *Serving business interests*

Businesses want highly qualified, competent graduates, and institutes of higher education are expected to deliver that talent. Unfortunately there is more speculation than research about what makes for a successful transition from higher education to professional life and then the needs of business change constantly. Organizations have often complained that graduates had sufficient knowledge but lacked important competencies such as teamwork and presentation skills. The converse is also true; since the introduction of competence-based education, companies have felt that graduates may have the necessary competencies but now lack sufficient knowledge. It is no wonder that institutions of higher education have struggled to meet the demands of these organizations. Organizations benefit from career research because knowledge about characteristics that are related to early career success enables them to select those students who will make the largest contributions to the success of their organization.

### *Serving higher education institutions interests*

Quality criteria that are commonly used as benchmarks for higher education institutes (HEI) include the time it takes for students to graduate and the rate and level of employment for graduates. HEI's compete with one another for government funding and do so in two ways: a) by increasing the number of students who graduate and b) by taking in more students. The government is taking a closer look at the performance of HEI's, especially in terms of graduation rates. Therefore, it is important to understand the factors that influence academic achievement. In the past, a significant portion of an HEI's funding was tied to the number of students it graduated - a so called 'diploma bonus'. This has led to a concern that HEI's have been lowering their academic standards to graduate more students.

### *Serving students' interests*

All full time students, 18 and older, are eligible for student financial support in The Netherlands. Students are entitled to a basic grant, and to additional grants depending on their income and that of their parents. Student grants are considered a loan. If students meet *study progress requirements*, then the loan is forgiven. This means that students must pass all exams of the 'propedeuse' within the first two years. Subsequent loans for succeeding years, up to five years maximum, are forgiven by the government if students complete their degrees within ten years (IBG, 2008). However, as noted, the time students

take to graduate has been on the rise. The government has recently decided to increase the fee substantially for students who take more than five years to graduate from a four year program. Because only 58% of the students graduate within five years, there is much to gain by understanding what factors are related to time-to-graduation. Moreover, students need adequate information about factors that are related to academic success. Students are entitled to a good education and be able to develop their strengths and weaknesses in the fairly safe environment of an educational setting. Feedback to students about their potential can help them estimate their chances of success when pursuing a college degree and help choose useful interventions.

## 1.2 Higher education in the Netherlands

The Netherlands has a unique dual system in which high school graduates can attend either a research university or a university of applied sciences, the latter also being known as Hoger Beroeps Onderwijs (HBO - higher professional education). The 13 research universities in The Netherlands include general and specialized universities, and the Open University. The 51 HBO's include general institutions as well as institutions that specialize in a specific field such as agriculture, the fine and performing arts, and teacher training. Whereas research universities are primarily responsible for offering *research-oriented* programmes, HBO's offer programmes of higher *professional* education that prepare students for particular professions. HBO's are more *practitioner* oriented compared to research universities. In this dissertation, the terms university and college are used when referring to research universities and HBO's, respectively.

In the 2009-2010 academic year, there were 403,212 college students, of whom 104,620 were starting their first year (HBO-raad, 2010), and 231,823 research university students, of whom 42,436 were in their first year (VSNU, 2010). There has typically been an annual 2% increase in student attendance over the years, however the number of students attending higher education increased by 5% in 2009-2010. A 25% increase in student enrolment is anticipated over the next ten years, which means 800,000 students will be studying at a research university or college.

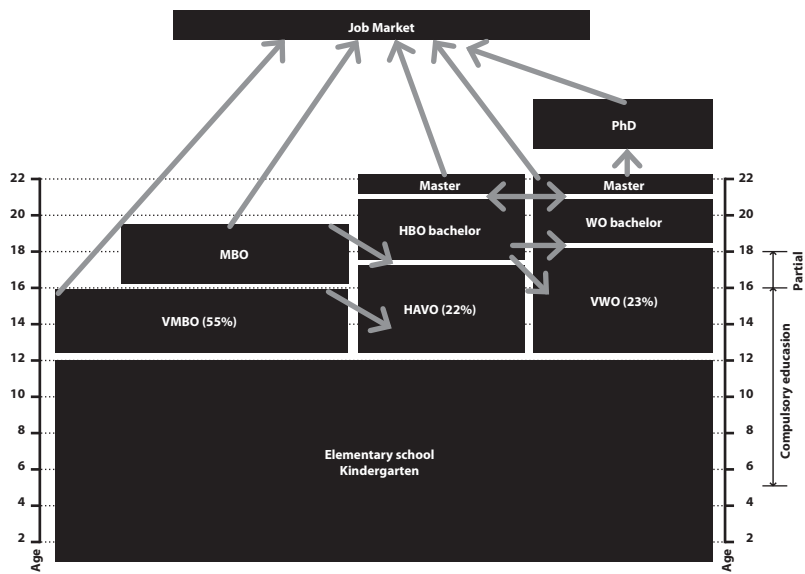


Figure 1.1 Educational system in the Netherlands

Dutch students typically start university or college when they are 17 to 18 years old. Prior to this, student abilities are assessed after primary school and students are streamed into one of several different types of secondary education. Students who have general secondary education (HAVO) diplomas or secondary vocational education (MBO) diplomas are eligible for admission to college. Students are eligible to enter university if they hold a pre-university (VWO) diploma. They may also switch to a university after receiving a first year certificate (propedeuse) from a college. Most college students (81%) pursue higher education studies full-time. The college curriculum usually requires four years of full-time study and is divided into a preliminary year of foundation courses called 'propedeuse' and three additional years called the 'main phase'. A bachelor's degree is awarded when a student completes 240 European credits (EC's), and graduates obtain degrees in specific fields such as Bachelor of Engineering or Bachelor of Human Resource Management. According to Dutch law, one credit represents 28 hours of work and 60 credits are earned during a 42 week, full-time year of study.

### **1.3 Current challenges of universities of applied sciences**

The research in this dissertation took place at a large college in the Netherlands. There were several reasons for choosing a college: enrolment at colleges is larger than at research universities, retention rates are lower, and only a few studies have focused on the academic achievement of college students. Dutch colleges currently find themselves in a fast changing environment. They face the simultaneous challenges of increased student enrolment, longer graduation times, and this during a time when the government is planning serious cuts in educational budgets. Moreover, colleges must also contend with two important educational reforms, namely, having to implement a bachelor-master structure, and introducing competence based education. The latter challenge is discussed below because it informs on the context in which the present study took place.

#### *Making competence based education work*

Over the past decade, competence-based education (CBE) has been the leading paradigm for innovation in vocational and professional education in the Netherlands as well as in other European countries (Arguelles & Gonczi, 2000; Descy & Tessaring, 2001; Van Merriënboer, Van der Klink, & Hendriks, 2002). Many programmes in secondary vocational education (MBO) are described as competence-based (Mulder, 2003). Most Dutch colleges, as well as some universities, have adopted the competence based learning approach. However, the term competence is broad, and competence-based education programs vary with respect to theoretical orientation, scope, intentions, and scientific focus (e.g., Mulder, 2004). Definitions of CBE date back to work by Grant, Elbow, Ewens, Gamson, Kohli, Neumann, Olesen, and Riesman (1979, p.6) who defined competence-

based education as, “. . . a form of education that derives a curriculum from an analysis of a prospective or actual role in modern society and that attempts to certify student progress on the basis of demonstrated performance in some or all aspects of that role”. Current characteristics of competence based education include, a) authentic, open problems and learning materials that have personal meaning for students and are presented in a variety of formats, b) teaching methods that arouse interest, activate prior knowledge, clarify meanings, and model appropriate learning strategies and reflective processes, c) small group learning that facilitates competencies such as teamwork, debating and reflection, and d) changes in the goal and use of assessments such as integrating instruction, learning, and assessment (e.g., Entwistle, 2000).

Nevertheless, there is no clear guideline as to whether a curriculum can be characterized as competence-based. There may even be quite large differences between different programs within an institution of higher education concerning how competence-based the curricula are. Most full-time programs do however have common elements: lectures, skills training, projects, internships, and a bachelor thesis during the last year of study. In CBE more time is spent on a) project work which helps students develop interpersonal competencies and teamwork skills, and b) internships help students practice and develop the necessary competencies in actual business settings. The increase in number of projects and internships has often been at the expense of the time students spent on lectures, that is the time students attend classroom based instruction by a teacher. In CBE it is expected that, although students spent less time in the classroom instruction environment, they acquire the same level of knowledge through self-study and by working on projects. Critics have argued that the emphasis in CBE on competencies has been at the expense of level of knowledge student acquire (e.g., Kayzel, 2005). Critics have also argued that CBE is nothing more than ‘old wine in new bottles’, implying that colleges only implement minor changes in their curricula, such as altering the amount of time spent on internships. This has led to large differences between colleges and within colleges between different programs. Some programs put the competence development of their students at the center of their program and integrate competencies into all courses and activities and also introduce new types and modes of assessments. Other programs may have chosen to integrate competence development into selected courses or into their student counselling program.

Competence-based learning is popular because it presumably reduces the gap between what is taught in school and what is needed on the job (Biemans, Nieuwenhuis, Poell, Mulder, & Wesselink, 2004). In addition to up-to-date knowledge, current work practice demands competencies such as the ability to work in teams, coping with uncertainty, social and communication skills, and taking calculated risks (e.g., Mulder, 2004). The underlying premise is that CBE enables students to acquire the competencies needed in their future professions, as well as in society in general. Moreover, they should

continue to develop their competencies so that they are able to react to, and anticipate, future developments in their work (Jenewein, Knauth, & Zülch, 2002).

Competence-based education originated from the behaviorist model of training and learning. During the 1970s, the 'competency movement' in the U.S. (see Friedlander, 1996; Parry, 1998) was characterised by a detailed analysis of the behavioral aspects of professional tasks at work. Tasks were dissected into component parts, resulting in long lists of fragmented behavioral elements. This approach became unfruitful and interest in the original competency movement declined by the 1980's (Mulder, 2003). Cognitive learning theories replaced the programmed instruction models of the 1970's, and in turn, were replaced by the constructivist learning theory. The social constructivist approach emphasized the social, collaborate aspects of learning. Knowledge is acquired and transferred within a context and with others, such as experts, teachers or peers. It is this social constructivist approach that once again triggered an interest in the importance of competence. According to this approach, effective learning occurs when students analyze and solve actual business problems, present and debate one's opinions, and show verbal and writing skills. The new form of competence thinking appears to be making its way back into education in the United States (U.S. Department of Education, National Center for Educational Statistics, 2002).

Unfortunately, there is a serious lack of scientific research and theory to support the claims of competence based education (Kayzel, 2005; Korthagen, 2004). From the earlier competency movement in the US that peaked during the 1970's, it is known that the risk of bureaucratisation is very real. The fate of many educational reforms is that they must cope with resistance to change, as well as with enthusiasts who take the concept a step too far. For example overemphasis on the importance of behavioral skills leads to the development of long, unworkable, behavioral checklists. The question is, whether competence-based education, now based on a more holistic approach, will succeed this time around. The reason that competence-based education is so popular in the Dutch policy arena is because of its alleged capacity to reduce the gap between education and work. There is a belief among policy-makers that competence-based graduates, compared to students with traditional educational qualifications, transition better into modern organisations.

#### *Implementing new assessment practices*

Experts agree that traditional exams are not sufficient in a competence-based curriculum (Van der Sanden, De Bruijn, & Mulder, 2003) and that new techniques to assess student learning are needed. The Dutch Educational Inspection Board (Inspectie van het Onderwijs) has described assessment as "tests or a set of tests or exams that are application directed and situated in a real live context". In addition to assessing traditional knowledge, assessing how knowledge is applied, especially in an actual business setting, is very important. CBE assessments are more integrated, complex, and often

multidisciplinary. Unlike exams, there is no one best answer to a question, but rather there are several possible ways to solve a problem. This has led to a discussion about how to 'assess the quality of assessments' (Dierick, Van de Watering, & Muijtjens, 2002).

The social sciences frequently use psychometric criteria, but educational researchers have questioned the usefulness of such criteria for educational assessments. This issue has led a group of educational researchers (e.g., Dochy & De Rijke, 1995; Segers, Dochy, & Cascallar, 2003) to develop a set of so called edumetric quality criteria that refer to metrics designed specifically for educational settings. At first glance the edumetric framework seems more extensive with up to ten criteria: authenticity, cognitive complexity, meaningfulness, fairness, transparency, consequences, directness, reproducibility, comparability, and costs. The importance of reliability (reproducibility) is downgraded because assessments are less standardized compared to traditional tests. Instead, the consequences of the assessment for learning are stressed (e.g., Hendriks & Schoonman, 2006; Van der Vleuten & Driessen, 2000). In other words: did the assessment have an effect on student learning? A reasonable question is what such a new framework adds to our understanding of assessment practices. We therefore compared the extensive validity framework of Messick (1994, 1995) with the edumetric criteria and found that the edumetric and psychometric criteria are very much related (see also Baartman, Bastiaans, & Kirschner, 2004). The most important contribution of the edumetric criteria for educational assessment practice is that assessment developers, teachers, and staff are encouraged to think about assessment from the perspective of students and how they learn.

Assessments of competencies, in contrast to traditional exams, are more labour-intensive and time-consuming. Assessments often involve interviews or structured observation rather than classroom exams, and there are two examiners instead of one. However, teachers are much more familiar with traditional exams. To prevent unproductive discussions about the merits of traditional exams and assessments, the Dutch educational council (Onderwijsraad, 2008) has stated that there is no single, best way to assess the qualities of students. All types of exams and assessment have their strengths and limitations. The council therefore recommends using a mixture of exam formats.

### **1.4 The context of the present study**

The present study was conducted with students who were preparing for a degree in human resource management (HRM) at a large Dutch college. The HRM program educates for a broad spectrum of HR jobs and graduates start their careers at large companies as junior HR advisers in personnel policy or at employment offices. The program consists of a variety of competency based learning environments such as classroom tuition, team



projects, and skills training. A large amount of time is devoted to internships in actual business settings during the second, third and fourth years, and students are required to write a thesis during the last six months of their study. Students participate in a student counselling program in which they reflect on their progress throughout the four years. They receive a book in each course that lays out the goals and relevant competencies. Most courses are assessed with traditional exam formats (multiple choice, short essay questions), as well as with project reports and products (e.g., manuals, advisory reports). Moreover, students write reports in which they reflect on what they have learned and these are also assessed. All assessments are graded according to the typical Dutch ten point rating scale (1 = very poor, 6 = pass, 10 = excellent). The HRM program contains two large scale assessment centers. Students participate in the assessment centers at the end of their first and third years and are rated on nine competencies: analyzing, creativity, advising, teamwork, communication, reporting, justification, presenting, and learning goals.

### **1.5 The individual differences approach to studying academic achievement and early career success**

Precisely why students differ in academic achievement has been under investigation for some time. There have been numerous approaches, models, and conceptual frameworks (Bean, 1980; Bean & Metzner 1985; Evans, 2000; Pascarella & Terenzini, 1991; Tinto, 1987). In general, these models divide the variables that affect academic achievement into input, process, and context variables (Huitt, 1994). Input variables are the characteristics of both students and teachers. Intelligence, prior knowledge, gender, age, motivation, self-efficacy, personality, study behavior, and learning style are examples of relevant student characteristics. Values, beliefs, knowledge, communication skills, performance skills, and personality are considered important teacher characteristics. The process variables concern the interaction between the students and the academic environment, such as the teacher-student interaction. Context variables are variables outside of the classroom that affect teachers and students. This includes family, community, peer groups, and characteristics of the school such as organisational structure, size, and school climate. This dissertation examines how the student input variables of intelligence, broad and specific personality traits, anxiety, motivational variables, and learning styles, affect academic achievement and early career success.

There are two main research approaches concerning how student input variables are related to academic achievement. The first approach has examined the direct influence of individual differences on academic achievement. Studies that use this approach have examined the relationship between personality and academic achievement and meta-analytic results have indicated a moderate relationship between personality traits and

academic achievement (e.g., Poropat, 2009). The second approach has examined process variables such as the strategies that students use when learning (e.g., Bruinsma, 2003) or has studied the behavior of students (Eggens, 2011). Figure 1.2 shows the two main approaches.

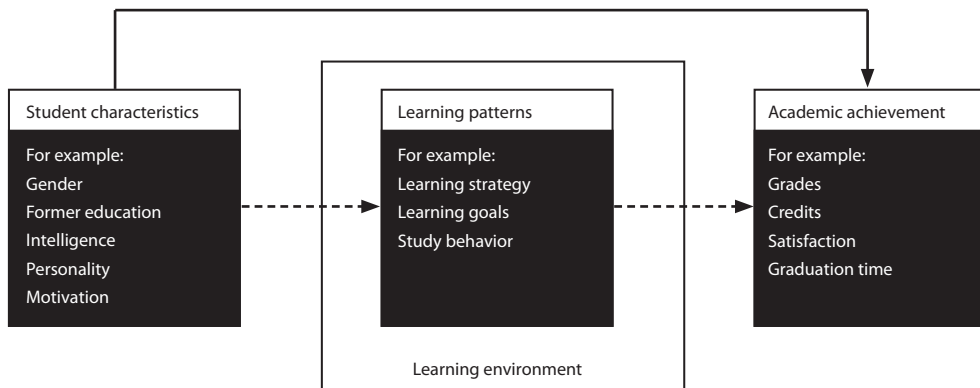


Figure 1.2 Direct (straight line) and indirect (dotted line) research approaches on studying student characteristics in relation to academic achievement

Based on a large scale study with first year undergraduates, Van Bragt (2010) proposed a model in which learning patterns were regarded as an input, rather than a mediator between individual differences and learning outcomes. The studies presented in this dissertation use the direct approach. The suggestion by Van Bragt (2010) to incorporate learning patterns as input variables is included in the working model of this dissertation. The model is presented in Figure 1.3 and lists the student characteristics (input), academic measures (mediating), and early career measures (outcome). Progress has been made in identifying specific individual differences that affect academic achievement (Riding, 2005). Gender differences (e.g., Skaalvik & Rankin, 1994), personality (e.g., De Raad & Schouwenburg, 1996; O’Conner & Paunonen, 2007), intelligence (e.g., Mackintosh, 1998), and cognitive and learning styles (e.g., Busato, Prins, Elshout, & Hamaker, 2000; Coffield, Moseley, Hall, & Ecclestone, 2004) have all been explored. However much variance in the performance of students remains to be accounted for. To advance the field further, researchers have suggested that multiple predictor variables be integrated into an overall system so that valid conclusions can be drawn concerning both individual and incremental contributions of each variable (e.g., De Raad & Schouwenburg, 1996), and that specific, instead of global, criterion measures of academic success be used (e.g., O’Conner & Paunonen, 2007). This dissertation incorporates both suggestions. Moreover, a longitudinal dataset is used to examine how intelligence, personality, anxiety,

motivation, and learning style affect academic achievement and how academic achievement, in turn, affects subsequent career success.

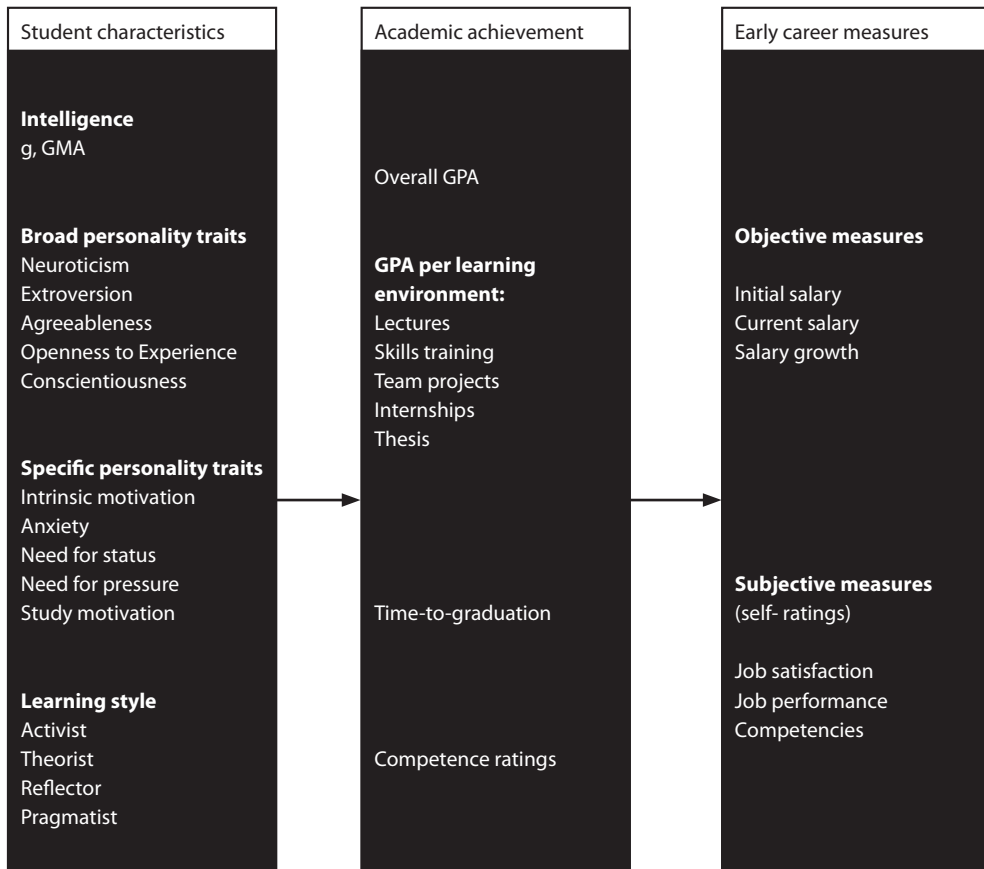


Figure 1.3. Research model

## 1.6 Research questions

After decades of systematic research, there are still many questions about how well cognitive and non-cognitive individual differences predict academic achievement and job performance. What is known about the role of individual difference variables, such as intelligence and personality, needs to be replicated across educational levels, subjects, and assessment methods. The research in this dissertation addresses three main research questions.

1. *How well do student characteristics predict grades?*
2. *How well do student characteristics predict time-to-graduation?*
3. *How well do student characteristics and academic achievement predict early career success?*

The first two research questions are addressed by three empirical studies that are presented in article format in three separate chapters. The first article examines whether the Big Five personality traits are related to student achievement. Although previous research has shown that the trait of conscientiousness is related to academic achievement, findings for the other four traits have been tenuous. One possible reason for the lack of findings is that overall GPA, arguably an insensitive criterion measure, is used by most researchers. Therefore, in addition to GPA, I gathered data on five specific criterion measures of academic achievement: classroom lectures, skills training, team projects, internship training, and a written thesis; and correlated these measures with the Big Five personality factors. The second article considers how well the combined set of variables of intelligence, motivation, and numerous personality traits, predict overall GPA, the five specific criterion measures, and time-to-graduation. An important question in the second article is whether personality traits and motivation predict academic achievement after controlling for intelligence. The third article investigates whether academic achievement can be improved by matching individual learning style to an appropriate learning environment. The five specific criterion measures are matched to four popular learning styles, namely, Activist, Theorist, Pragmatist, or Reflector. Finally, the third research question about how student characteristics and academic achievement affect early career success is examined in chapter 8. Furthermore, the predictive validity of GPA and competence ratings, as well as their unique contributions over intelligence and the personality variables, are also examined.

### *Relevance of the study*

The research in this dissertation has several unique features. First, the use of the five specific achievement measures (e.g., lectures, projects, internships), in addition to overall GPA, makes it possible to exam predictor-criterion relationships in more detail than has

normally been the case. Second, the use of a longitudinal research design that stretches from entry, through college, to early career, provides the opportunity to study complete student careers. Third, next to GPA the inclusion of time-to-graduation adds an economically important learning outcome measure. Fourth, whereas the majority of research in this area has been conducted with university students in English speaking countries, this research is conducted with college students in The Netherlands. Significant study findings here will add to the robustness of general findings in the field. Taken together, these features help advance our understanding of how individual differences are related to different dimensions of academic achievement. The findings of this research serve the stakeholders of science, institutes of higher education, business organizations, and first and foremost, our students.

## **1.7 Overview of the chapters**

This dissertation has nine chapters. Chapter one contains background and context information, the research model, and introduces the main research questions. Chapter two examines the measurement properties of GPA. The multidimensional character of GPA is the motive to determine whether GPA can be decomposed into meaningful facets of academic achievement. Chapter 3 examines the relationship between the Big Five personality traits and academic achievement. The Big Five personality traits are correlated with overall GPA, five specific measures of academic achievement, and with time-to-graduation. Chapter 4 reports on how well intelligence and personality variables predict academic achievement. The individual and incremental validities of the intelligence and personality variables can be determined by examining the variables in an integrated model. Chapter 5 shows what happens when learning styles are matched to correspondingly suitable measures of academic achievement. Chapter 6 considers the question of whether the research findings in this dissertation can be generalized beyond HRM students to other student disciplines such as health and economics. Chapter 7 reports on the topic of competencies. The relationships between traditional exams and competence ratings, and between competence ratings and intelligence and personality variables, are examined. Chapter 8 looks at how well student characteristics such as intelligence and personality, and academic measures such as GPA and competencies, predict early career success. The shared and unique contributions of the diverse predictor variables are reported and discussed. Finally, chapter 9 summarizes the theoretical and practical implications of all of the research.



## Chapter 2

### The reliability of overall, and specific, GPA

#### Abstract

Because much of this dissertation is concerned with the validities of various measures, it is important to consider the reliabilities of the criterion measures. One must be certain that any potential, non significant findings concerning validity are not caused by the unreliability of those measures. Two important measures in this dissertation are overall grade point average (GPA), and facets of GPA that stem from specific learning environments, namely, classroom lectures, skills training, team projects, internship training, and thesis. This study sought to determine the reliability of overall GPA, as well as the reliabilities of specific GPA measures. Grades on 148 HRM students were available. The main findings are that overall GPA is a reliable measure with  $\alpha = .75$  and that classroom lecture, internship training, and team project grades are also reliable.

Two important measures in this dissertation are overall grade point average (GPA), and specific GPA's that stem from grades in specific learning environments. Because much of this dissertation is concerned with the validity of these measures, it is important to consider the reliability of these criterion measures. One must be certain that nonsignificant findings cannot be attributed to unreliability of the criterion measures. Therefore, two research questions are addressed in this chapter: a) how reliable is overall GPA and b) can overall GPA be divided into facets of student performance without adversely affecting reliability? The educational research literature was reviewed to determine the various ways in which estimates of GPA reliability have heretofore been reported.

### 2.1 Literature

Grade point average (GPA) is the arithmetic mean of grades across a wide range of classes in a student's curriculum. It is assumed that GPA reflects a student's capacities and potential, and is therefore a good measure of success in school. In the United States, a student's GPA is often used when judging the suitability for subsequent entry to vocational college, university, or graduate school. It is therefore not surprising that GPA has been the most widely used measure in educational research. However, many researchers have failed to assess the reliability of their GPA measures and have instead relied on reliability estimates reported by others (Bacon & Bean, 2006). That reliability estimates should be reported is evident from the fact that there are many potential sources of rater bias that can affect the GPA measure. Vickers (2000) has noted that GPA, a) does not distinguish between easy and difficult courses, b) is insensitive to the number of courses taken, c) does not distinguish between different skills, and d) lacks a uniform arithmetic system for averaging individual course grades.

Measurement reliability is "the extent to which a set of measurements is free from random-error variance" (Guion, 1998, p. 217). In more practical terms, reliability refers to the consistency of a set of measures. Unfortunately, very few studies have reported reliability estimates for GPA. Authors have typically referred to reliability estimates in other studies rather than reporting their own estimates. And of the studies that have reported reliabilities, the majority have done so using grades from the first year only (Humphreys, 1968; O'Conner & Paunonen, 2007; Schmitt et al., 2007). The methods that have been used to report the reliability of GPA data include test-retest, split-half (Cook, 2004), and Cronbach's alpha (Bacon & Bean, 2006).

#### *Test-retest: Identical and Parallel Forms*

The test-retest method is a well known strategy for estimating the reliability of psychological tests and is also suitable for assessing the reliability of course grades.



Individuals are tested at time one (T1) and are retested with the identical test at time two (T2). The correlation coefficient between T1 and T2 is the reliability estimate. Whereas the identical forms test-retest method seems appropriate for estimating the reliability of a single test or exam, the parallel forms test-retest method seems more feasible for examining GPA reliability.

Contrary to the test-retest method which uses identical test items during T1 and T2, the parallel forms test-retest method is a repeated measures estimate that uses an equivalent set of questions. A large set of items that define the domain of a particular construct is created and the items are randomly divided into two tests – one test for T1 and the second test for T2. The correlation coefficient between the two parallel exams is the reliability of the test or exam. A variant of the parallel forms method has been used by researchers to obtain the reliability of GPA. For example, researchers have correlated the average of all students' grades earned in semester one (T1) with the average of grades earned in semester two (T2) and then accepted the correlation as a reliability estimate of a student's GPA in one semester (Baird, 1985). This approach was used by Humphreys (1969) who collected grade point averages for 1600 undergraduate students and reported stable semester-to-semester correlations over eight semesters, with an average of .58.

There are several problems with the parallel forms strategy for estimating reliability. Both course content and exam format can vary dramatically from one semester to another, in which case the T1 and T2 measures can not be considered parallel. Another problem concerns the possibility of 'method variance', which can occur when the same teacher grades the student's performance in different classes in different semesters or years. This causes reliability estimates to inflate because the teacher is influenced by the student's initial performance during T1 and then grades the student at T2 according to a self-fulfilling prophecy. This parallel method for examining the reliability of GPA can at best be regarded as test-retest if all courses, and by extension course content, are continued from one semester to the next. This is however, almost never the case and it is for this reason that the approach has been questioned. Alternatively, researchers have used correlations between courses that require a common aptitude or competence, such as mathematics or language skills (De Groot, 1978). For example Van de Griend (1964) found a correlation of .59 across grades for courses in which English was an important competency. More recently Bacon and Bean (2006) reported a coefficient of .69 between grades in business courses across three years and of .67 between grades in marketing courses.

### *Split-half Reliability*

In the split-half method, an entire set of items (i.e., set of grades) is randomly divided into two separate sets. The correlation coefficient between the two sets is regarded as the split-half reliability estimate. A study by Barritt (1966) amongst 2,315 first semester freshmen reported a .84 split-half reliability for first semester GPA. Barritt used course grades which were rated on a three point scale. A problem with split-half reliability

estimates is that one never knows which two sets to compare. One can calculate split-half estimates for all possible sets of grades and then average the outcome. However such an approach is often unpractical. A better known, and mathematically equivalent approach, is available and is discussed next.

### *Internal Consistency*

Cronbach's alpha (Cronbach, 1951) provides an estimate of reliability by assessing the internal consistency of a set of items in a scale or test. Based on the assumption that students' grades capture the same latent construct to some degree, and can therefore be combined into a scale, researchers have been using Cronbach's alpha to estimate the reliability of GPA (Bacon & Bean, 2006). The internal consistency reliability of a scale ( $\bar{r}_{kk}$ ) is a function of the average interitem correlation ( $\bar{r}_{ij}$ ) and the number of items in the scale ( $k$ ) and can be estimated using the Spearman-Brown prophecy formula (Nunnally 1978, p. 211):

$$\bar{r}_{kk} = \frac{k\bar{r}_{ij}}{1 + (k-1)\bar{r}_{ij}}.$$

As can be seen in the formula, it is expected that GPA reliability will increase as more courses are included in the GPA (higher  $k$ ), as long as the courses do not diminish the average correlation among the grades ( $\bar{r}_{ij}$ ). However, as course grades are added to the GPA calculation, the course grades can reflect performances in different content areas, different skills, and at different points in time. In that case one can expect the  $\bar{r}_{ij}$  to decline as courses are added.

In the context of test and scale construction, an alpha coefficient of .80 is considered appropriate in an applied setting as long as the number of items (i.e., grades) is not too large (Van de Brink & Mellenbergh, 2006). Bacon and Bean (2006) used Cronbach's alpha to examine the reliability of a longitudinal set of college grades over a four year period. They reported an alpha of .84 for GPA in year 1 (GPA1), .90 for GPA2, .93 for GPA3 and .94 for overall GPA. A comparable estimate of reliability of .85 for GPA was reported by Reilly and Warech (1993).

### *The reliability of facets of overall GPA*

Traditional GPA measures have been criticized because they do not distinguish among different types of courses (Vickers, 2000). For example, a grade on an easy, introductory statistics course is averaged along with a grade from a difficult statistics course about advanced multivariate models. Thus, some students can boast of earning a high GPA simply by taking easy courses when in fact they are not very smart, whereas other

students may have a lower GPA because they took difficult courses but may score much higher on a standard measure of intelligence. The variance of GPA can therefore be affected by an extraneous factor that is unrelated to the meaning of the construct.

Vickers (2000) also noted that overall GPA does not distinguish among different skills. Most students in higher education receive skills training in general competencies such as leadership, communication styles (i.e., feedback) and advising, as well as training in skills specific to a particular course of study. For example, HRM students learn how to interview job applicants and learn mediation skills for solving conflicts between employer and employee. They also receive skills training in advanced English to learn the English HRM jargon and to present in English.

A potential solution to the problems of overall GPA noted above would be to decompose GPA into multiple and specific measures of academic achievement. Previous non-significant findings between GPA and other measures of interest, such as learning styles and personality traits, may be attributed to the fact that overall GPA is plagued with sources of irrelevant variance. In fact, other researchers have called for decomposing broad criterion measures such as overall GPA into multiple components (O'Conner & Paunonen, 2007; Poropat, 2009). Correlations between personality traits, or learning style, and academic achievement could be examined in greater depth and more accurate hypotheses could be formulated (O'Conner & Paunonen, 2007). For example, it would be interesting to know if students who scored high on openness to experience and extroversion outperform their more introverted and narrow minded colleagues on team work. Moreover, the association between intelligence and academic achievement could be studied by examining whether intelligence is differentially related to performance in more traditional learning environments like lectures as compared to more novel learning environments like team work or internships. The relatedness of specific GPA's to other academic achievement criteria could also be examined, such as to time-to-graduation or student satisfaction. In the current Dutch situation, traditional exams are being augmented by new types of assessments such as portfolio and performance assessments. It would be interesting to know with which GPA's the newer types of assessments are most closely related. For example, one could examine if ratings on performance assessment correlate stronger with grades on internship training than with grades on lectures. The first two issues will be addressed in chapter 3 and 4 and the latter issue in chapter 7 of this dissertation.

At this point it is prudent to determine the reliability of multiple and specific measures of academic achievement. Therefore the following two research questions were posited, a) how reliable is overall GPA and b) can overall GPA be divided into facets of academic achievement without to much affecting reliability?

## 2.2 Method

### *Sample*

Undergraduate human resource management (HRM) students in a professional school of higher educational learning participated in a naturally occurring field study. Students studied such HRM issues as how to reintegrate the long term unemployed into the workforce, help employees plan their careers, as well as legal steps involved in employee termination. The current program promoted competency based learning and used a variety of learning environments including classroom tuition, team projects, and skills training to do so. A large amount of time was devoted to internship training in actual business settings during the second, third, and fourth years, and students were required to write a thesis during the last six months of their study.

Students had finished their secondary school education and were generally in the 18 to 22 year age bracket. Data were collected of 148 students who had completed the HRM program at the time the data were collected. Of these, 26% were male and 74% were female.

### *Multidimensional GPA Criteria*

Five specific criterion measures were obtained for each student: classroom lectures, skills training, team projects, internship training, and written thesis. Each criterion score represented an average of multiple assessments throughout the student's curriculum. The number of independent assessments for each student for each criterion was as follows: classroom lectures (25), skills training (10), team projects (10), internship training (3), and thesis (2). All students in the HRM program followed the same curriculum.

Academic achievement data (grades) were obtained from the student administration office. Because the administration office did not collect the grades for research purposes the independence of the measurements can not be guaranteed. Students received grades ranging from one to ten for each assessment they completed with a grade of six representing a pass. However, as is typical in the Dutch education system, students who failed an assessment on their first attempt were permitted to try again, and many did so. However, only grades of six and above were available because first-try failed grades were not recorded by the school.

*Lectures.* Teachers graded students on multiple choice and essay exams and the grades reflected how well students had attended to classroom lectures and assigned textbook readings. Exams were taken during class time and were monitored by the teacher. The teacher performed the role of subject matter expert in this traditional learning environment and communication was mostly one-way.

*Skills training.* Students learned general skills such as negotiating and debating, as well as curriculum specific skills such as how to conduct an employment interview. The teacher used multiple techniques including theoretical presentations, acting, role playing,

and videos to help students master the skills. This learning environment entailed considerable social interaction between the students and the teacher who was a subject matter expert and a coach. Students wrote short reports after class (e.g., at home) in which they reflected on their training experiences and these reports were graded by the teacher. Thus, although the learning environment encompassed social skills, the evaluation measure was based on introspection.

*Team projects.* Students worked in teams of four or five over a ten week period and developed HRM programs (e.g. selection and training) or products such as an HRM handbook or a Health & Safety manual. Students were permitted to plan their own projects but were required to submit a project plan within two weeks. Teachers served as resources, but offered minimal feedback and were generally non-directive. The teamwork required students to learn the subject matter, as well as how to communicate, plan, and resolve conflicts. The quality of the overall team project was evaluated by an independent rater who provided a single grade for the project as a whole. Therefore group members were dependent on each others' contributions and performance.

*Internship training.* Students worked during their third and fourth year of college as junior employees in actual business settings in curriculum relevant areas such as recruiting and applying government employment regulations. Students were trained by experienced supervisors and had only occasional contact with their teachers during this time. Internship criterion scores were obtained from discussions between the workplace supervisors and the students' teachers.

*Thesis.* Students submitted a 30 page written thesis at the end of the curriculum in which they discussed practical business problems and solutions, often stemming from their internship experiences. They had eight to ten contact hours with their supervisor during which they received feedback on their progress. Students managed their own time but were advised to finish their thesis within a six month time-frame. The student's supervisor, as well as an independent faculty member, rated the thesis and corresponding oral presentation, and agreed to a final grade after discussion.

*Overall GPA.* Overall GPA was the cumulative average of grades across all subjects during a student's entire tenure at school. Course grades were weighted by multiplying them by their corresponding European Credits Transfer System (ECTS) credit values. EC values ranged from 1 EC for a short course up to 30 EC for a long period of internship training. One EC represents 28 hours of work and 60 credits are earned during a 42 week full-time year of study, totalling 240 credits for a bachelor degree. The weighted grades were summed for each student and then divided by the total number of grades, which in this study was equal to 51. The same weighing procedure was used in calculating the GPA's for the five specific criterion measures.

### Analyses

Of the various methods for establishing reliability reviewed above, the internal consistency method was used. A homogeneity coefficient ( $\alpha$ ) will be reported for overall GPA, as well as for each of the five specific criteria: classroom lectures, skills training, team projects, internship training, and written thesis.

## 2.3 Results

Descriptive statistics for the overall and specific GPA's are listed in Table 2.1. Tests for skewness, kurtosis, and Kolmogorov-Smirnov normality were conducted to ascertain whether the measures were normally distributed.

Table 2.1 *Descriptive statistics of performance measures (N = 148)*

	Mean	SD	Range		Skewness	Kurtosis	K-S Z*	Ng**
Lectures	6.92	.391	6.0	7.9	.681	.040	1.21, $p = .11$	25
Skills training	7.15	.473	6.0	8.7	.305	-.033	.61, $p = .85$	11
Projects	7.10	.365	6.0	8.0	-.164	.199	.43, $p = .99$	10
Internship	7.27	.709	6.0	9.0	-.453	-.819	2.28, $p = .00$	3
Thesis	6.84	.785	6.0	8.9	.288	-.962	1.94, $p = .00$	2
GPA	7.09	.419	6.1	8.0	-.218	-.660	.72, $p = .68$	51

\* Kolmogorov-Smirnov Z and p-value, \*\*Ng=number of grades

Results showed that indices of skewness and kurtosis were  $< 1$  for all measures. Kolmogorov-Smirnov tests for normality revealed that the internship and thesis measures were not normally distributed. Moreover, the averages of all performance criteria were clustered around the grade of 7 and the scores had small standard deviations.

The intercorrelations among the five specific GPA's and overall GPA are presented in Table 2.2. Performance on lectures correlated strongly with performance on skills training ( $r = .65$ ,  $p < .01$ ). Other correlations were in the range of .29 to .47. Of course, the specific GPA's correlated strongly with overall GPA, ranging from .56 for training to .88 for internship. A homogeneity coefficient ( $\alpha$ ) was calculated for each of the five academic achievement criteria. These coefficients are shown in parenthesis along the diagonal in Table 2.2. The homogeneity coefficient for lectures was based on 25 grades and results indicate that student performance on this criterion is quite stable. Cohen (1975) has suggested a value of .70 as a minimum acceptable level of alpha. The skills training criterion which is based on 11 grades and the internship criterion which is based on only three grades are the least reliable criteria.

Table 2.2 *Intercorrelations and alpha values of specific and overall academic achievement measures*

	Lectures	Skills training	Projects	Internship	Thesis	GPA
Lectures	(.79)					
Skills training	.65	(.60)				
Projects	.47	.31	(.73)			
Internship	.42	.45	.42	(.56)		
Thesis	.32	.29	.33	.39	(.72)	
GPA	.66	.56	.63	.88	.65	(.75)

Note. Alpha coefficients are in parenthesis; all correlations are significant at  $p < .01$

Correlations are often attenuated due to measurement error. Correcting for attenuation shows what the correlation would have been if the measurements were perfectly reliable (Jensen, 1998). The GPA intercorrelations were corrected for attenuation and results are reported in Appendix A. As shown, the highest corrected correlation occurred between lectures and training, namely .94.

Once correlations have been corrected for attenuation they can be assessed with Lisrel to determine if the measurements are distinct or not and therefore should be regarded as separate measures or as one. When for example the performance of students on lectures is equal to their performance on skills training they can not be regarded as unique and distinguishable measures of student performance. Because the second research question of this study aimed to answer the question whether overall GPA could be divided into facets of student performance, it was necessary to examine whether the specific GPA's were statistically distinguishable. Using Lisrel it was tested if the corrected correlations between the specific GPA's would equal 1 ( $H_0$ ). First the highest corrected correlation between lectures and skills was tested, and as could be expected based on the high corrected correlation coefficient, the outcome was not significant ( $\chi^2 = .97$ ,  $df = 1$ ,  $p = .33$ ). This result indicated that the two criteria were not statistically distinguishable. Next, the second highest corrected correlation of .78 between skills training and the internships was tested. The outcome was significant ( $\chi^2 = 5.80$ ,  $df = 1$ ,  $p = .02$ ) and therefore the hypothesis that the correlation between the two criteria could equal 1 was rejected. This meant that the two criteria could be distinguished statistically. In other words students did not perform exactly the same in skills training as they did in their internship training period. Because the other corrected correlations were lower there was no need for additional testing. In sum, the results of the Lisrel analyses indicated that the specific GPA measures were distinct – with the exception of the interrelation of skills training with lectures.

## 2.4 Discussion

This study sought to determine the reliability of overall GPA, as well as the reliabilities of the five specific GPA measures. The main findings are that, a) overall GPA is a reliable measure and b) three of the five specific measures are reliable: lectures, team projects, and skills training.

The findings show that GPA is a reliable measure ( $\alpha = .75$ ) and indicate that a student's performance is reasonably consistent during college; good students tend to perform well and less capable students perform more poorly. The reported estimate of reliability may be an underestimation because Cronbach's alpha is known to lead to a lower bound estimate of reliability (Sijtsma, 2010).

Students also perform consistently within different learning environments, especially on lectures, internship training, and thesis. In other words a student with high grades on his or her first lectures can be expected to achieve high grades on subsequent lectures. Moreover lectures, projects, and thesis had higher reliabilities than the .70 standard advocated by Cohen (1975). Among the five specific GPA measures, the lectures measure is the most reliable with  $\alpha = .79$  and the skills training measure is the least reliable with  $\alpha = .57$ . This could be caused by unreliability in the individual grades due to different raters or due to differences in types of skills practiced in the skills training environment. For example students receive training in advanced English, in general skills like leadership and oral and written communication, and in specific HRM skills like how to interview an applicant or mediate in a conflict. Internship training also appears to be a less reliable performance measure. This could be caused by the small number of grades on which the internship measure is based ( $N_g=3$ ). This is similar to measuring the reliability of subscales of a questionnaire: subscales with fewer items are often less reliable. This result can be expected because Cronbach's alpha is sensitive to the number of 'items', or the number of grades, as in the current study. Alternatively, unreliability may also stem from the fact that external workplace raters have little expertise and unequal norms in grading students or the fact that teachers did not actually see a student perform during an internship but do have an equal vote in determining the rating for the internship.

That specific GPA's have lower reliability than overall GPA was also found by Bacon and Bean (2006) who report homogeneity coefficients of .62 for student performance in business courses and .52 for performance in marketing courses. In their study of undergraduates in marketing education, they find higher reliabilities for yearly GPA's and overall GPA (.94).

Not only is the reliability of the five specific GPA's of interest, but whether these GPA's are distinguishable measures of student performance is also important. If not, one may question the additional value of the use of specific measures above overall measures. Lisrel analyses indicate that although in general student performance is fairly stable, student performance can differ in varying learning environments. The result is in line with the



common aptitude theory (De Groot, 1978) in which an aptitude is expected to lead to moderately stable performance of a student, which in turn manifests itself in stable grades. It is noteworthy that students' performance on lectures and skills training was highly related and statistically not distinguishable. Students who do well on lectures also tend to perform well on skills training.

In sum the results of this study indicate that GPA is a reliable measure of student performance and that several specific GPA's have sufficient reliability and can be regarded as distinct measures of student performance. As mentioned in the introduction, the use of specific GPA's makes it possible to examine predictor/criterion relationships in more detail - for example, to examine whether extroverts outperform introverts in team work projects and visa versa in lectures. If such specific associations exist it may be worthwhile to develop specific interventions for such groups of students.



## Chapter 3

### Using multiple and specific criteria to assess the predictive validity of the Big Five on academic performance\*

#### Abstract

Multiple and specific academic performance criteria were used to examine the predictive validity of the Big Five personality traits. One hundred thirty three students in a college of higher learning in The Netherlands participated in a naturally occurring field study. The results of the NEO-FFI were correlated with grades on five specific measures of academic achievement: classroom lectures, skills training, team projects, internship training, and a written thesis. Results show that conscientiousness is an important predictor for academic achievement in higher education, regardless of which achievement measure was used, and that neuroticism is positively related to achievement when the grading conditions are less stressful.

\* This chapter is based on  
Kappe, F. R., & Van der Flier, H. (2010).  
Using multiple and specific criteria to assess the  
predictive validity of the Big Five personality factors on  
academic performance. *Journal of Research in  
Personality*, 44, 142-145.

### 3.1 Literature

Of the Big Five personality traits, conscientiousness has been the only trait that has been consistently associated with academic achievement (e.g., Bidjerano & Dai, 2007; De Vries, De Vries, & Born, 2011; Komarraju, Karau, Schmeck, & Avdic, 2011; Nofle & Robins, 2007). Poropat (2009) conducted a meta-analysis with a cumulative sample size of 70,000 subjects from primary to tertiary levels of education and reported a population correlation of  $r = .24$  between conscientiousness and academic achievement, however correlations for the other four personality factors were much lower. For example, openness to experience and agreeableness correlated .12 and .07 respectively, with grade point average (GPA).

Most predictive validity studies in educational settings have used generalized criterion measures such as GPA. However, GPA typically reflects arithmetically averaged components of performance and meaningful variance may well have been averaged out in such a multidimensional criterion. Therefore, weak and nonsupportive findings might have occurred because insensitive criterion measures were being used. It is for this reason that Poropat (2009), along with O'Connor and Paunonen (2007), have called for decomposing broad criterion measures such as overall GPA into more specific components. The purpose of this study was to test the predictive validity of the Big Five traits using five specific academic performance criteria: classroom lectures, skills training, team projects, internships, and a written thesis. Hypotheses were formulated based on outcomes of specific personality-performance studies for neuroticism, extroversion, openness to experience, agreeableness, and conscientiousness.

Neurotics have lower academic achievement because they worry about exams beforehand, feel pressured during exams, and dislike being observed. Research has suggested that neurotics might perform better under more relaxed, non-observed conditions, (Geen, 1985; Zeidner, 1998) and on take-home exams (Chamorro-Premuzic & Furnham, 2003a). Therefore,

*Hypothesis 1.* Neurotic students will score higher on an academic performance measure when the measure is free of time constraints or direct observation by the teacher.

Extroverts are proactive, gregarious, social, assertive and impulsive. Poropat (2009) reported that extroversion was more strongly and positively related to academic achievement at the primary ( $r = .18$ ), rather than higher levels of education ( $r = -.01$ ). He reasoned that primary school students were being assessed with informal, social tasks whereby extroverts would have the advantage over introverts. Conversely, extroverts have a disadvantage at higher education levels because assessment is typically associated with more formal, complex, and analytical tasks that require individual preparation. In fact, research has reported that extroverts scored higher on participation and oral expression in college seminar classes (Furnham & Medhurst, 1995), and less well on knowledge tests (Rolfhus & Ackerman, 1999). Therefore,

*Hypothesis 2.* Extrovert students will score higher on an academic performance criterion when the criterion entails social interaction and teamwork.

Openness to experience types tend to be imaginative, nonconservative, have broad interests, and prefer variety over routine. Poropat reported a population correlation of .12 with academic achievement. O'Connor and Paunonen (2007) also conducted a meta-analysis and, although their findings overlap with those of Poropat because they reviewed the same literature, they reported a 90% confidence interval with a sizable spread around the population value they presented and suggested that one or more unknown moderators were at work. Possible moderators stem from research that suggests that working in groups could disfavor open individuals (McCrae & Costa, 1994) and that open individuals have difficulty with discipline and following rules (Chamorro-Premuzic & Furnham, 2005). Therefore,

*Hypothesis 3.* Openness to experience is significantly negatively related to team work that requires conformity and deadlines.

Agreeable types are pleasant and accommodating in social situations, friendly, modest, and cooperative. Poropat (2009) reported a population correlation of .07 with academic performance, however the relationship varied as a function of educational level. The relationship was .30 at the primary school level, where academic achievement ratings were more often based on interaction and functioning as a group, and was .06 at the higher levels of education. For this trait O'Connor and Paunonen (2007) again reported a population value and a 90% confidence interval with a sizable spread, and suggested once more that one or more unknown moderators were at work. Farsides and Woodfield (2003) reported a positive correlation of  $r = .14$  with a sample of undergraduates, and suggested that agreeable students would thrive when instruction and assessment occurred in the context of a collaborative social interaction. Therefore,

*Hypothesis 4.* Agreeableness is positively related to performance in teams (4a) and internship training (4b).

Conscientious individuals are well organized, efficient, determined and persistent. The association between conscientiousness and academic achievement has been reported at all educational levels and across different disciplines and ways of assessing achievement. Conscientiousness has also been found to relate to academic achievement when measured as a student's study pace or as the number of credits students acquired within a certain time period (Busato, 1998). Therefore,

*Hypothesis 5.* Conscientious students will score higher on all academic performance criteria and will graduate in a more timely manner.

## 3.2 Method

### *Participants and Setting*

Undergraduate human resource management (HRM) students participated in a naturally occurring field study. Data were collected from four cohorts of students (2000-2004,  $n = 34$ ; 2001-2005,  $n = 44$ ; 2002-2006,  $n = 57$ ; 2003-2007,  $n = 39$ ). One hundred and forty-eight students, between the ages of 18 to 22, completed the HRM program and 133 accounted for the current data-set after deletions due to missing data. Of these, 26% were male and 74% were female.

Academic achievement data were obtained from the student administration office. Students received grades ranging from one to ten for each assessment they completed with a grade of six representing a pass. Because first-try failed grades were not recorded by the school, only grades of six and above were available for analyses.

Five criterion measures were obtained for each student and each criterion score represented an average of multiple assessments throughout the student's four year curriculum. The criteria and number of independent assessments were as follows: classroom lectures (25), skills training (10), team projects (10), internship training (3), and thesis (2). Having multiple assessments provided criterion score stability, reduced error variance, and increased power (Kenny & Kashy, 1992).

All students were sent a letter some time after having graduated in which they were asked to rate in retrospect the stress they felt while being assessed on each of the five academic criteria. Ninety-four graduates responded and the mean scores for each criterion, as rated on a 5-point Likert type scale, were as follows: lectures (1.72), skills training (0.13), team projects (1.27), internship training (1.07), and written thesis (2.35) whereby higher scores designated more stress. A one-way analysis of variance across the five criteria was significant,  $F(4,90) = 23.42$ ,  $p < .01$ , therefore post-hoc analyses were conducted. Of note is that the difference between the means of written thesis and skills training was significant,  $t = 8.96$ ,  $df = 93$ ,  $p < .01$ .

### *Measures*

*Big Five.* The NEO-FFI short version of the NEO Personality Inventory was used to measure The Big Five. All students completed the Dutch language version of the NEO-FFI as part of an assessment center at the end of their first year of college.

*Lectures.* Teachers graded students on multiple choice and essay exams and the grades reflected how well students had attended to classroom lectures and assigned textbook readings. Exams were taken during class time and were monitored by the teacher who performed the role of subject matter expert in this traditional learning environment.

*Skills training.* Students learned general skills such as negotiating and advising, as well as curriculum specific skills such as how to conduct an employment interview. They subsequently wrote short reports outside of class in which they reflected on their training

experiences, and that were graded by the teacher. Thus although the learning environment encompassed social skills, the evaluation measure was based on introspection.

*Team projects.* Working in teams of four or five over a ten week period, students planned and developed HRM programs such as selection or training, or products such as a HRM handbook or a Health & Safety manual. The team presented a project plan within the first two weeks, a concept report after another seven weeks, and a finalized program or product three weeks thereafter. Teachers served as resources but offered minimal feedback and were generally non-directive. The teamwork required students to communicate, plan, assign work, and resolve conflicts. The project was evaluated by an independent rater resulting in a team score, therefore students were dependent on each others' contributions and performance.

*Internship training.* Students worked during their third and fourth year at college as junior employees in actual business settings. Students were trained by experienced supervisors and had only occasional contact with their teachers during this time. Internship criterion scores were obtained from discussions between the workplace supervisors and the students' teachers.

*Thesis.* Students submitted a 30 page written thesis at the end of the curriculum in which they discussed practical business problems and solutions. They also presented and defended the thesis in public, which was subsequently graded by the student's supervisor as well as by an independent faculty member.

*GPA.* GPA was added for comparison purposes with extant research and was the cumulative measure of grades across all subjects during the student's entire tenure at school.

*Time-to-graduation.* This was the number of months required for a student to complete the curriculum and to graduate from the program.

A power analysis was conducted to determine the probability of finding an expected effect size. With  $N = 133$  and an alpha of .05, the power for conscientiousness, assuming a medium effect size of .30 as based on findings from the meta-analyses, was .93, whereas it was .63 for the other traits, assuming a smaller effect size of .20.

### 3.3 Results

Descriptive statistics, intercorrelations, and reliabilities for all variables are listed in Table 3.1. Reliabilities for the Big Five traits varied between .59 and .79 and coincided with those in the NEO-FFI manual.

As predicted in hypothesis 1, the correlation between neuroticism and academic performance was positive and significant ( $r = .18^*$ ) when the low stress criterion, skills training was used to measure academic achievement. As noted above, students had rated written thesis as most stressful and skills training as least stressful. Further tests (Steiger, 1980) showed that the correlations between neuroticism and thesis ( $r = -.02$ ), and between neuroticism and skills training ( $r = .18^*$ ), differed significantly from one another,  $Z(133) = 2.16, p = .05$  (two-tailed).

Extroversion correlated positively ( $r = .12$ ) with performance on team projects, the criterion that involved the most social interaction. However the correlation was not significant, therefore hypothesis 2 was not supported.

Openness to experience was negatively associated with team projects ( $r = -.19^*$ ) thereby supporting hypothesis 3. Whereas team projects required the most social interaction, the criteria of lecture and written thesis required the least. Comparisons of the correlations between openness to experience and team projects ( $r = -.19^*$ ) lectures ( $r = -.05$ ), and written thesis ( $r = -.07$ ), revealed no significant differences,  $Z(133) = 1.63, p = .10$  (two-tailed), and  $Z(133) = 1.20, p = .23$  (two-tailed), respectively.

Hypotheses 4a and 4b were rejected because correlations between agreeableness and teamwork ( $r = .09$ ), and between agreeableness and internship training ( $r = .13$ ), were non-significant.

Hypothesis 5 was fully supported. Conscientiousness was consistently and significantly related to all academic criteria, with correlations ranging from  $r = .25^{**}$  to  $r = .41^{**}$ ,  $r = .46^{**}$  with GPA, and  $r = -.37^{**}$  with time-to-graduation. The relationship between gender and GPA could be explained by conscientiousness because the GPA variance accounted for by gender dropped from  $r^2 = .078^{**}$  to  $r^2 = .026^{ns}$  when conscientiousness was partialled ( $t = 1.90, p = .06$ ). Finally, along with conscientiousness, the personality traits of neuroticism  $r = -.28^{**}$  and openness to experience  $r = -.29^{**}$  were significantly correlated with time-to-graduation.



Table 3.1 *Intercorrelations of predictor and criterion variables.*

	1	2	3	4	5	6	7	8	9	10	11	12
1 Neuroticism	(.83)											
2 Extroversion	-.53**	(.77)										
3 Openness	.03	.12	(.72)									
4 Agreeableness	-.20*	.18*	.11	(.64)								
5 Conscientiousness	-.22*	.05	-.09	.26**	(.84)							
6 Lectures	-.05	.06	-.05	.10	.34**	(.79)						
7 Skills	.18*	-.17*	-.12	-.06	.31**	.59**	(.60)					
8 Projects	-.04	.12	-.19*	.09	.25**	.51**	.32**	(.73)				
9 Internship	-.05	.03	-.04	.13	.41**	.37**	.41**	.43**	(.56)			
10 Thesis	-.02	-.10	-.07	.12	.33**	.31**	.31**	.33**	.38**	(-)		
11 GPA	-.06	.05	-.08	.14	.46**	.66**	.54**	.67**	.87**	.64**	(.75)	
12 Time-to-graduation	.28**	-.14	.29**	.13	-.37**	-.17	-.17	-.18	-.25*	-.35**	-.40**	(-)
13 Gender	.14	-.15	-.05	.24**	.31**	.28**	.14	.10	.21	.17	.28**	-.30**
Mean	30.98	44.09	37.33	43.70	42.87	6.91	7.15	7.11	7.30	6.87	7.10	50.80
SD	7.41	5.71	5.75	4.89	6.51	.391	.432	.353	.699	.790	.413	4.89
Min.	15	27	22	27	26	6.1	6.0	6.2	6.0	6.0	6.2	48
Max.	51	56	55	54	58	7.9	8.2	8.0	9.0	8.9	8.0	60

Note. Males were coded 1, females 2. Correlations in parenthesis along the diagonal represent internal consistency reliability estimates.

\*\*  $p < .01$ , \*  $p < .05$ .

### 3.4 Discussion

This study investigated whether heretofore nonsignificant and weak associations between the Big Five and academic achievement would materialize if specific, differentiated academic achievement criteria were substituted for the traditional overall GPA criterion. The most important findings of this study are that a) neuroticism is positively related to performance when the assessment conditions are less stressful, b) openness to experience is negatively related to academic performance when deadlines and team conformance are required, c) conscientiousness is an important predictor of performance in higher education regardless of how performance is measured and d) conscientiousness is an important predictor for time-to-graduation.

The population correlation between neuroticism and GPA is  $-.06^{ns}$ , (Poropat, 2009), however it turns positive when students are assessed with a low stress criterion. Geen (1985) and Zeidner (1998) suggest that neuroticism might be positively related to academic achievement under more relaxed conditions in which students are not observed. Students who score high on neuroticism might very well benefit from internet based education whereby they can study and be appraised at home.

The negative relationship between openness to experience and the teamwork criterion confirms the findings of Chamorro-Premuzic and Furnham (2005) and indicates that open individuals perform less well in group situations with deadlines and rules. Instead of stimulating each other to learn and to discuss multiple insights, angles and aspects of a problem or situation, students focus on completing the team assignment on time.

There is no significant association between extroversion and the teamwork criterion ( $r = .12$ ). Although extroverts may have an advantage when there is social interaction with the teacher (Furnham & Medhurst, 1995), the current findings show that this does not extend to social interaction with peers. Neither are there significant associations between agreeableness and the more socially embedded criteria of teamwork and internship training. Being nice and cooperative does not lead to higher performance ratings in these learning contexts. The importance of conscientiousness coincides with meta-analytic findings and can be explained by the fact that conscientious students are more organized, disciplined, and determined. The finding of a consistent association between conscientiousness and diverse academic performance criteria including time-to-graduation adds to the already large body of empirical evidence on the relationship between conscientiousness and educational and occupational success. Identifying students who score low on conscientiousness and intervening with planning and time-management courses may lead to better grades and quicker time-to-graduation rates.

This study has several limitations. Although the sample size of 133 provides adequate power for testing medium effect sizes, tests for smaller effects are underpowered. Further, although our findings suggest that, in some instances, the type of criterion moderates the

relationship between personality and academic performance, research using a classic 2 x 2 factorial design should be conducted. Finally, whereas test fatigue required using the FFI, using the full version NEO will reveal how the multiple performance criteria might interact with the full complement of 30 personality facets.

The results have implications for practice. Institutions that use admission tests should consider including personality tests, especially in the Dutch system where intelligence scores have less predictive power. Because conscientiousness is clearly related to multiple academic achievement criteria, students who score low on conscientiousness should be identified early and encouraged to follow additional courses (e.g., time-management, planning) or to accept coaching. Neurotic students might have the advantage in a learning environment that uses less stressful types of assessments, such as take-home exams or written portfolio's. Neurotic students are at risk of prolonging their studies too long. Neurotics may also benefit from institutions that offer distance education, like the Open University which offers distance learning in which students can study at home, choose when they study, and be assessed with written take-home exams. The results also indicate that introverts, as opposed to extroverts, might also benefit from such a system.



## Chapter 4

### Predicting academic achievement in Higher Education: What's more important than being smart?\*

#### Abstract

This study investigated the combined predictive validity of intelligence and personality factors for multiple measures of academic achievement. Students in a college of higher education in The Netherlands completed a survey that measured intelligence, the Big Five personality traits, motivation, and four specific personality traits. Student performance was measured with grade point average and time-to-graduation, as well as with five specific performance measures: regular exams, skills training, team projects, internships, and a written thesis. Results show that 33% of the variance in GPA and 30% of the variance in time-to-graduation can be explained by combining intelligence, personality, and motivational predictors. Conscientiousness is the best predictor across a broad spectrum of academic achievement measures and explains five times as much variance in GPA as does intelligence. The practical implications are that institutes of higher education should collect personality data on students at the outset and then help students accordingly. Highly conscientious students who are organized and internally motivated might potentially be offered more challenging honours programs with corresponding special commendations on their diplomas, whereas students who score low on conscientious would receive more structure through student study groups, frequent deadlines, shorter assignments, group assignments, clearly defined learning goals, and less second chances for passing examinations.

\* This chapter is based on  
Kappe, F.R., & Van der Flier, H. (2011).  
Predicting Academic Achievement in Higher Education:  
What's More Important Than Being Smart?  
Conditionally accepted for publication.

## 4.1 Literature

Harris (1940) and Cattell (1965) long ago noted the importance of personality factors for predicting academic achievement. Harris (1940) discussed the importance of persistence, in the guise of effort, and concluded that the essential factors for scholastic achievement were: a) ability, otherwise known as intelligence or scholastic aptitude, b) effort, also known as drive or degree of motivation, and c) personal, economic, social, and academic circumstances. Years later, Cattell (1965) suggested that, for university students who had already been selected on intelligence, personality and motivation would be just as important for predicting academic achievement. Recent research has shown that personality accounts for variance in academic achievement over and above intelligence (Bratko, Chamorro-Premuzic, & Saks, 2006; Gilles & Bailleux, 2001; Nettle & Robins, 2007; Poropat, 2009), and that personality may have even more predictive power than intelligence at the post-secondary levels of education (Conard, 2006; Di Fabio & Busoni, 2007; Furnham & Chamorro-Premuzic, 2004; Furnham, Chamorro-Premuzic, & McDougall, 2003; Petrides, Chamorro-Premuzic, Frederickson, & Furnham, 2005).

O'Connor and Paunonen (2007) recently reviewed the studies concerning how well intelligence and personality factors predicted academic achievement at the post-secondary level and reiterated the earlier recommendations of Harris and Cattell. They offered a template for subsequent research and called upon researchers to: a) use multiple predictors beyond intelligence, such as personality, motivation, and study habits, when predicting academic achievement; b) use narrow measures of personality, in addition to the Big Five; c) use multiple and specific measures, in addition to GPA, to assess academic achievement; and d) examine the direct (based on zero-order correlation) predictive validity of each predictor, as well as its incremental validity, in a multivariate context. In this chapter the results are reported of a study with post-secondary students in which the template suggested above was followed. Predictor measures included intelligence, broad and narrow measures of personality, and motivation. Moreover, academic achievement was measured in six different ways in addition to GPA. Finally, multivariate statistics were used to assess the unique and combined effects across the multiple predictors and criteria.

O'Connor and Paunonen (2007) have noted that intelligence has not been a good predictor of academic achievement at the post-secondary education level. This finding has been confirmed in two studies of higher education research in The Netherlands (Busato, Prins, Elshout, & Hamaker, 2000; Resing & Drenth, 2009). Resing and Drenth (2009) reported a modest relationship between academic achievement and intelligence in a general sample of university students, and Busato et al. (2000) reported a minimal correlation of  $r = .13$  in a sample of university psychology students. Based on studies performed in the U.S. and the U.K., Chamorro-Premuzic and Furnham (2005) noted that the correlation between intelligence scores and academic achievement decreased as students became older; declining from .60 to .50, to .40, and to .30, at the elementary,

secondary, university undergraduate, and postgraduate levels respectively. This decline has often been attributed to 'restriction of range' because students have already been selected on intelligence at each successive stage in their education (Boekaerts, 1995). Additional factors, such as personality, are therefore needed to predict academic achievement at post-secondary levels.

O'Conner and Paunonen (2007) and others (Poropat, 2009) have recommended that the Big Five personality factors be used to predict academic achievement in post-secondary education. O'Conner and Paunonen summarized the research across 23 studies and reported an average correlation of  $r = .24$  between conscientiousness and academic achievement. Similarly, Poropat (2009) conducted a meta-analysis and reported an average correlation of  $r = .22$  between conscientiousness and academic achievement. Additional studies have shown that the Big Five have predicted academic achievement, even after controlling for intelligence. For example, conscientiousness accounted for an additional 10% of the variance in GPA in a sample of Italian High School students (Di Fabio & Busoni, 2007); conscientiousness and extroversion explained an additional 17% of the variance in academic achievement in a sample of young adolescents in Croatia (Bratko et al., 2006); and conscientiousness, extroversion, and openness to experience accounted for an additional 12% of the variance in grades on two statistics examinations at the university level (Furnham & Chamorro-Premuzic, 2004).

Whereas the Big Five factors have predicted academic achievement, the use of more narrow personality measures has also been recommended by O'Conner and Paunonen. For example, conscientiousness can be subdivided into the facets of competence, deliberation, dutifulness, order, achievement-striving, and self-discipline – and the latter two have been particularly good at predicting academic achievement. Achievement striving involves being ambitious, diligent, and persistent whereas self-discipline involves the motivation to finish tasks and resist distractions. Reported correlations for these two facets have ranged from  $r = .15$  to  $r = .39$  for achievement striving, and from  $r = .18$  to  $r = .46$  for self-discipline (O'Conner & Paunonen, 2007). Openness to Experience can be subdivided into ideas, fantasy, and aesthetics, however empirical evidence with academic achievement is mixed. Studies report modest correlations, mostly negative, but in some cases positive. Similarly, mixed results have been reported for the facets of extroversion. The facets for neuroticism include impulsivity and anxiety and Chamorro-Premuzic and Furnham (2003b) have reported negative correlations with GPA for impulsivity ( $r = -.26$ ) and for anxiety ( $r = -.29$ ). No significant relations between the facets of agreeableness and GPA have been reported. Additional recommended personality measures include achievement, dominance, exhibition, defence, abasement, nurturance, play, understanding, organization, risk-taking, interpersonal warmth, conformity, and anxiety.

Thirdly, O'Conner and Paunonen (2007) suggested that the effects of intelligence and personality would depend on how academic achievement was operationalized and therefore recommended that multiple measures of academic achievement be used. They

cited work by Rothstein, Paunonen, Rush, and King (1994), who compared the effects of intelligence and personality on academic achievement and reported that written performance was more strongly predicted by intelligence whereas classroom participation was more strongly predicted by personality - specifically extroversion. That there has been a consistent lack of variety in how academic achievement has been measured in post-secondary studies is evident from the research reviewed by O'Conner and Paunonen. They listed 23 published studies between 1991 and 2006, and reported how each study had measured academic achievement: GPA (16 studies), exam grade (6), essay grade (2), written performance (1), thesis research (1), and classroom participation (1).

Although the use of diverse, multiple measures of academic achievement has been scarce, a recent study by Kappe and Van der Flier (2010) assessed the predictive validity of the Big Five for multiple criteria of academic achievement. In addition to performance on exams, the authors measured curriculum-appropriate skills acquisition, team work on group projects, written and oral performance on a thesis, and performance during a student internship. They reported that conscientiousness was an important predictor of academic achievement regardless of how it was measured, and that neuroticism was positively related to achievement when the assessment conditions were less stressful. However, they did not measure intelligence, nor narrow measures of personality, nor motivation, and they reported only bivariate statistics.

A final recommendation by O'Conner and Paunonen (2007) was to examine the direct predictive validity of each predictor, as well as its incremental validity, in a multivariate context. They noted that conclusions concerning the relations between personality and achievement have often been drawn exclusively on the basis of zero-order predictor-criterion relations (O'Conner and Paunonen, 2007). Such zero-order correlations are useful for determining an initial link but should be followed by examining the unique or incremental validity of each personality variable in relation to other variables such as motivation and cognitive ability. Based on the literature review above, the following hypotheses were formulated:

*Hypothesis 1:* Personality variables explain additional variance in academic achievement over and above intelligence.

*Hypothesis 2:* Conscientiousness explains additional variance across a wide range of academic achievement measures over and above intelligence.

In summary, this study answered the call from O'Conner and Paunonen (2007) to use multiple predictors and criteria, narrow measures of personality in addition to Big Five factors, and to combine the results in a multivariate context. And, whereas the majority of research in this area has been conducted in English speaking countries, the current study was conducted in The Netherlands, which would, in the event of significant findings, add to the robustness of research concerning how well intelligence and personality factors predict academic achievement.



## 4.2 Method

### *Sample*

Undergraduate human resource management students in a post-secondary, professional school of higher educational learning participated in the research. Students studied such issues as how to reintegrate the long term unemployed into the workforce, or help employees plan their careers, as well as the legal steps involved in employee termination. The current program used a variety of learning environments such as classroom tuition, team projects, and skills training, which promoted competence-based learning. A large amount of time was devoted to internship training in actual business settings during the second, third, and fourth years, and students were required to write a thesis during the last six months of their study.

Students had finished their secondary school education, had entered as first year freshmen, and were generally between 18 to 22 years old. The data came from four cohorts of students. Cohort 1 started their education in 2000 and cohorts 2, 3, and 4 started their educations in 2001, 2002, and 2003 respectively. The entire sample, across all four cohorts, consisted of 174 students. However, 26 students dropped out prematurely and could not be included in the study because the predictor measures were collected at the end of the first year of the program. There were a total of 34 students in the 2000 cohort, of whom 4 dropped out prematurely, therefore 88% of the enrolled students participated in the current study. Drop-out and participation figures for the other three cohorts were as follows: 9/79% for the 2001 cohort, 7/88% for the 2002 cohort, and 6/85% for the 2003 cohort. Of the 148 students who participated in the study, a further 11 students were deleted because complete predictor and criterion data were unavailable. The distribution of gender favoured females (74%) over males (26%).

A power analysis with a two-tailed significance test,  $\alpha = .05$ , and  $n = 137$  yielded a value of .81. An estimated effect size of  $r = .30$  was posited, which was based on a review of several studies that had reported incremental validities of the Big Five factors with academic achievement, above and beyond intelligence.

### *Predictor Measures*

*Intelligence.* Student intelligence was measured with the short version of the Multiculturele Capaciteiten Test voor Hoger Niveau (Multicultural Test of Intellectual Ability for Higher Education) by Bleichrodt and Van den Berg (1999). This test was developed especially for students in higher education, is considered to be a psychometrically sound instrument by the Dutch Testing Committee, the COTAN (Evers, Van Vliet-Mulder, & Groot, 2000), and is used widely in Dutch higher education. The COTAN assessed the MCT-H on the following criteria: construct and criterion validity, reliability, principles of test construction, quality of test material, norms, and quality of the manual. As an example of construct validity the correlations between gender and

intelligence in this test are alike correlations reported in the manuals of other international intelligence tests. The correlation with age was also examined. Results show a negative correlation for fluid intelligence and a positive correlation for crystallized intelligence, both in line with international research. The following three subtests, listed with their corresponding reliabilities, were used: double word analogies items such as, shoe is to .... as glove is to .... ( $\alpha = .95$ ), exclusion items in which two pictures, out of a series of six, were excluded ( $\alpha = .84$ ), and speed and accuracy tests in which word or number pairs were compared ( $\alpha = .96$ ). This shorter version of the MCT-H was used because of time constraints and to minimize test fatigue. The short version correlates as high as  $r = .86$  with the long version, whose overall reliability is .97.

*Big Five Personality.* All students completed the short version of the NEO, the NEO-FFI (Costa & McCrae, 1992). The NEO-FFI is composed of 60 self-descriptive statements; 12 statements for each of the five factors. Items were rated on a five point Likert scale, ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). A reference to the Dutch language version can be found in Hoekstra, Ormel, and De Fruyt (1996). Cronbach alphas for the current study were as follows: conscientiousness (.84), neuroticism (.83), extroversion (.77), agreeableness (.64), and openness to experience (.72). These values coincided closely with reliability values that have been reported in the NEO-FFI manual.

Items for the remaining personality variables of intrinsic motivation and anxiety were taken from the WerkStijlenVragenlijst (Questionnaire About One's Style of Working) by Akkerman and Van der Linden (2000), which in turn was based on the work of Hermans (1967, 1976). Four additional variables: anxiety, need-for-status, motivation-to-study, and environmental press, were also taken from the WerkStijlenVragenlijst and were included in the current research on an exploratory basis. Items on all scales were rated on seven point Likert scales, ranging from 1 (*strongly disagree*) to 7 (*strongly agree*).

*Intrinsic motivation.* The intrinsic motivation subscale consisted of 12 items. Sample items include: "I feel that everything I do must be of a high standard", "I often do more than is expected of me", and "People should not put off until tomorrow that which can be done today". The user manual has reported the following intercorrelations: extroversion (.27), neuroticism (-.27), and conscientiousness (.56). The internal reliability ( $\alpha$ ), as measured in the current study was .79.

*Anxiety.* Anxiety was defined as underachieving due to uncertainty and stress when tasks are experienced as difficult or must be performed under stressful conditions. Sample items include: "I am the type of person who can develop interesting ideas, but who has difficulty acting upon them" and "The thought that a decision might turn out wrong makes me freeze in my tracks". The scale contained 27 items and registered an internal reliability ( $\alpha$ ) of .86. A high score implied the experience of strong feelings of fear and stress. O'Conner and Paunonen listed anxiety as a variable of interest and studies have shown it to correlate -.29 with school grades (Hembree, 1988), -.28 with oral exams (Diaz, Glass, Arnkoff, & Tanofsky-Kraff, 2001), and -.29 with scholastic achievement (Chamorro-Premuzic

& Furnham, 2003b). As might be expected, anxiety is also correlated with other predictors of academic achievement such as intelligence,  $r = -.33$  (Ackermann & Heggstad, 1997), and neuroticism,  $r = .53$  (Busato et al., 2000).

*Environmental press.* Environmental press was defined as the need for additional tension and challenging tasks to reach optimal performance. Sample items included: "I work a lot better when under pressure", "I am easily distracted when there is no pressure", and "I only really get into gear as a deadline approaches". The scale contained 13 items and registered an internal reliability ( $\alpha$ ) of .84. A high score implies that additional pressure is needed to motivate an individual to perform.

*Need for status.* The need-for-status was defined as an internal desire to define success by comparing oneself to others. It consisted of 15 items, such as: "What others think of my accomplishments is important to me", "I often compare my accomplishments to those of others", and "I strive for positions in life that provide recognition". The user manual reported the following correlations: extroversion (.33), agreeableness (-.39), and conscientiousness (.21). The internal reliability ( $\alpha$ ), as measured in the current study was .86.

*Motivation-to-study.* Motivation-to-study was defined as the motivation to learn new knowledge and skills and the desire to study for one's personal development. Sample items for this 12 item scale included: "I would rather read a textbook than watch a movie" and "One needs a good education in order to advance in life". The scale has shown the following intercorrelations: extroversion (.27), neuroticism (.26), conscientiousness (.46), openness to experience (.27), and intrinsic motivation (.52). The internal reliability ( $\alpha$ ) was .76.

### *Measures of Academic Achievement*

Grades on five different measures of academic achievement were obtained for each student: classroom lectures, skills training, team projects, internship training, and written thesis. Students were graded by their teachers on a scale from 1 to 10, with 6 representing a pass, for each completed course or learning module. Lecture, skills training, team project, and internship training grades were collected throughout the curriculum, and the thesis was graded at the end of the curriculum. Cumulative GPA was included as an overall measure of academic achievement as was time-to-graduation.

*Lectures.* Teachers graded students on multiple choice and essay exams and the grades reflected how well students had attended to classroom lectures and assigned textbook readings. In total, over the first three years of the HRM curriculum students had 25 lecture courses. Therefore GPAlectures was based on 25 grades.

*Skills training.* Students learned general skills such as negotiating and debating, as well as curriculum specific skills such as how to conduct an employment interview. Teachers used theoretical presentations, acting, role playing, and videos to help students master the skills and considerable social interaction was required between the students

and the teacher. Students wrote short reports after class (e.g. at home) in which they reflected on their training experiences and these reports were graded by the teacher. In total, students were graded on 10 different training courses during the four year HRM curriculum.

*Team projects.* Students worked in teams of five over a ten week period and developed HRM programs (e.g. selection and training) or products such as an HRM handbook or a Health & Safety manual. Students planned their own projects and the teamwork required students to learn how to communicate, plan, and resolve conflicts with one another. The project was scored as a whole, therefore students were dependent on each others' contribution and performance. Whereas first year students were assigned to project teams, students chose with whom they wanted to work in all subsequent years. Students were graded on 10 different team projects during the four year HRM curriculum.

*Internship training.* Students worked as junior employees in actual business settings in curriculum relevant areas such as recruiting and applying government employment regulations. Students were trained by experienced supervisors and grades were obtained after discussion between the workplace supervisor and the student's teacher. Students received six weeks of training during their second year, and 20 weeks during the third and fourth years.

*Thesis.* Each student submitted a 30 page thesis towards the end of the curriculum in which practical business problems and solutions were discussed, often stemming from internship experiences. The student's supervisor, as well as an independent faculty member, rated the thesis and corresponding oral presentation, and agreed to final grades after discussion.

*GPA.* Grade point average was the weighted, cumulative average of grades across all subject matter during the student's entire tenure at school.

*Time-to-graduation.* Time-to-graduation was the number of months required for a student to complete the curriculum and graduate from the program. The formal curriculum was meant to be completed in four years, however the system was somewhat flexible and a number of students took longer to complete their studies.

### 4.3 Results

Bivariate correlations for all predictor and criterion variables, as well as descriptive statistics, have been listed in Table 4.1. As is often reported intelligence was not significantly correlated with any of the personality variables. As expected in our sample of post-secondary students, intelligence showed only small correlations with measures of academic achievement. In contrast, conscientiousness showed large correlations with the five specific academic achievement criteria, GPA, and time-to-graduation. Intrinsic motivation was correlated with almost all of the criteria, however intrinsic motivation and

conscientiousness were substantially intercorrelated ( $r = .45$ ). Motivation-to-study was obviously relevant in a college level sample and it correlated consistently with academic achievement, however it correlated highly with intrinsic motivation ( $r = .62$ ) which shows that it too was not a purely independent construct. Anxiety was unrelated to any of the performance criteria, however it correlated as expected with neuroticism ( $r = .32$ ). Need-for-status was unrelated to any of the academic achievement criteria, however environmental press showed small inverse correlations with a few of the criteria. The Big Five personality traits showed low intercorrelations amongst themselves, except for a rather high correlation between extroversion and neuroticism ( $r = -.41$ ). Finally, the academic achievement criteria were consistently intercorrelated amongst themselves in the expected directions: good students did well consistently across a broad spectrum of academic criteria.

Table 4.1 Descriptives and bivariate correlations ( $N=137$ )<sup>a</sup>

	IQ	Neur	Extr	Open	Agree	Cons	IM	AN	Press	Status	Study	Lect	Train	Team	Intern	Thesis	GPA	TTG
IQ	-																	
Neur	.06	-																
Extr	-.12	-.41**	-															
Open	-.14	.00	.16 <sup>†</sup>	-														
Agree	-.04	-.18*	.16 <sup>†</sup>	.11	-													
Cons	-.01	-.17*	.00	-.11	.24**	-												
IM	.10	.09	.03	.01	.03	.45**	-											
AN	.02	.32**	-.27**	-.05	-.04	-.20*	.09	-										
Press	.11	.03	.11	.20*	-.21*	-.24*	-.03	.18*	-									
Status	.11	.21*	.02	.04	-.29**	-.03	.23**	.30**	.25*									
Study	.04	.15	-.04	.25**	-.10	.33**	.62**	.07	.06	.28**	-							
Lect	.25*	.06	.02	-.08	.09	.37**	.22**	-.07	-.06	-.01	.28**	-						
Train	.31**	.16 <sup>†</sup>	-.16 <sup>†</sup>	-.16 <sup>†</sup>	-.04	.36**	.24**	.14	-.04	.07	.22**	.66**	-					
Team	.15 <sup>†</sup>	.07	.08	-.20*	.08	.26*	.17*	-.03	-.09	.08	.18*	.50**	.36**	-				
Intern	.14 <sup>†</sup>	-.06	.00	-.09	.12	.44**	.34**	-.04	-.13	-.01	.23*	.42**	.44**	.42**	-			
Thesis	.24*	.01	-.10	-.08	.12	.33**	.26**	-.05	-.23*	.01	.23**	.33**	.33**	.31**	.39**	-		
GPA	.23*	.03	.00	-.12	.12	.47**	.36**	-.07	-.18*	.01	.30**	.67**	.59**	.64**	.88**	.64**	-	
TTG	-.16 <sup>†</sup>	.22*	-.06	.30*	-.12	-.41**	-.26*	.07	-.14	.07	-.03	-.27*	-.22*	-.22**	-.41**	-.32**	-.44**	-
Gender	-.10	.12	-.20*	-.09	.24**	.27**	.15 <sup>†</sup>	.14	-.14	-.18*	.12	.29**	.21*	.07	.22*	.22**	.27**	-.30**
Mean	94.03	31.78	44.01	37.37	43.66	43.31	59.71	114.89	58.29	62.86	53.09	6.92	7.15	7.12	7.30	6.88	7.12	50.80
SD	3.58	7.49	5.36	5.58	4.85	6.53	8.23	18.98	11.36	11.85	9.12	.40	.45	.36	.70	.79	.41	4.89

<sup>a</sup> IQ = Intelligence; Neur = Neuroticism; Extr = Extroversion; Open = Openness; Agree = Agreeableness; Cons = Conscientiousness; IM = Intrinsic motivation; AN = Anxiety; Press = Environmental pressure; Status = Need for status; Study = Motivation to study; Lect = lectures; Train = skills training; Team = team projects; Intern = internship training; Thesis = thesis; GPA = overall grade point average; TTG = time-to-graduation. Note. Males were coded 1, females 2. \*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$ .

The first hypothesis stated that personality would explain additional variance in academic achievement over and above intelligence. This hypothesis was supported for GPA, time-to-graduation, and for each of the five separate measures of academic achievement. The Big Five traits together registered incremental validities, over and above intelligence, of  $\Delta R^2 = .24$  for GPA and  $\Delta R^2 = .25$  for time-to-graduation. Intrinsic motivation registered incremental validities of  $\Delta R^2 = .11$  for GPA and  $\Delta R^2 = .06$  for time-to-graduation. The four narrow traits of anxiety, motivation-to-study, need-for-status, and environmental press together registered an incremental validity of  $\Delta R^2 = .14$  for GPA, but no significant effect for time-to-graduation. The proportions of variance explained by intelligence were  $R^2 = .05$  for GPA and  $R^2 = .03$  for time-to-graduation.

The second hypothesis, which stated that conscientiousness will account for additional variance across a wide range of academic achievement measures over and above intelligence, was also supported. Conscientiousness showed the following incremental validities: GPA ( $\Delta R^2 = .22$ ), time-to-graduation ( $\Delta R^2 = .17$ ), lectures ( $\Delta R^2 = .14$ ), skills training ( $\Delta R^2 = .13$ ), team projects ( $\Delta R^2 = .07$ ), internship training ( $\Delta R^2 = .19$ ), and thesis ( $\Delta R^2 = .11$ ).

We further examined the predictive value of the total set of predictor variables and examined their unique contribution in relationship to each other for GPA, TTG and the five specific measures of academic achievement. The results are listed in Table 4.2. Results show that 33% of the variance in GPA and 30% of the variance in time-to-graduation can be explained by combining intelligence, personality, and motivational predictors. Conscientiousness was by far the most important predictor with significant and large beta values for all dependent variables. Intelligence was the second predictor of importance with significant beta values for GPA, lectures, skills training and thesis. The unique contribution of the other personality traits, like intrinsic motivation or study motivation, was limited.

Table 4.2 Results of multiple regression analyses for all variables

	GPA	TTG	Lectures	Skills training	Team projects	Internship training	Thesis
	$\beta$	$\beta$	$\beta$	$\beta$	$\beta$	$\beta$	$\beta$
Intelligence	.23**	-.14 <sup>†</sup>	.26**	.30**	.13	.14 <sup>†</sup>	.25**
Neuroticism	.10	.15 <sup>†</sup>	.15 <sup>†</sup>	.15 <sup>†</sup>	.17 <sup>†</sup>	-.01	-.01
Extroversion	.09	-.06	.14	.02	.20*	.03	-.05
Openness	-.08	.22*	-.09	-.07	-.25*	-.03	-.02
Agreeableness	.04	-.01	.03	-.11	.09	.02	.07
Conscientiousness	.37**	-.32**	.34**	.44**	.20 <sup>†</sup>	.34**	.19 <sup>†</sup>
Intrinsic motivation	.06	-.19 <sup>†</sup>	-.11	-.05	-.10	.16	.05
Anxiety	.01	-.06	.00	.20*	-.02	.04	.00
Need for pressure	-.10	.02	.00	-.01	-.03	-.05	-.20*
Need for status	-.05	.06	-.10	-.08	.05	-.06	.02
Study motivation	.16	.11	.26*	.08	.21 <sup>†</sup>	.04	.15
$R^2$	.33	.30	.26	.31	.19	.24	.23
$F(11, 125)$	5.61**	4.95**	3.93**	5.15**	2.58**	3.62**	3.40**
Adjusted $R^2$	.27	.24	.19	.25	.11	.18	.16

Note. GPA = overall grade point average; TTG = time to graduation. \*\*  $p < .01$ , \*  $p < .05$ , <sup>†</sup>  $p < .10$ .



Because the sample favored females we investigated whether gender would account for variance in the dependent measures above and beyond intelligence, personality, and motivation. The females had higher GPA's ( $r = .27$ ), graduated faster ( $r = -.30$ ), were more conscientious ( $r = .27$ ), more agreeable ( $r = .24$ ), although less extroverted ( $r = -.20$ ), and showed less need-for-status ( $r = -.18$ ). Results showed that gender accounted for an additional 2% of the variance in GPA and an additional 5% in time-to-graduation. The effects of other predictor variables in the equation remained stable when gender was added. Differential validity - that is whether intelligence, personality, and motivation predicted academic achievement differently for males than for females - was also investigated. Multivariate analyses for GPA and time-to-graduation, in which gender functioned as a fixed factor and intelligence and personality functioned as covariates, revealed no significant interaction effects.

Finally, as recommended by O'Conner and Paunonen (2007) we used structural equation modeling. We used the observed predictor and criterion data of the 137 participants and entered the data as manifest variables. Although we already had a large set of predictor variables and two criterion measures of academic achievement, we added gender as a variable. For structural equation modeling to have additional informational value above and beyond that of correlation and regression analyses, a stepwise, backward approach was used. We started out by testing a model in which all predictors were included for GPA as well as for time to graduation. We then excluded the predictor variable with the smallest t-value. To illustrate, the model with all variables totalling 22 possible paths showed a perfect fit, however some of the paths were nonsignificant. We deleted the path with the smallest t-value and subsequently retested the model. We repeated this procedure until the model consisted of only significant paths with t-values larger than 2.0. The path diagram of the final model is shown in Figure 4.1 ( $\chi^2 = 1.23$ ,  $df = 3$ ,  $p = .75$ , RMSEA = 0.00). Significant paths were found between conscientiousness, intelligence, gender for both GPA and time-to-graduation. The results also show that study motivation is important for high grades whereas neuroticism and openness are important for timely graduation, and that study motivation is more important than intrinsic motivation. The latter is explained by the contextual specificity of study motivation in comparison with intrinsic motivation.

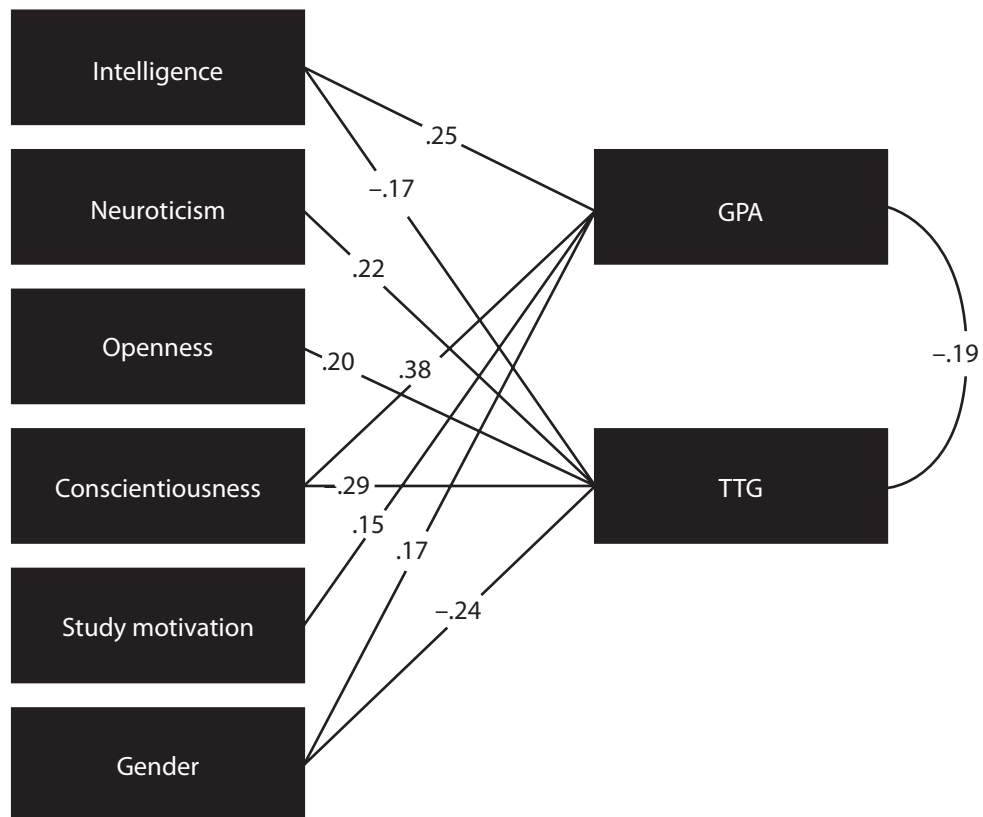


Figure 4.1 Path diagram

#### 4.4 Discussion

The data show clearly that a combination of variables can accurately predict academic achievement. In fact, 33% of the variance in GPA and 30% of the variance in time-to-graduation can be explained by combining intelligence, personality, and motivational predictors. Conscientiousness is the best predictor across a broad spectrum of academic achievement measures and explains five times as much variance in GPA as does intelligence. Although intrinsic motivation is an important predictor on its own, and explains twice as much variance as intelligence, it explains only minimal variance in academic achievement beyond personality, especially conscientiousness.

The findings in this study support the writings of Harris (1940), Cattell (1965), and O'Conner and Paunonen (2007) that personality is an important predictor of academic achievement in post-secondary education. Intelligence, when entered as the first variable in the regression equation, accounts for only 5% of the variance in GPA in the current study, which coincides with the 4% reported by Farsides and Woodfield (2003) and 3% reported by Furnham and Chamorro-Premuzic (2004). Conscientiousness is the most consistent and strongest predictor of academic achievement. Not only does conscientiousness, after controlling for intelligence, account for 22% of the variance in GPA and 17% of the variance in time-to-graduation, it is also consistently related to each of the five specific performance criteria. The current findings dovetail nicely with those of Mount, Barrick, and Steward (1998) who report that conscientiousness correlates with various job related performance criteria across diverse occupations. Conscientious individuals perform better because they persevere longer and are more organized than their counterparts.

The findings in this study show that of all personality traits only the effect of conscientiousness on academic achievement is robust. Although not all findings need to be replicated across cultures, it is advisable to do so when cultural contingencies raise doubts about whether findings from important initial studies are generalizable. Dutch higher education practices differ from those in other countries in three notable ways. Admission standards are much less stringent. Students who complete their high school education are automatically admitted to an institute for higher education and selection criteria such as high school GPA, SAT or ACT scores are not generally used. Second, higher education in The Netherlands is substantially subsidized by the government and is therefore much less of a financial burden on students and their families. Therefore, there is presumably less external pressure on students to succeed. Yet, even in this context, the effect of conscientiousness on academic achievement is still quite large. Third, the curriculum of the HRM program in this study is quite fixed in the sense that all students follow the same courses, which is different from a typical university liberal arts curriculum in which students have more flexibility after having chosen a major.

Intrinsic motivation is related to both GPA and time-to-graduation and by itself explains almost twice as much variance in GPA as does intelligence. The magnitude of this

effect coincides with the meta-analytic mean population correlation of .26 reported by Robbins, Lauver, Le, Davis, Langley, and Carlstrom (2004), and of .24 found by Vallerand, Pelletier, Blais, Briere, Senecal, and Vallieres (1993) in a comparable sample of 220 undergraduate college students. Intrinsic motivation also correlates significantly with four of the five specific academic achievement criteria in this study. However, a caveat is in order in that motivation and conscientiousness are strongly related. Whereas the bivariate correlations between motivation and the achievement measures are sizeable, the partial correlations after controlling for conscientiousness are much smaller. A simple reason for this finding is that a number of items between the two scales of conscientiousness and intrinsic motivation have quite similar content. In the following examples the first item is from the conscientiousness scale, whereas the second item is from the intrinsic motivation scale: a) "I am a productive person who always finishes what I start/I always finish whatever I start."; b) "I am quit good at motivating myself to finish things on time"/"People should not put off until tomorrow that which can be done today"; c)"I work hard to reach my goals"/"I keep working until a problem is solved". In all, seven items between the two scales have highly similar meaning, which is considerable because both scales have only 12 items. A factor analysis of the conscientiousness and intrinsic motivation items yields a three factor solution. The factor analysis shows that the mentioned item pairs load on the same factor. Clearly, components of conscientiousness and intrinsic motivation are similar. Therefore, adding motivational factors in a prediction model with personality traits can lead to the inaccurate perception that motivation is less important.

Although a strength of the current study is the realism and face validity of the context, a limitation is that it is necessary to work with the measures at hand, namely those used by the school of higher education that participated in this research. Thus, one clear limitation in this study is that the narrow personality factors such as need-for-status and environmental press are not, as is recommended by O'Conner and Paunonen, among the type of 30 narrow facets of Big Five factors that are measured by the NEO-PI-R (Costa & McCrae, 1992). It is unclear why anxiety, which is mentioned as one of the 30 facets, did not correlate appreciably with the achievement criteria. This may be related to how anxiety is operationalized, or to the sample being Dutch. Additional research is needed concerning the degree to which narrow personality measures account for variance in academic achievement above and beyond the Big Five factors (O'Conner & Paunonen, 2007). A second limitation concerns the 15% of the enrolled students who did not finish the program. This is however a natural occurrence in higher education and might make the conclusions more conservative. Had the drop-outs been included it is possible that they would have scored lower on conscientiousness as well as had lower GPA's. A third limitation is that there is no measure of class attendance or self-regulation in the current study, variables which account for significant variance in GPA beyond intelligence (Farsides & Woodfield, 2003). Finally, it should be noted that although the measure of intelligence

used in this study is endorsed by experts in psychometrics, is widely used in The Netherlands, and shows some evidence of construct validity, there is no convergent validity data of this test with more commonly used IQ tests.

The results of this study have practical utility for admission boards and counsellors. Students who are intelligent and conscientious, and who therefore do not need additional pressure to study, can participate in a special honours program in which they are challenged more during their coursework and for which they receive an additional commendation on their diploma that distinguishes it from a normal diploma. The diploma with commendation should give such students an edge when soliciting for employment. Conversely, students who score low on conscientiousness might benefit from programs that improve motivation, such as participating in study groups. A study group consists of several students who happen to be taking the same course and can be either self-supporting or coached by an advanced student who has already mastered the subject. The groups meet frequently, for example after lectures, and students share notes and discuss the subject in order to clarify difficult concepts. Students who are low in conscientiousness may also benefit from programs that provide more structure. This is especially important in the first year of college because a number of students experience difficulties when transitioning from secondary to higher education. Schools might consider a curriculum that has more deadlines, shorter assignments, more group assignments, clearly defined learning goals, and less second chances for passing examinations. Students will not need to plan as much when structure is provided and should therefore achieve better grades, finish their first year on time, and have lower drop-out rates.

The practical utility of the time-to-graduation criterion is of interest to policy makers because education is costly. Post-secondary education in The Netherlands is almost entirely subsidized by the government, and both students and schools receive financial support from the government. The normal duration of a college education is four years, however students may receive funding for one additional year to complete their studies if needed. Students who drop out must return their financial support to the government. Schools receive funding from the government for every student they graduate, commonly referred to as a 'diploma bonus'. Thus, students who fail to graduate become expensive liabilities, both to themselves as well as to their schools. Because conscientiousness, neuroticism, and openness to experience are related to time-to-graduation, students might be identified on these characteristics. Several post-secondary schools in The Netherlands are experimenting with screening students before they embark on their studies and use personality tests to assess incoming freshmen in an effort to increase graduation rates.

Harris noted in 1940 that "ability and effort are two essential factors for student achievement" (Harris, 1940, p.151). At the post-secondary level of education, the findings of the current study show clearly that effort, especially in the form of conscientiousness, overshadows ability as a predictor of grades and of how quickly a student will graduate. More important than what a student *can* do, is what a student is *willing to* do.

## Chapter 5

### Learning styles and academic achievement\*

#### Abstract

Multiple and specific learning criteria were used to examine the predictive validity of the Learning Style Questionnaire (LSQ). Ninety-nine students in a school of higher education in The Netherlands participated in a naturally occurring field study. The students were scored on four LSQ dimensions, namely, Activist, Theorist, Pragmatist, and Reflector, and they were also graded on five different learning criteria throughout their four year education: classroom lectures, skills training, team projects, internship training, and a written thesis. Although learning styles were matched to correspondingly suitable learning criteria, the LSQ revealed no predictive validity, however we can report good test-retest reliabilities over a two year time period. Using the LSQ to stimulate learning in college students is debatable given the lack of positive findings.

\*This chapter is based on  
Kappe, F. R., Boekholt, L., Den Rooyen, C., & Van der Flier, H. (2009).  
A Predictive Validity study of the Learning Style Questionnaire  
(LSQ) Using Multiple, Specific Learning Criteria.  
*Learning and Individual Differences, 19*, 464-467.

## 5.1 Literature

People differ dramatically in how quickly and easily they learn new material. One theory often promoted to improve learning efficiency is learning style, which posits that people learn best when their particular learning styles are matched to correspondingly suitable learning environments (Coffield, Moseley, Hall, & Ecclestone, 2004; Ford & Chen, 2001; Pheiffer, Holley, & Andrew, 2005). Learning style theory has become quite popular commercially, but predictive validity studies show that empirical support for the theory is thin (Coffield et al., 2004). An important methodological problem of many predictor studies in educational settings has been the use of the ubiquitous, generalized student GPA as a criterion measure. The purpose of the present study was to provide a more accurate test of the learning style model by pairing the various learning styles with multiple and specific performance criteria suited to each learning style preference.

One model of learning styles that has generated a significant amount of research is that of David Kolb (1984). While teaching management students he noticed that some students preferred learning through experiences whereas others preferred the traditional classroom lecture. His subsequent theory of experiential learning proposed that, while learning, people resolved conflicts between a) active experimentation and b) reflective observation along one axis and between c) concrete experience and d) abstract conceptualization along another axis. His model yielded four quadrants and he stated that, over time, people developed learning style preferences that can be categorized into one of the four quadrants. Kolb developed the Learning Style Inventory (LSI) to measure peoples' individual learning styles. By knowing the learning styles of their students and by creating learning environments matched to those learning styles, educators could enhance learning. Whereas commercially popular pedagogical tools have been generated by Kolb's work, empirical support for construct and predictive validity have been lacking (Coffield et al., 2004).

Difficulty with some of Kolb's theoretical ideas and low face validity for the LSI prompted Honey and Mumford (1986) to develop their own learning style theory as well as a new measure, namely, the Learning Style Questionnaire (LSQ). The LSQ also identified four types of learners: Activists, Theorists, Pragmatists, and Reflectors. Activists are individuals who enjoy new experiences, tend to make decisions intuitively, but who dislike structured procedures. Theorists focus on ideas and systemic logic and are distrustful of intuition and emotional involvement. Pragmatists like practical, down to earth approaches and debate, but tend to avoid reflection and deep levels of understanding. Reflectors observe and describe processes, try to predict outcomes and try to understand meaning.

Despite its commercial success, empirical support for the LSQ has been mixed. Allinson and Hayes (1988) concluded that the temporal stability and internal consistency of the instrument were well established and provided some evidence of construct validity but were unable to find concurrent or predictive validity. Whereas some researchers have



reported good test-retest reliabilities (Duff, 2000; Honey & Mumford, 2000; Swailes & Senior, 1999; Veres, Sims, & Locklear, 1991), others have questioned the internal consistency of the LSQ or reported low test-retest values (Jackson & Lawty-Jones, 1996; Ruble & Stout, 1993; Sims, Veres, & Shake, 1989). Duff and Duffy (2002) employed exploratory and confirmatory factor analysis but failed to find four separate learning styles and reported that learning style was a weak predictor of academic performance. Price and Richardson (2003) used the LSQ to predict student performance, study techniques, and recall processes, but failed to find significant results. Finally, Zwanenberg, Wilkinson, and Anderson (2000) were unable to predict learning outcome scores for any of the four LSQ subscales and surmised that the use of generalized, instead of specific, criterion measures might have accounted for their lack of findings.

Most predictive validity studies in educational settings have used generalized criterion measures (O'Conner & Paunonen, 2007), namely, grade point average (GPA). An overall indicator such as GPA typically reflects arithmetically averaged components of performance and one could well have averaged out meaningful variance with such a criterion. It is quite possible that the weak and nonsupportive findings mentioned above were attributable to the use of insensitive criterion measures. It is for this reason that O'Conner and Paunonen (2007) have called for decomposing the broad criterion measures, such as overall GPA, into more specific components. The purpose of this study was to test learning style theory by using a number of specific, different criteria that could be logically related to the four learning styles measured by the LSQ.

Honey and Mumford (2000) have developed a table in which they describe how individuals who score high on each of the four different learning styles prefer to learn. This table served as the basis for linking the learning styles to the multiple performance criteria available to us in this study. For example, the table states that Activists prefer action learning, job rotation, role playing, business game simulations, discussion in small groups, training others, and outdoor activities. We therefore hypothesized that Activists would score high on a measure of internship training as well as score high on a measure of practical skills training. Similarly, Theorists and Reflectors, who prefer listening to lectures, reading, e-learning, and self-directed learning should score high on a criterion that measured attentiveness to lecture material. Finally, Pragmatists, who prefer action learning, discussion in small groups, problem solving workshops, group tasks where learning is applied should score high on a criterion that measured how well students, working in groups, developed in-depth programs that met practical business needs or how well students perform during internships.

## 5.2 Method

### *Participants and Setting*

Undergraduate students in a professional school of higher educational learning, majoring in human resource management (HRM), participated in a naturally occurring field study. Students studied such HRM issues as how to reintegrate the long term unemployed into the workforce, or how to help employees plan their current careers, as well as the legal steps involved in employee termination. The current HRM program uses a variety of learning environments such as classroom tuition, team projects, and skills training. Also a large amount of time is devoted to internship training in an actual business setting during the second, third and fourth year, and students are required to write a thesis in the last six months of their study.

Students had finished their secondary school education and were generally in the 18 to 22 year age bracket. Data were collected from four cohorts of students (2000-2004,  $n = 34$ ; 2001-2005,  $n = 44$ ; 2002-2006,  $n = 57$ ; and 2003-2007,  $n = 39$ ). Complete academic performance data as well as complete data on the two measures with the LSQ were available of 99 students. Of these, 31% of the participants were male and 69% were female.

A power analysis was conducted to determine the probability of finding an expected effect size. With a posited medium effect size of .30 and an alpha level of .05, a sample size of 99 yields a sufficient power value of .85 (Cohen, 1975).

### *Measures*

*Predictor.* The Learning Style Questionnaire (LSQ) was chosen as a measure of learning style because it, a) has shown face validity, b) has been popular commercially, and c) was originally developed for managers and should therefore generalize to the participants in our study, namely, students who were being trained to become human resource managers. All students completed the LSQ twice; once at the end of their first year (T1) and then again at the end of their third year (T2), both times as part of an Assessment Center. The T1 score was used to compute the LSQ score for each student. LSQ test-retest reliability scores were based on correlations between T1 and T2 scores.

The LSQ by Honey and Mumford (1992) consists of four learning style subscales: Activist, Reflector, Theorist, and Pragmatist; each containing 20 items for a total of 80 items. Respondents either agreed or disagreed with the statements and subscale scores were determined by summing the positive responses, thus yielding a theoretical range of 0 to 20. As noted by Honey and Mumford, the LSQ has been translated into many languages and the Dutch language version was used in this study.

*Criterion.* Academic achievement data were obtained from the student administration office. Students received grades ranging from one to ten for each assessment they completed with a grade of six representing a pass. However, as is typical in the Dutch

education system, students who failed an assessment on their first attempt were permitted to try again. However, only grades of six and above were available to the researchers because first-try failed grades were not recorded by the school.

Five criterion measures were obtained for each student, one each for: classroom lectures, skills training, team projects, internship training, and written thesis. Each criterion score represented an average of multiple assessments throughout the student's four year curriculum. Except where noted, students were appraised by their teachers. Attentiveness to classroom lectures and assigned textbook readings was measured with multiple choice and essay exams. Skills training measured how well students mastered such skills as negotiating and debating, as well as curriculum specific skills such as how to conduct an employment interview. For team projects, students worked in teams of five, each typically lasting ten weeks, to develop HRM programs (e.g. selection and recruitment) for businesses. For internship training, students worked as junior employees in actual business settings in curriculum relevant areas such as recruiting and applying government employment regulations. A grade was given after the workplace supervisor and the student's teacher discussed the student's work. Finally, written theses, typically 30 pages in length, were required in which students discussed practical business problems, often stemming from their internship experiences, and for which they offered potential solutions. Teachers and independent judges rated and discussed the written theses and oral presentations and then graded the work.

The number of independent grades for each student for each criterion was as follows: classroom lectures (25), skills training (10), team projects (10), internship training (3), and thesis (2). In the same way that the mean of a distribution of scores is a more stable index than is a single score from that distribution, obtaining multiple grades provided additional stability for each criterion score.

### **5.3 Results**

Descriptive statistics for all predictor and criterion variables are listed in Table 5.1. Tests for skewness, kurtosis, and Kolmogorov-Smirnov normality were conducted to ascertain whether the criterion data were normally distributed. Results showed that indices of skewness and kurtosis were  $< 1$  for all criteria except thesis. Furthermore, Kolmogorov-Smirnov tests for normality revealed that three of the five criteria were normally distributed, the exception being internship training ( $p = .001$ ) and thesis ( $p = .014$ ).

Table 5.1 *Descriptive statistics predictor and criterion*

	Mean	SD	Range		R <sub>tt</sub>	Skewness	Kurtosis	K-S Z*
			Min	Max				
<u>Predictor</u>								
Activist	9.94	3.22	4	18	.70	.354	-.424	1.182, <i>p</i> = .123
Reflector	12.28	3.65	3	20	.63	-.037	-.571	.961, <i>p</i> = .314
Theorist	11.04	2.54	4	17	.50	-.380	-.059	1.317, <i>p</i> = .062
Pragmatist	10.94	2.80	5	17	.46	.243	-.366	1.070, <i>p</i> = .202
<u>Criterion</u>								
Lecture	6.92	.382	6.3	7.9		.876	.237	1.081, <i>p</i> = .193
Skills	7.19	.419	6.5	8.2		.321	-.565	.622, <i>p</i> = .834
Projects	7.09	.330	6.4	7.9		.126	-.455	.474, <i>p</i> = .978
Internship	7.28	.697	6.0	8.0		-.589	-.896	1.943, <i>p</i> = .001
Thesis	6.80	.775	6.0	8.3		.274	-1.045	1.579, <i>p</i> = .014

\* Kolmogorov-Smirnov Z and p-value

The main hypothesis of this study was tested with bivariate correlations and results are presented in Table 5.2. Contrary to expectations, none of the correlations between learning styles and criterion measures were significant. In an effort to increase the power beyond 90%, 21 non-matriculating students who had dropped out of the program but had completed the first learning style questionnaire (T1), were included so that  $n = 120$ . The analyses were rerun, however none of the correlations were significant.

Table 5.2 *Correlations learning styles and learning environments.*

	Lectures	Skills	Team projects	Internship	Thesis
Activist	-.01	-.04	.00	-.05	.00
Reflector	.15	.13	-.03	-.07	.08
Theorist	.04	.04	.02	-.06	.00
Pragmatist	.05	.02	.02	.01	.05

\*  $p < .05$

Because there were no significant predictive validity results, we tried several tests using mismatched conditions to determine whether the LSQ would differentiate among groups. Several researchers have recommended testing the LSQ using mismatched conditions (e.g. Stellwagen, 2001). Results showed that Activists did not score higher than a combined Reflector/Theorist group on the criteria of internship training,  $t = -.262$ ,  $df = 78$ ,  $p > .05$ ; nor did they do so on the skills training criterion,  $t = .066$ ,  $df = 78$ ,  $p > .05$ . Reflectors did not score higher than a combined Activist/Pragmatist group on the classroom lecture criterion,  $t = .928$ ,  $df = 82$ ,  $p > .05$ ; nor did Theorists score higher than a combined Activist/Pragmatist group on this criterion,  $t = -1.902$ ,  $df = 51$ ,  $p > .05$ . Finally, no

significant results were found, on the criterion of classroom lecture, when students who scored high on either Reflectors or Theorists were combined into one group and compared to a group of students who scored low on either of these orientations,  $t = -.138$ ,  $df = 97$ ,  $p > .05$ .

On a post hoc basis we used an alternative criterion of academic achievement namely time-to-graduation (TTG), that is the months students needed to obtain their diploma. We tested the bivariate correlations between the four learning styles and time-to-graduation and results were non significant with  $r = -.05$  for Activist,  $r = .14$  for Reflector,  $r = .11$  for Theorist, and  $r = .07$  for Pragmatist.

We also classified students into one of the four learning style subgroups according to within-student rank-order scores. Thus if a particular student had the following scores: 7/activist, 13/reflector, 17/theorist, and 12/pragmatist, then he or she would be classified as being a theorist. Results indicated that 20% of the students preferred an activist learning style, with other preferences as follows: Reflector (46%), Theorist (17%), and Pragmatist (17%). We can report positive findings concerning test-retest reliabilities for the LSQ. Test-retest reliabilities for the four subscales, over the two year period, were: Activists ( $r = .70$ ), Reflectors ( $r = .63$ ), Theorists ( $r = .50$ ), and Pragmatists ( $r = .46$ ). These values, taken in context, are quite respectable because test-retest values above  $r = .50$  are normally considered acceptable. Whereas most researchers report acceptable overall test-retest reliabilities ranging from .60 to .90 across a two to eight week time period, with Duff (2000) extending that to one year and reporting an overall value of .74, values in the current study ranged from  $r = .46$  to  $r = .70$  over a two year period.

## 5.4 Discussion

The most important finding of this study is, in essence, a non finding. Despite the use of specific performance criteria tailored to the four learning styles, LSQ subscores do not predict academic achievement nor timely graduation. The lack of findings can not be attributed to unreliability in the LSQ, nor to a lack of power in this research.

That we are unable to report positive findings is perhaps not surprising. Smith, Sekar, and Townshend (2002) provide a review of the literature regarding the matching hypothesis and report that for every study that supports the hypothesis there is a study which rejects it. Such a difference in findings can sometimes be explained by a hidden moderator variable and were one to conduct a meta-analysis, in which studies are grouped according to levels of a moderator variable, one might explain why some studies support whereas others reject a particular finding (Wiersma, 1992). Generalized GPA scores versus specific performance scores may be such a moderator and further research might benefit by grouping studies according to how specific the criteria are, however results of our study are not encouraging.

Limitations in the current study may account for the lack of positive findings and considering our claim of non findings it is important to consider any limitations carefully. An argument can be made about restriction of range in the performance criteria but this argument becomes debatable upon closer review. Theoretical scores on our criteria can range from six to ten, but we acknowledge that a ten is almost never awarded and a nine is also quite rare. Thus the practical range is from six to eight. However, this is not all that different from, for example, the American higher educational system. Grades in the U.S. can range from 'A' to 'F', in which an A is worth four points and an F is worth zero. A straight 'A' student would receive a GPA of 4.0. Here too, the actual grade distribution is restricted because students often drop courses that they are in danger of failing, and grade inflation, embodied in the gentleman's 'C' for subpar work, is common on college campuses. One is also left essentially with a three point scale from A to C. Thus, from a practical point of view, it seems that future field tests of the learning style model may very well be confronted with this fact.

From a methodological viewpoint, restriction of range is a concern because it restricts the amount of variance that can be predicted and therefore the size of obtainable correlations. This we acknowledge, however we wish to note that we obtained multiple grades per criterion which lessens the 'noise' around each individual criterion score. Basing a classroom lecture criterion on 25 grades across a four year curriculum yields a much more stable criterion score than one based on one or two grades because single assessments contain much more error variance. By the same token, a distribution of criterion mean scores reduces the total observable variance but the variance which remains is less effected by error variance.

Honey and Mumford have responded to the lack of empirical support for their theory by stating that the LSQ was never meant to be a psychometric instrument but rather a self-developmental tool that makes people aware of how they learn.

## Chapter 6

### Using multiple samples to replicate findings\*

#### Abstract

A limitation of studies reported in this dissertation concerns the size and specificity of the sample. This limitation is addressed in the current chapter. The purpose of this study is to test whether comparable outcomes are found with respect to first year results of the HRM sample, and whether these results can be generalized to another HRM sample and other samples from different disciplines like health and economics. Based on the literature and earlier empirical findings reported in the previous chapters, a positive association between first year academic performance and intelligence, conscientiousness, intrinsic and study motivation is expected across the samples. This is, in the main, supported by the results of correlation and regression analyses.

\* This chapter is based on:  
Kappe, F.R., & Pluijter, M.M. (2009, May)  
Student characteristics related to first year withdrawal  
and student success across four disciplines.  
Paper presented at the European First Year Experience (EFYE) conference  
“Enhancing the first-year experience: Theory, research, practice”,  
Groningen, the Netherlands.

The purpose of this dissertation is to examine how well intelligence, personality, anxiety, motivation, and learning styles predict academic achievement. Findings from previous chapters have shown that the academic achievement of undergraduate HRM students is moderately associated with intelligence, strongly related with the personality trait conscientiousness, moderately but consistently associated with intrinsic and study motivation, and not related with anxiety or learning style. Although these findings coincide closely with the literature, a convenience sample rather than a random sample was used, therefore generalizability is limited. That limitation is addressed in this chapter. The purpose of the present study is twofold. First, to test whether comparable outcomes are found with respect to first year results of the HRM sample. The second purpose is to test whether these results can be generalized to another HRM sample and other samples from different disciplines. This chapter contains a brief literature paragraph because most of the relevant literature has been addressed in the previous chapters. The discussion centers around studies that distinguish first year performance in higher education from performance in later years, and that report validity coefficients of intelligence and personality variables with academic achievement.

### 6.1 Literature

Predictive validity studies on academic success often focus on the first academic year. This is partly because students are dealing with academic, social, emotional and other challenges as they transition from High School to college. Whereas the majority is able to deal with these challenges, a large proportion of students drop out and leave higher education. The average drop-out rate in the first year in higher education in the Netherlands is about 18% and increases to 22% and 23% in the second and third years respectively. This means that a quarter of the students that enter higher education decide to stop pursuing a degree within three years. Of the students who remain in college only 58% graduate with a diploma within five years (HBO-raad, 2009). These figures have been constant over decades and show that more than a third of all students who enter higher education in the Netherlands fail to obtain their degrees.

#### *The role of cognitive variables in higher education*

Studies have shown that previous performance (e.g., high school GPA) and scores on general tests of scholastic aptitude such as the SAT and ACT, as well as scores on specific and general cognitive ability tests such as the Raven and WAIS, are associated with a variety of student and educational characteristics that include educational aspirations (Neisser et al., 1996) and GPA (Jensen, 1998). In fact, in higher education, intelligence is the most widely used variable in studies on individual differences in student performance (Chamorro-Premuzic, 2007; Deary, Whiteman, Starr, Whalley, & Fox, 2004; Gottfredson,



2002). The association between intelligence and academic success has usually been explained in terms of the ability to learn more efficiently and effectively, which leads to learning enjoyment and successful problem solving (Furnham & Chamorro-Premuzic, 2004). However, research has also shown that the correlation between intelligence and academic performance at the higher levels of education has either been small or failed to reach statistical significance levels altogether. A study by Chamorro-Premuzic and Furnham (2005) revealed that the correlation between intelligence scores and academic performance decreased from .60 to .50, to .40, and to .30, respectively, at the elementary, secondary, university undergraduate, and postgraduate levels. This decline has often been attributed to 'restriction of range'. Because only the brightest students enter higher education there is less variance in IQ scores at the higher levels of education and possibly even less at the later years in college because the less gifted students drop out. Other researchers theorized that the validity of intelligence in higher education tends to decrease over time because the criterion of academic achievement tends to shift over time. At first cognitive abilities such as critical thinking are important, but later personality or motivation are more important (Ackerman, Bowen, Beier, & Kanfer, 2001). Zyphur, Bradley, Landis, and Thoresen (2008) examined the extent to which intelligence and personality predicted initial as well as successive performance and performance change during seven semesters in college. They expected to find a decreasing role of intelligence as a predictor of academic performance and an increase in the importance of conscientiousness. They reported that intelligence and conscientiousness were predictors of initial performance, with bivariate correlations of .38 for intelligence and .26 for conscientiousness with first semester performance. In contrast to their expectations, they found that the effect of intelligence was consistent in all academic semesters, yielding an average correlation of .33. The role of conscientiousness, as expected, increased from .26 in the first semester to .42 in the seventh semester. A meta-analysis by Kuncel, Hezlett, and Ones (2004) showed that the Miller Analogies Test (MAT) had equal predictive validity for first year grade point average ( $\rho = .41$ ) and graduate GPA ( $\rho = .39$ ).

#### *The role of non cognitive factors in higher education*

Of the Big Five personality traits, only conscientiousness has been consistently associated with academic achievement (e.g., Bidjerano & Dai, 2007; Nofle & Robins, 2007). Poropat (2009) conducted a meta-analysis with 70,000 subjects from primary to tertiary levels of education and reported a population correlation of  $r = .24$  between conscientiousness and academic achievement. Correlations for the other four personality factors were much lower. For example, openness to experience and agreeableness correlated .12 and .07 respectively, with grade point average (GPA). These correlations weakened at the higher levels of education to .07 and .06 respectively. Two studies have shown that conscientiousness predicted performance in the first year as well as in successive years (Busato, 1998; Chamorro-Premuzic & Furnham, 2003a), and another

study reported an increase in the importance of conscientiousness as a predictor of performance in the later years in college (Zyphur et al., 2008). The studies by Busato (1998) and Zyphur et al. (2008) reported no evidence of an association between the other Big Five personality traits and first year or successive years academic performance. Only the study by Chamorro-Premuzic and Furnham (2003a) reported a stable association of neuroticism with first (-.28), second (-.31) and third (-.32) year GPA in an undergraduate sample and a significant positive association of agreeableness with first year GPA (.34) but non significant for second (.06) and third year GPA (.07).

There is a consensus amongst educational psychologists that motivational variables play a crucial role in student learning. Some have cited motivation to be the single best predictor of academic success (Arcuri, Daly, & Mercado, 1982) or at least 'half the battle' (Mouw & Khanna, 1993, p. 334). A meta-analysis by Robbins et al. (2004) however reported a mean population correlation of .26 between achievement motivation and GPA and of .11 between achievement motivation and remaining in school. Motivation has often been divided into intrinsic and extrinsic motivation. Several studies have shown that students with higher levels of intrinsic motivation generally outperform those with lower levels (Goldberg & Cornell, 1998; Gottfried, 1990; Mitchell, 1992; Vallerand, Pelletier, Blais, Briere, Senecal, & Vallieres, 1993). Busato (1998) reported a stable association between achievement motivation and performance in year one, two and three in the order of .15. Zyphur et al. (2008) proposed that the importance of motivational variables may increase in the later years in college. However, their suggestion has not yet been validated by empirical results.

The use of learning styles is widely practiced in pedagogy, however researchers have not found sufficient empirical support for an effect of learning styles on performance. One important methodological problem concerns the conceptual confusion of the learning style construct. This has led to a multitude of definitions, theoretical models, and learning style instruments. For some models, such as those by Kolb (1984) and Honey and Mumford (1986), there are studies that show empirical support. However for each study that reports positive findings, there is at least one other study that reports the opposite. In the Dutch higher educational context Vermunt (1992) reports a positive association between the meaning directed learning style and academic achievement and a negative association between the undirected learning style and academic achievement. Busato (1998) also used the Vermunts learning style typology and reported a stable negative association between only the undirected learning style and academic performance in year 1, 2 and 3 at university in the order of -.12. Despite empirical support for some specific learning styles, based on their review of the most used learning style models Coffield, Moseley, Hall, and Ecclestone (2004) conclude that empirical support for a broad association between learning styles and academic performance, e.g. predictive validity, is lacking.

The purpose of the present study is twofold. First, to test whether comparable outcomes are found with respect to first year results of the HRM sample. Based on the literature discussed and the results reported in the previous chapters, a positive association between first year academic performance and intelligence, conscientiousness, intrinsic and study motivation is expected. The second purpose is to test whether these results can be generalized to another HRM sample and other samples from different disciplines. In the additional samples with first year performance data, based on the literature on first year academic achievement, a positive association with intelligence, conscientiousness and motivation is expected and no association with learning styles.

## 6.2 Method

### Samples

*Sample 1.* This sample is the original HRM undergraduate sample. For the purpose of the present study, the data were re-analyzed using grade point average of year 1 (GPA1) as the criterion measure. Complete measures of intelligence as well as personality, motivation, learning styles, and overall GPA were available for 112 students. Of these students, 27% were male and 73% were female.

*Sample 2.* Two additional samples of undergraduate HRM student cohorts for years 2005 and 2006 were available at the Inholland institute in Haarlem. Measures of intelligence, personality, motivation and learning styles were available for 45 students. GPA1 was available as a criterion measure of academic achievement. Of these students 43% were male and 57% were female. The HRM program for the 2005 cohort was reclassified from a social domain to an economic domain, which might explain the larger percentage of male students.

*Sample 3.* This sample contained 95 students in the Sports, Health & Movement program and data were collected at the same institute for higher education, namely Inholland. The dataset included measures of intelligence, personality, motivation and learning styles. Grade point average for year 1 was available. A complete dataset was available for only 65 students due to missing data on the personality questionnaire. Of the participants, 68% were male and 32% were female.

*Sample 4.* The data of sample 4 were collected at the same institute and included 87 students from the school of Economics. The dataset contained measures of intelligence, motivation and learning styles, but not personality. Grade point average for year 1 was available. Of the participants, 59% were male and 41% were female.

*Sample 5.* This sample contained 113 students of the school of Communication of the Inholland institute. The dataset included measures of intelligence, motivation, and learning styles, but not personality. Grade point average for year 1 was available. Of the participants, 27% were male and 73% were female.

### Measures

In all samples, academic achievement was measured using the grade point average of the first year (GPA1), and intelligence was measured with the Multicultural Test of Intellectual Ability for Dutch higher education. In samples 1, 2, and 3, personality traits were measured using the NEO-FFI Personality Inventory (see chapter 3 for full description of the MCT-H and NEO-FFI). In samples 4 and 5, personality traits were not measured. In samples 1 and 2, motivational variables were measured using the “Questionnaire About How One Works” (see chapter 4 for full description) and in samples 3, 4 and 5 using the “Motivation and Learning Style Questionnaire for higher education (MLV-H)”. The MLV-H measures intrinsic, extrinsic and achievement motivation (see chapter 6 for full description). Learning style was assessed in samples 1 and 2 with the Learning Style Questionnaire (LSQ) and in samples 3, 4, and 5 with the MLV-H which measures meaning directed, reproductive, and applied learning styles (see chapter 5 for full description).

### Statistical analysis

Correlation analysis was used to determine the relationship between predictor and criterion measures and in addition multiple stepwise regression analysis was performed to determine the best combination of significant predictors within each sample.

## 6.3 Results

The results concerning the first research question are reported in Table 6.1. The correlation between intelligence and overall academic achievement (GPA) was .22 and with first year performance (GPA1) it was .19. Conscientiousness, intrinsic motivation, and study-motivation were positively associated with both GPA's, with a slight increase over the years for the importance of study-motivation.

Before the correlations were tested with Steiger's (1980) test for correlated data (two-sided) to determine if all pairs of correlations could be regarded as the same, the data were re-analyzed. This was done because first year grades (GPA1) are incorporated in the overall GPA measure. In the re-analysis, GPA was calculated without the first year grades. The re-analysis showed the same significant predictor-criterion associations namely .29 ( $p < .01$ ) for intelligence, .47 ( $p < .01$ ) for conscientiousness, .37 ( $p < .01$ ) for intrinsic motivation and .31 ( $p < .01$ ) for study motivation.

Also the Steiger test had comparable outcomes when GPA1 was or was not included in the overall GPA measure (see Table 6.1). The Steiger test for correlated data (two-sided) confirmed that all pairs of significant correlations had large  $p$  values indicating that the correlations could be regarded as the same. For example, the correlations of intelligence with GPA and GPA1 were .19 and .22 respectively and had an outcome of  $Z = .46, p = .65$ , and in the re-analysis with .19 and .29 had an outcome of  $Z = -.87, p = .39$ . However, significant outcomes for neuroticism ( $p = .01$ ) and agreeableness ( $p = .01$ ) were found indicating that these predictor-criterion correlations were different in the first year compared to overall. For extroversion the predictor-criterion correlation was significant when first year grades were included in the overall GPA measure ( $p = .01$ ), but not when first year grades were excluded from the overall GPA measure ( $p = .28$ ).

Table 6.1 *Correlations between predictors and academic achievement*

	GPA1	GPA	Steiger test	Steiger test
			GPA1 included in GPA	GPA1 excluded from GPA
Intelligence	.19*	.22*	$Z = .46, p = .65$	$Z = -.87, p = .39$
Neuroticism	.18	.02	$Z = 2.41, p = .01$	$Z = 3.02, p = .01$
Extroversion	-.17	.01	$Z = -2.70, p = .01$	$Z = -1.33, p = .18$
Openness to Experience	-.17	-.11	$Z = -.91, p = .36$	$Z = -.61, p = .54$
Agreeableness	-.09	.13	$Z = -3.30, p = .01$	$Z = -2.05, p = .01$
Conscientiousness	.42**	.47**	$Z = -.85, p = .40$	$Z = -.55, p = .58$
Extrinsic motivation	.07	.01	$Z = .09, p = .37$	$Z = -.24, p = .81$
Intrinsic motivation	.43**	.37*	$Z = .99, p = .32$	$Z = 1.06, p = .29$
Study motivation	.22*	.30**	$Z = -1.24, p = .21$	$Z = -.62, p = .54$
Activist	.02	-.06	$Z = 1.20, p = .23$	$Z = -.84, p = .40$
Reflector	.07	-.02	$Z = 1.35, p = .18$	$Z = 1.57, p = .12$
Theorist	.04	.06	$Z = -.30, p = .76$	$Z = -.24, p = .81$
Pragmatist	.04	.04	$Z = .00, p = 1.0$	$Z = -.48, p = .63$

Note. GPA1 = grade point average year 1, GPA = grade point average year 1 to 4. \*\*  $p < .01$ , \*  $p < .05$ .

The results concerning the second research question are reported in Tables 6.2 to 6.5. The correlation coefficients between the various predictor variables and GPA1 are reported in these Tables for the original sample (sample 1) and for the additional samples (samples 2 - 5).

Table 6.2 *Correlations between intelligence and GPA1*

	GPA1					
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	
Intelligence	.19*	.29*	.11	.30**	.30**	$\chi^2 = 2.86, df = 4, p = .58$

\*  $p < .05$ , \*\*  $p < .01$

Intelligence was significantly related to GPA1 in the original HRM sample ( $r = .19, p = .04$ ), the additional HRM sample ( $r = .29, p = .02$ ), the Economics sample ( $r = .30, p = .01$ ) and the Communication sample ( $r = .30, p = .00$ ), however it was not significant in the Sports, Health & Movement sample ( $r = .11, p = .28$ ). Another test for correlated data (two-sided) was used to determine if the predictor-criterion correlations were the same ( $H_0$ ). The outcome of this statistical analysis was non-significant and therefore  $H_0$  could not be rejected. The outcome indicates that the correlations between intelligence and GPA1 can be regarded as being of the same size across the five samples ( $\chi^2 = 2.86, p = .58$ ).

Table 6.3 *Correlations between personality traits and GPA1*

	GPA1					
	Sample 1	Sample 2	Sample 3	Sample 4	Sample 5	
Neuroticism	.18	.14	.15	-	-	$\chi^2 = .07, df = 2, p = .97$
Extroversion	-.17	-.03	-.05	-	-	$\chi^2 = .96, df = 2, p = .62$
Openness	-.17	.08	-.03	-	-	$\chi^2 = 2.15, df = 2, p = .34$
Agreeableness	-.09	-.02	-.05	-	-	$\chi^2 = .17, df = 2, p = .58$
Conscientiousness	.42**	.37**	.29*	-	-	$\chi^2 = 1.08, df = 2, p = .58$

\*  $p < .05$ , \*\*  $p < .01$ . Note. - indicates that the data were not collected

Conscientiousness was significantly related to GPA1 in the original HRM sample ( $r = .42, p < .01$ ), the additional HRM sample ( $r = .37, p = .01$ ), and the Sports, Health & Movement sample ( $r = .29, p = .02$ ). The significance of the differences of the correlations of conscientiousness with GPA1 across the three samples was tested and the outcome was non significant indicating that the correlations between conscientiousness and GPA1 across the three samples are the same size ( $\chi^2 = 1.08, p = .58$ ).

Table 6.4 *Correlations between motivation and GPA1*

	GPA1					
	Sample 1 (WSV)	Sample 2 (WSV)	Sample 3 (MLV-H)	Sample 4 (MLV-H)	Sample 5 (MLV-H)	
Intrinsic	.43**	.14	.04	-.07	-.09	$\chi^2 = 20.01, df = 4, p = .00$
Extrinsic	.07	-.02	-.26**	-.14	-.01	$\chi^2 = 6.46, df = 4, p = .17$
Study	.22*	.18	-	-	-	$\chi^2 = .22, df = 1, p = .82$
Achievement	-	-	-.09	-.22	-.04	$\chi^2 = 1.66, df = 2, p = .43$

Note. \*  $p < .05$ , \*\*  $p < .01$ . - indicates that the data were not collected

Intrinsic and study motivation was significantly related to GPA1 in the original HRM sample. However, the correlation was not significant in the additional HRM sample, nor in other samples in which a different questionnaire was used. A test of the correlations of intrinsic motivation with GPA1 across the samples confirmed the difference between the correlations ( $\chi^2 = 20.01, p = .00$ ).

The results in Table 6.5 show that learning styles were neither related to GPA1 in the original sample, nor to GPA1 in the other samples.

Table 6.5 *Correlations between learning styles and GPA1*

	GPA1					
	Sample 1 (LSQ)	Sample 2 (LSQ)	Sample 3 (MLV-H)	Sample 4 (MLV-H)	Sample 5 (MLV-H)	
Activist	.02	.04	-	-	-	$Z = -.11, p = .91$
Reflector	.07	.07	-	-	-	$Z = .00, p = 1.0$
Theorist	.04	.27	-	-	-	$Z = -1.29, p = .20$
Pragmatist	.04	.22	-	-	-	$Z = -1.00, p = .37$
Meaning	-	-	.05	-.03	-.03	$\chi^2 = .40, df = 2, p = .82$
Reproductive	-	-	-.01	-.18	-.09	$\chi^2 = 1.30, df = 2, p = .52$
Applied	-	-	-.08	-.12	.11	$\chi^2 = 3.05, df = 2, p = .22$

Note. \*  $p < .05$ , \*\*  $p < .01$ . - indicates that the data were not collected

Multiple stepwise regression analyses were performed for each sample. Conscientiousness, neuroticism, and intelligence together explained 27% of the variance in GPA1 in sample 1,  $F(3,100) = 12.09, p = .00$ . Intelligence and conscientiousness explained 24% of the variance in GPA1 in sample 2,  $F(2,41) = 6.31, p = .01$ . Conscientiousness, intelligence, and extrinsic motivation explained 21% of the variance in GPA1 in sample 3  $F(3,61) = 5.24, p = .00$ . Intelligence explained 9% of the variance in GPA1 in samples 4 and 5 with  $F(1,67) = 6.72, p = .01$  and  $F(1,93) = 9.49, p = .00$  respectively.

## 6.4 Discussion

The purpose of this study is to test a) whether comparable outcomes are found with respect to first year results of the HRM sample and b) whether these results can be generalized to another HRM sample and other samples from different disciplines. Based on the literature and earlier empirical findings reported in the previous chapters, a positive association between first year academic performance and intelligence, conscientiousness, intrinsic and study motivation is expected across the samples.

The results in Table 6.1 show positive associations between intelligence, conscientiousness, intrinsic motivation, and study motivation on the one hand and GPA and GPA1 on the other. The Steiger test for correlated data reveals that the correlations are not significantly different from one another. This indicates that intelligence, conscientiousness, intrinsic motivation, and study motivation are stable predictors of students' performance from the start to finish in the HRM sample. This result is in line with earlier studies that report that the effects of intelligence, conscientiousness (e.g., Zyphur et al., 2008), and motivation (Busato, 1998) are stable throughout the four year curriculum. Intelligent, motivated, and conscientious students perform better in the first year and keep on performing better during successive years.

A second purpose of this study is to test whether results using first year data can be generalized to other samples, and this is the case in the main. Intelligence is positively related to GPA1 in the original HRM sample (.19,  $p < .05$ ) and in the additional HRM sample (.29,  $p < .05$ ). In the three samples from other disciplines intelligence is also positively related to academic achievement, with especially strong and significant correlations in samples 4 (.30,  $p < .01$ ) and 5 (.30,  $p < .01$ ). Conscientiousness is, as expected, positively related to first year academic achievement in all samples in which personality traits are measured. Intrinsic and study motivation are positively related to GPA1 in the original HRM sample, however these correlations are not replicated in the additional HRM sample. There is also little evidence of a link between motivation and GPA1 in samples 3, 4, and 5, however motivation in these samples was measured with a different questionnaire. It would be interesting to examine these motivational variables in relation to academic achievement in the later years and test Zyphur et al's. (2008) idea that motivation plays an increasingly important role during successive college years. There is no evidence of an association between learning styles and GPA1 in any of the additional samples.

Multiple stepwise regression analysis is used to examine if the same combination of predictors across the first year samples will occur. Conscientiousness is the most prominent predictor of academic achievement in samples 1, 2, and 3, where information about personality traits is available. Intelligence is the first or second most important predictor of academic achievement in four of the five samples (1, 3, 4, 5). Motivational variables explain a significant percentage of variance in academic achievement only in



sample three. The results of the various multiple regression analyses are consistent with the literature concerning the roles of intelligence and personality in predicting academic achievement (e.g., Poropat, 2009), motivation (Robbins et al., 2004) and learning styles (e.g., Coffield et al., 2004).

A limitation of the current study is the use of different measurement instruments across the samples. In the case of intelligence the same measuring instrument was used (MCT-H), but for the other constructs two different measuring instruments were used.



## Chapter 7

### Academic assessment centers and student competencies\*

#### Abstract

This chapter reports on the use of assessment centers (AC) in higher education. Students (N = 137) in the human resource management program participated in two AC's; once at the end of their first year and then again at the end of their third year of study. They were assessed on nine competencies. In addition, the following measures for each student were available: GPA, intelligence, personality, motivation, and learning style. Results show that, a) the nine competencies could be collapsed into three competency dimensions and an overall factor, b) students received higher AC ratings in their third year, c) the correlation between GPA and AC ratings was significantly higher for third, compared to first, year students, and d) students perceived the AC experience as being helpful to their learning. Whereas AC's measure broader competencies that are learned throughout the curriculum, traditional exams measure short term, task or course related performance. Because AC's can be cost effective, educational institutions should consider using existing exams and AC's alongside one another.

\* Part of this chapter featured as:

Kappe, F. R. (2009, May)

*Quantitative versus Qualitative criteria for assessing student success*

Poster session presented at the annual ORD dagen [Educational Research Days],

Leuven, Belgium.

## 7.1 The assessment center method

Traditional exams offer a number of important advantages: they are widely used and understood, easy and cheap to administer, generally presumed to be reliable and valid, and yield standardized grades that make it easy to compare students within and across schools. Researchers often use a student's average score (GPA) as an overall measure of academic achievement. GPA reflects how well students performed on different, often short, and highly content-specific exams. However, as a consequence of competence-based education which emphasises the development of skills and competencies, new methods of assessing student performance are needed (Bartels, Bommer, & Rubin, 2000; Dochy, Segers, & De Rijdt, 2002; Riggio, Mayes, & Schleicher, 2003). The purpose of the present study is to examine the efficacy of using the AC methodology as a means for evaluating undergraduate student competencies.

The assessment center (AC) refers to a method that involves, a) a set of competencies upon which individuals are assessed, b) a set of exercises to measure the competencies, and c) a team of assessors to evaluate the candidates. Individual role plays assess interpersonal communication, sensitivity, or client focus. The in-basket exercise assesses problem solving, decisiveness, and stress tolerance. Group exercises assess leadership and the ability to work in a team. Tests of cognitive ability and personality are also often administered during an AC. Psychologists and upper level managers often work together to assess candidates.

The AC is in fact a standardized test. That is, although the hallmark of an AC is its behavioral or performance-based exercises, the AC is administered, and exercises are scored, under standardized conditions (Waldman & Korbar, 2004). Assessment centers can measure an array of competencies that cut across various learning modules within a curriculum. Additionally, AC's can provide students with meaningful behavioral feedback about their strong and weak competencies, and counselling can provide recommendations for subsequent personal development. Organizations around the globe have used AC's for over 60 years to select and develop employees. Mayes (1997) reported that an estimated 80% of *Fortune* 500 companies used AC's somewhere in their organizations. Previous research has substantiated the reliability, validity, and fairness of AC's for management progression and job satisfaction. However, a recent review by Thornton and Gibbons (2009) concerning the current state of affairs of AC's in business settings showed that several methodological, instrumental, and practical issues still need further clarification. These include a) the number of unique dimensions measured in an AC, b) the relationships of AC ratings with other criteria, c) how to build exercises that maximally elicit behavior, d) the level of standardization, e) the effect of retesting and reusing exercises, and f) the different ways of providing feedback.

This chapter considers a number of important issues, that parallel those mentioned by Thornton and Gibbons (2009), when an AC is used with students in an academic

learning environment. The first issue concerns the number of distinct dimensions that are actually being measured by the assessors. This is especially important because, unlike AC's in business settings in which the focus is on selection, a primary purpose of an academic AC is to provide students with feedback. A second issue is whether an AC, when repeatedly applied, can effectively measure a student's personal development over several years. Third is the issue of how well AC performance ratings correlate with other measures of academic performance such as GPA, or with cognitive ability tests and personality inventories (Bartels et al., 2000). The final important issue concerns how students evaluate their learning experience in an AC. The literature for each of these issues is reviewed and discussed below.

#### *7.1.1 Internal structure of the AC*

The internal structure of AC ratings, that is the number of distinct dimensions that are captured by the exercises, has puzzled researchers and Collins, Schmitt, Sanchez-Ku, Thomas, McDaniel, & Le (2003) have noted that there is sparse empirical evidence of construct validity for AC ratings. Thornton and Gibbons (2009) have argued that if it can not be proven that the AC measures specific, separate dimensions, then giving dimensional, competency-specific feedback to candidates will be questionable.

A common strategy for determining the internal structure of AC dimensions has been through the multi-trait, multi-method approach (Campbell & Fiske, 1959; Thornton & Gibbons, 2009). When applied to AC's, dimensions are analogous to traits and exercises are analogous to methods. For AC ratings to be valid, ratings of the same dimension across different exercises should be highly correlated and show convergent validity, whereas ratings of different dimensions within a single exercise should be correlated to a much lesser extent and thus show discriminant validity. Empirical support of convergent and discriminant validity in AC's has been lacking (Jansen & Stoop, 2001; Thornton & Gibbons, 2009).

However, there is a growing consensus that the multi-trait, multi-method model, which treats exercise effects as measurement error, is not appropriate for examining the construct validity of AC's (Thornton & Gibbons, 2009). It has been suggested that both dimensions and exercises may have meaningful roles to play when determining overall assessment center performance. That is, inconsistencies in the behavior of a candidate across exercises may in fact represent real variation in performance and that variation may be linked to job-relevant variables. Exploratory and confirmatory factor analyses have also been used to investigate the properties of AC dimensions (Kleinmann & Köller, 1997).

#### *7.1.2 Learning competencies*

Using an AC more than once can help determine how well students learn competencies over time. Providing feedback at multiple points helps students develop their weaknesses and helps those who design school curricula because it measures the

effects of curriculum change. Only one study has used a longitudinal design in which students repeatedly participated in an AC. Ebert (2009) assessed students' project management knowledge and competencies and then did so again 10 weeks later. Results indicated that students' knowledge had increased significantly with an effect size ( $d$ ) of 1.10. Students' competencies also increased, but the increase was marginal with an effect size ( $d$ ) of .57. In a related study, Riggio et al. (1997) used a cross-sectional design and reported that senior year undergraduates performed significantly better than freshmen in an academic AC. The authors regarded this result as evidence of the validity of the AC because one would expect seniors, who have had more years of education, to outperform freshman who are just starting their education. It would be interesting to replicate the longitudinal study by Ebert (2009) and to use a longer time interval than 10 weeks.

### 7.1.3 *AC ratings and GPA*

Bartels et al. (2000) examined whether academic AC ratings were related to traditional measures of student performance. Using a sample of 347 undergraduate students, they reported a mean correlation of .23 between an overall assessment center rating (OAR) and GPA, of .24 between the OAR and specific course exam grades, and of .13 between the OAR and grades on group discussions. Riggio et al. (2003) also reported that the AC was correlated with GPA ( $r = .27, p < .01$ ) in a sample of 754 undergraduate business students. However Waldman and Korbar (2004) reported a nonsignificant correlation between an OAR and GPA in a sample of 66 undergraduate business students. They reasoned that GPA captured predominantly students' knowledge and that the behavioral component, as is measured by an AC, was absent in GPA.

### 7.1.4 *AC ratings, intelligence, and personality*

Research in organizations has shown that AC ratings are correlated with cognitive ability tests and with personality inventories. A meta-analysis by Collins et al. (2003) reported, based on 65 correlations, the following (corrected) correlations with OAR's, .67 for intelligence, .17 for agreeableness, .50 for extroversion, .35 for emotional stability, and .25 for openness to experience. However a caveat is in order. The test scores are often read by assessors which may bias their ratings of candidates on the AC exercises. If assessors heavily weigh intelligence and personality in their ratings, then correlations between the OAR, and intelligence or personality, will be artificially inflated. Collins et al. (2003) have advised researchers to use a design in which test scores on intelligence and personality are not read in advance by assessors.

Correlations between AC ratings and intelligence are often explained by the inclusion of an in-basket exercise which has a strong analytical and problem solving component. Correlations between AC ratings and personality can be explained by the fact that personality influences the way an individual behaves across different situations. Assessors observe a candidate's behavior in multiple exercises and therefore may indirectly be

obtaining information about the candidate's personality. In other words, the behaviors that assessors observe during AC exercises are being influenced or shaped by individual differences in fundamental traits (Collins, Schmitt, Sanchez-Ku, Thomas, McDaniel, & Le, 2003).

A few studies have been conducted in an educational setting. Bartels et al. (2003) used a sample of 347 business students and reported a correlation of .24 ( $p < .01$ ) between an OAR and intelligence. Furthermore, Riggio et al. (1997) reported that students, who scored higher on need for achievement and dominance, performed significantly better on the in-basket, the mock hiring interview, and in leaderless group discussions. Correlations between personality and AC exercises were in the range of .20 to .33. Conversely, Waldman and Korbar (2004) reported nonsignificant correlations between an OAR and measures of personality. However none of the cited studies stated clearly if and how the test scores of cognitive ability and personality were used by assessors while assessing candidates on AC exercises. In sum, there is some evidence that AC ratings, in business as well as in educational settings, correlate with measures of intelligence and personality, however these results are mostly based on studies in which personality scores were likely incorporated into the AC ratings.

#### 7.1.5 *Relevance of participating in an AC*

Thornton and Gibbons (2009, p.173) have noted that with the evolution of the holistic view of validity, there has been a new emphasis on the consequences of assessment. Consequences in this context mean the intended and unintended outcomes of the assessment for student learning (Dierick, Van de Watering, & Muijtjens, 2002). For example, classroom exams are known to have a so called backwash effect (Biggs, 2003), which is regarded as a negative consequence because students focus their learning only on what is measured by the test. In contrast to *backward* feedback, the AC has the potential to offer *feed-forward* feedback concerning one's strong and weak competencies. A study by Riggio et al. (1997) reported that 60% of the students indicated positive consequences of having participated in an AC, namely, "being better prepared for the business world". In a follow-up study, Riggio et al. (2003) reported that 92% of students that participated in the AC rated it as a "valuable learning experience".

#### 7.1.6 *Research questions*

Data for the current research came from college students who had participated in an assessment center during their first year (AC1) and then again second during their third year (AC2). Several research questions were investigated. First, it was important to assess the psychometric properties of the competency ratings. That is, will the ratings show multiple and distinct competencies, a smaller set of dimensions, or perhaps a single overall factor? Second, will students exhibit performance increases from AC1 to AC2? Third, do AC ratings correlate with other measures of academic success such as grades on exams,

internships, and written theses? Fourth, how well do AC ratings correlate with measures of intelligence and personality? Finally, how will students evaluate their experience in an AC?

## 7.2 Method

### *Sample*

Undergraduate students in a professional school of higher educational learning, who were majoring in human resource management (HRM), participated in this study. The HRM program used a variety of learning environments such as classroom tuition, team projects, and skills training, which promoted competence-based learning. Much time was devoted to internship training in actual business settings during the second, third, and fourth years. Students were also required to write a thesis. Students were in the 18 to 22 year age bracket, and data were collected from four cohorts of students. The total sample of 148 student candidates was reduced to 137 due to missing data.

Fourth year HRM students, who were about to graduate, acted as assessors in the AC. Twenty-two student assessors were trained in 2000, 24 in 2001, 22 in 2002, 32 in 2003, 39 in 2004, and 43 in 2005. Assessors worked in two-person teams and assessed a single student candidate during an entire day. The student assessors worked between three to six days, and were unaware of their candidates' test scores on intelligence, personality, motivation, or learning style.

### *Measures*

*AC ratings.* Students were assessed on nine different competencies. An organization that specialized in AC's for students designed and implemented the AC procedure and exercises, and they were joined by faculty members from the HRM program. The AC consisted of various simulated work exercises such as role plays, group exercises, written assignments, and a self-reflection exercise. The assessors used five point based, behaviorally anchored rating scales (BARS) to rate competencies such as analyzing, communicating, and advising. The norm for students in AC1 was a level 3 rating, whereas the norm in AC2 was a level 4 rating. At graduation students should perform conform a level 5 rating. The two assessors observed their candidate during the exercises and chose the BAR descriptions that best described the student's behaviors for each competency. Intraclass correlation coefficients for the nine competencies were all  $>.80$  for AC1 and  $>.70$  for AC2, indicating sufficient interrater reliability of the AC ratings. Steps were taken to motivate the student candidates. Although participation was obligatory, all students received one credit point (ECTS), a written assessment report, and 30 minutes of oral feedback from both assessors a week after participating in the AC. Candidates and assessors alike completed evaluation forms at the end of each AC. Candidates answered questions about the relevance and difficulty of the AC exercises, the treatment by the



assessors, the quality of the feedback, and the relevance of the AC exercises. Assessors answered questions about the assessor training, usefulness of the AC for future work, and how they valued the AC experience as a whole.

*GPA.* An overall grade point average was calculated, as well as GPA's within a particular year. The overall GPA was the weighted cumulative average of grades across all subjects during the student's entire tenure at school. GPA1 and GPA2 were the weighted cumulative averages of all grades in the student's first and third years of study. Grades were based on academic performance in five different types of learning environments: classroom lectures, skills training, team projects, internship training, and written thesis.

The curriculum stimulated students to develop competencies as they worked through various learning modules. For example, team projects named "labor relations" and "reintegration into the work force" emphasized learning about the labor market and equal employment opportunity legislation; helped students acquire interpersonal and communications skills; and taught problem-solving and decision-making skills. Internships provided further opportunities to practice competencies in an actual place of work. However, no single learning module, whether internship, lecture, or team project, can assess the entire set of competencies; as is possible with an assessment center.

*Intelligence.* Cognitive ability tests and personality inventories were administered during the first assessment center. The short version of the Multiculturele Capaciteiten Test voor Hoger Niveau (Multicultural Test of Intellectual Ability for Dutch higher education) measured intelligence. This test is widely used in Dutch higher education and has been regarded as psychometrically sound by the Dutch Testing Committee (Evers, Van Vliet-Mulder, & Groot, 2000). The following three subtests were used with corresponding alpha's for the current study: Double Word Analogies ( $\alpha = .95$ ), Exclusion ( $\alpha = .84$ ), and Mental Speed ( $\alpha = .96$ ). This shorter version of the MCT-H was used because of time constraints, to minimize test fatigue, and because it has been shown to correlate as high as  $r = .86$  with the long version, whose overall reliability is .97.

*Personality.* The Dutch version of the NEO-FFI was used to assess the Big Five personality traits: conscientiousness, neuroticism, extroversion, agreeableness, and openness to experience. Cronbach's alphas for the current study were: .84, .83, .77, .64, .72, respectively.

*Motivation.* Motivation was measured with the "Questionnaire About How One Works" by Akkerman and Van der Linden (2000) which was based on earlier work by Hermans (1967). This questionnaire measures intrinsic motivation ( $\alpha = .79$ ), anxiety ( $\alpha = .86$ ), environmental press ( $\alpha = .84$ ), need for status ( $\alpha = .86$ ), and motivation-to-study ( $\alpha = .76$ ).

*Learning style.* Learning style was measured with the Learning Style Questionnaire (LSQ) of Honey and Mumford (1992). The LSQ consists of four learning style subscales: Activist, Reflector, Theorist, and Pragmatist that registered alphas in the current study of .70, .63, .50, .46, respectively.

*Relevance.* Students were asked about their experiences with the AC exercises like whether they felt the AC exercises were realistic and if the feedback had been useful. There were five items in the scale, and sample items were: “Did the AC provide adequate information about your performance and personal development?”, “Was the AC useful?”, and “Are you now able to write a better personal development plan?”. Students scored each statement on a five point scale from 1 (*totally disagree*) to 5 (*totally agree*). A single overall item asked, “Overall, how would you rate the AC experience?” and was scored on a 10 point scale from 1 (*very disappointing*) to 10 (*excellent*). The relevance scale registered an alpha of .80, indicating sufficient reliability.

### 7.3 Results

The findings are divided into three sections. First the internal structure of the AC is investigated, and differences in student performance between the first and second assessment center are examined. Then correlations of AC ratings with other measures such as grades, intelligence, and personality are explored. Finally, the results concerning the relevance of the AC for student learning are presented.

#### 7.3.1 *Internal structure and competency learning*

An important research question concerned the internal structure of the AC and whether nine distinct competencies were actually being measured. Separate factor analyses were conducted on the ratings from AC1 and AC2. As shown in Table 7.1, exploratory factor analysis (EFA) revealed a three factor structure in both assessment centers, and with one exception, the loadings of the various competencies on the factors were comparable. A three factor structure is often found in assessment centers measuring a comparable amount of competencies. The taxonomy of Thinking, Feeling, and Firmness is well known in the Dutch assessment practice (see Kolk, 2001). This taxonomy was used to label the factors: 1) Thinking (analyzing, communicating, advising), 2) Feeling (creativity, teamwork, justifying), 3) Firmness (presenting, reporting, learning goals).

Table 7.1 Factor analysis<sup>a</sup> for AC1 and AC2 competency ratings

	First year assessment center (AC1)			Third year assessment center (AC2)		
	1	2	3	1	2	3
Creativity	<b>.627</b>	.468	-.166	<b>.776</b>	.167	.092
Teamwork	<b>.722</b>	.079	.243	<b>.781</b>	.033	.025
Justifying	<b>.790</b>	.220	.174	<b>.716</b>	.244	.117
Analyzing	.010	<b>.883</b>	.145	.080	<b>.838</b>	.134
Advising	.248	<b>.838</b>	.009	.215	<b>.803</b>	.143
Communicating	.318	<b>.686</b>	.156	.175	<b>.845</b>	-.034
Reporting	-.129	.177	<b>.803</b>	.108	-.041	<b>.902</b>
Learning goals	.251	-.091	<b>.660</b>	.092	.351	<b>.518</b>
Presenting	.318	.180	<b>.545</b>	<b>.707</b>	.102	.074

<sup>a</sup> Extraction Method: Principal Component Analysis. Rotation Method: Varimax with Kaiser Normalization

As has been advised by Kleinmann and Köller (1997), confirmatory factor analysis was also performed. Several models were tested using Lisrel (Jöreskog & Sörbom, 1988). The best fit for data from each AC was found for a model consisting of only one competency factor (see Appendix B). Previous factor analytic studies have also reported one underlying competency factor (Sackett & Harris, 1988; Turnage & Muchinsky, 1982). The internal structure of the AC ratings was assessed further by considering convergent and discriminant validity. Table 7.2 lists the correlations among all competency ratings. As expected, the ratings showed significant correlations within and across the AC's. The same competencies across AC's did not show higher correlations (convergent validity) than different competencies within or across AC's (discriminant validity).

Table 7.2 Descriptives and intercorrelations of AC1 and AC2 ratings

	AC2		1	2	3	4	5	6	7	8	9	Mean	SD
AC1													
1 Analyzing			(.08)	.29**	.68**	.17	.49**	.18*	.25**	.23	.01	2.9	.76
2 Creativity			.29**	(.23)	.50**	.26**	.41**	-.03	.47**	.13	.09	2.7	.64
3 Advising			.68**	.50**	(.29)	.24**	.49**	.08	.37**	.18*	.11	2.8	.66
4 Teamwork			.17	.26**	.24**	(.22)	.32**	.17	.46**	.23**	.14	3.0	.57
5 Communicating			.49**	.41**	.49**	.32**	(.23)	.17*	.36**	.26**	.08	3.0	.68
6 Reporting			.18*	-.03	.08	.17	.17*	(.12)	.14	.23**	.29**	2.9	.56
7 Justifying			.25**	.47**	.37**	.46**	.36**	.14	(.23)	.31**	.16	2.9	.55
8 Presenting			.23*	.13	.18*	.23**	.26**	.23**	.31**	(.06)	.22**	2.9	.63
9 Learning goals			.01	.09	.11	.14	.08	.29**	.16	.22*	(.00)	2.9	.65
Mean			3.6	3.4	3.7	3.7	3.7	3.5	3.5	3.6	3.7		
SD			.70	.71	.69	.65	.67	.78	.60	.68	.67		

\*\*  $p < .01$ , \*  $p < .05$

Based on the results of the confirmatory factor analysis, OAR scores were used to determine whether students showed improvement from AC1 to AC2. The overall level of competency of students increased from an average of 2.9 ( $sd = .38$ ) in AC1 to an average of 3.7 ( $sd = .37$ ) in AC2. This increase was significant when tested with a paired sample (two sided) t-test,  $t = 19.39$ ,  $df = 113$ ,  $p < .01$ . As expected, students developed their competencies in the two year time lag, probably by participating in relevant courses, training sessions, projects and internships, or the first year AC experience. Although the students' competencies increased, they in fact underperformed according to the *norm* set by faculty. The norm for students in AC1 was a level 3 rating, whereas the norm in AC2 was a level 4 rating. Single sample t-tests showed a significant difference between the mean observed ratings and the expected (norm) ratings for year AC1,  $t = -3.09$ ,  $df = 136$ ,  $p = .01$ , and for AC2,  $t = -10.87$ ,  $df = 147$ ,  $p < .01$ .

### 7.3.2 Relatedness with other measures

Correlations among the three AC dimensions, the OAR, and annual GPA's are presented in Table 7.3. First year grades were not related to first year AC performance ( $r = .11$ ,  $p = .22$ ), however third year grades and third year AC performance were related ( $r = .25$ ,  $p = .02$ ).

Table 7.3 *Intercorrelations among AC dimensions, OAR, and GPA per year*

			1	2	3	4	5	6	7	8	9
AC1	1	Thinking	-								
	2	Feeling	.50**	-							
	3	Firmness	.25*	.27*	-						
	4	OAR1	.83*	.78**	.63**	-					
	5	GPA1	.07	.02	.16†	.11	-				
AC2	6	Thinking	.25**	.24*	.13	.28**	.14	-			
	7	Feeling	.18†	.23*	.03	.20*	.19*	.36**	-		
	8	Firmness	.17†	.20*	.10	.21*	.30**	.35**	.44**	-	
	9	OAR2	.28**	.34**	.12	.33**	.26**	.79**	.73**	.70**	-
	10	GPA2	.16†	.17†	.15	.20*	.50**	.15†	.12	.34**	.25**

Note. OAR1 and OAR2 = overall assessment ratings for AC1 and AC2, respectively; GPA1 = overall grade point average of year 1; GPA2 = overall grade point average of years 1 to 3. \*\* $p < .01$ , \* $p < .05$ ,

†  $p < .10$ .

The results in Table 7.4 show that the overall OAR, which is the arithmetic average of the OAR1 and OAR2 ratings, was related to all five specific student performance criteria and had the strongest correlation with internship training ( $r = .32, p < .01$ ). Firmness, which is the arithmetic average of the firmness rating of AC1 and firmness rating of AC2, correlated significantly with all five specific performance measures. The OAR significantly correlated with overall GPA ( $r = .41, p < .01$ ).

Table 7.4 *Intercorrelations among competency dimensions, OAR, the specific GPA's, and overall GPA*

		1	2	3	4	5	6	7	8	9
1	Thinking	-								
2	Feeling	.52**	-							
3	Firmness	.44**	.50**	-						
4	OAR	.84**	.82**	.77**	-					
5	Lectures	.12	.10	.27*	.20*	-				
6	Skills	.17*	.14	.20*	.21**	.65**	-			
7	Projects	.09	.09	.29**	.19*	.47**	.31**	-		
8	Internship	.25**	.21*	.35**	.33*	.42**	.45**	.42**	-	
9	Thesis	.09	.13	.26**	.19*	.32**	.29**	.33**	.39**	-
10	GPA	.29**	.27**	.46*	.41*	.66**	.56**	.63**	.88**	.65**

Note. OAR = the average overall assessment rating across two assessment centers; GPA = overall grade point average. \*\* $p < .01$ , \* $p < .05$ .

Table 7.5 lists the correlations among the three AC dimensions, the overall OAR, intelligence, personality, motivation, and learning style. The OAR was related to the personality traits of openness to experience ( $r = .20, p = .02$ ), and conscientiousness ( $r = .25, p < .01$ ), as well as to intrinsic motivation ( $r = .24, p < .01$ ) and to study motivation ( $r = .24, p < .01$ ), however it was unrelated to intelligence ( $r = .07, p = .39$ ). Students who are generally more disciplined, curious, and motivated do better in academic AC's.

### 7.3.3 *Relevance of the academic AC*

The student assessors evaluated their training as well as their experience with the AC program in general. Of the 149 responding assessors, 97% indicated that they had learned new skills in the assessor training. The assessors gave the AC program a 7.4 ( $sd = .63$ ) on a 10 point scale. A small percentage (11%) of the student assessors thought that faculty and staff members would have made better assessors.

Student candidates were also asked about their experiences with the AC exercises. Results indicated that students perceived the AC program and its exercises (e.g., role play, group discussion, in-basket, presentation) as being useful for providing information about their level of competency. They also felt that the AC was realistic and that competencies

measured during the AC would be required in their future profession. The students felt positive about the potential of the AC for future learning purposes ( $M = 3.9$ ,  $sd = .69$ ). Additional items with high mean scores were, "I found it useful to participate in the assessment center" ( $M = 4.1$ ,  $sd = .82$ ); and "The AC yields additional information concerning my performance level and development" ( $M = 3.9$ ,  $sd = .96$ ). Ninety-three percent of the students agreed with the statement that the AC should not be deleted as part of the curriculum. Overall, the candidates gave the AC program a 7.4 ( $sd = .77$ ) on a ten point scale. However, similar to the student assessors, 11% felt that faculty and staff members would have made better assessors.

Table 7.5 *Intercorrelations among competency dimensions, OAR, intelligence, personality, motivation, and learning style*

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1 Thinking	-																	
2 Feeling	.52**	-																
3 Firmness	.44**	.50**	-															
4 OAR	.84**	.82**	.77**	-														
5 Intelligence	.00	.04	.14†	.07	-													
6 Neuroticism	.06	.12	.13	.12	.09	-												
7 Extroversion	.02	.04	.01	.03	-.13	-.43**	-											
8 Openness	.18*	.19*	.12	.20*	-.12	.03	.12	-										
9 Agreeableness	.08	.08	-.09	.04	-.04	-.20*	.19*	.09	-									
10 Conscientiousness	.14†	.19*	.30**	.25**	.01	-.14	.01	-.11	.23**	-								
11 Intrinsic motivation	.12	.14†	.36**	.24**	.11	.08	.04	.00	.03	.47**	-							
12 Anxiety	-.04	-.14†	-.23**	-.15†	.02	.34**	-.27**	-.04	-.05	-.21**	.06	-						
13 Need for pressure	.08	.02	-.01	.04	.10	.02	.12	.18*	-.20*	-.23**	-.03	.19*	-					
14 Need for status	-.03	-.02	.02	-.02	.10	.21*	.01	.03	-.29**	-.06	.15†	.32**	.25**	-				
15 Study motivation	.09	.17*	.35**	.24**	.06	.15†	-.02	.25**	-.10	.32**	.65**	.07	.06	.23**	-			
16 Activist	.08	-.01	.14†	.08	.05	-.11	.41**	.10	-.24**	.00	.12	-.29**	.24**	.11	.03	-		
17 Reflector	.06	.03	.05	.06	-.03	.16†	-.26**	.16†	.14†	.15†	.11	.24**	-.12	.01	.18*	-.60**	-	
18 Theorist	.07	.12	.15†	.14†	-.02	.13	-.13	.08	-.12	.25**	.23**	.07	-.04	.22**	.30**	-.20*	.40**	-
19 Pragmatist	.04	.11	.18*	.13	-.02	-.08	.09	.01	-.32**	.16†	.18*	-.09	.10	.26**	.23**	.20**	-.07	.54**

Note. OAR = overall assessment rating; \*  $p < .05$ , \*\*  $p < .01$ , †  $p < .10$



## 7.4 Discussion

The present study addresses four key issues concerning how to use assessment centers in higher education. The main findings are: a) competency ratings tend to cluster into three AC dimensions and one overall factor, b) third year students perform better than first year students, c) overall AC ratings are meaningfully correlated with overall GPA, and AC ratings are especially related to academic performance criteria that have a more practical orientation, and d) candidates and assessors alike are positive about their experiences with the AC.

The issue of internal structure is important because it influences the form in which candidates receive their feedback. The AC feedback is separate and distinct from the feedback that students receive from their normal coursework. Student candidates say that the behavioral feedback at the end of the AC is rich and they feel good about being able to subsequently practice their competencies within the relative safety of the college environment. Converging data from multiple sources such the classroom and the AC validates for students how they are seen by their instructors as well as by their peers (Kottke & Schultz, 1997).

The internal structure of the competencies in the current research is best expressed as three dimensions, or as one overall factor, a finding that is in line with previous studies. Critics argue that giving feedback on specific dimensions is misleading when one is unsure of how many dimensions are actually being measured (Sackett & Dreher, 1982). However there are several reasons to continue using dimensions in an academic AC. First, dimensions are not exclusive to the AC method, but are widely used in human resource management functions such as in performance appraisal (Rupp, Thornton, & Gibbons, 2008). Second, the alternative of basing feedback on AC exercises instead of on dimensions, has limitations - one being the inability of an exercise based AC to capture the full complexity of the target job (Thornton & Gibbons, 2009). The scope of an AC in an educational setting should be broad because a particular curriculum, such as the HRM program in this instance, prepares students for a variety of HRM jobs. The need to generalize to a wide variety of possibly unknown situations is one of the original reasons for using dimensions in AC's (Thornton & Gibbons, 2009). That said, feedback on AC dimensions should be given with care and those giving the feedback should be informed that individual competencies cluster into a smaller number of dimensions. The focus should be on giving feedback about what behavior is being observed and how that behavior relates to a competency dimension.

As expected, students do better in an AC as they progress through their education. The overall AC rating in year one shows no significant correlation with GPA in year one, however the correlation between OAR2 and GPA2 is significant. The difference in the correlations can be explained by the fact that students have many short courses during the first year of the program, whereas courses in the third year are larger, more complex, and

more integrated with other courses. Moreover, students spend a disproportionate amount of time on internships during their third year. Students are graded on twenty-two, short, self-contained courses during their first year at school. It is argued that in such short courses at best only 'elements of competencies' are assessed, and that such elements do not relate strongly to the broader competencies that are measured in an AC. Conversely, later on during their third year of study, students are graded on only eleven, large, highly integrated, and practically oriented courses (e.g., projects, internship training). In fact, the overall OAR is most strongly correlated with performance during the student internship in which competencies are necessary for success. AC and internship performance ratings are each based on behavioral observations in actual work situations, albeit in the AC in simulated form.

Not just students, but those who develop and implement the curriculum of a school can learn from the feedback provided by an AC, especially when the AC is used more than once. For example, if a given cohort receives low ratings on a specific competency (e.g., creativity), then faculty can assign new class work or team projects to help students develop that competency, reassign an existing course to a different year, or delete an existing course altogether. Once changes have been implemented, student learning can be monitored with the second AC and a continuous feedback loop can provide ongoing information about the benefits of the changes.

This study shows only small correlations between AC ratings and various measures of personality, which is in line with earlier studies that reported only marginal correlations between intelligence and personality (Waldman & Korbar, 2004), but in contrast to Collins et al. (2003) who reported moderate correlations. Assessors in the current study are not aware of the candidates' personality test scores, which is not the case in most other studies. Knowledge of such scores can artificially influence the correlation between AC ratings and personality, because assessors may be influenced by the scores when they assess the candidates after the exercises.

An important feature in the current study concerns the use of students as assessors. Critics argue that students should not be used as assessors because they have insufficient experience evaluating others. Although students may be new to being AC assessors, they do have experience rating others during their curriculum, such as when they rate their team members during project work. Senior students have passed through the curriculum and can therefore provide feedback to the candidates concerning how and when specific competencies should be learned. Senior students may be more receptive to, and have more time for, assessor training than staff members. They may provide more objective ratings of candidate behavior than staff because they are not acquainted with the prior achievements of the candidates. Student assessors are likely highly motivated because they receive academic credits for their effort and because they feel a heightened sense of responsibility while rating others. Indeed, student assessors report that they learn new skills in the assessor training and are positive about the whole assessor experience. A

disadvantage of using senior students who are ready to graduate is that new student assessors need to be trained each year. However, as is the case in the present study, neither extra time nor faculty effort are needed when assessor training is incorporated into the curriculum. Moreover, the findings of this study show that ratings of student assessors are reliable. Schleicher, Day, Mayes, and Riggio (1999) report sufficient interrater reliabilities for student assessors, and Riggio et al. (2003) denote that trained student assessors do about as good a job as trained faculty assessors. In fact, the interrater reliability coefficients in the current study are generally higher than those of studies listed in the meta-analysis by Collins et al. (2003), in which faculty members act as assessors.

Critics further argue that AC's are expensive, complex and time consuming, and are not suitable for evaluating the effectiveness of a curriculum (see Ebert, 2009). AC's are in fact not appropriate for evaluating individual courses within a curriculum, but AC's have additional value when they measure broader competencies that are learned throughout the curriculum. Cost is a issue but can be considerably reduced when senior students are used as assessors and when AC exercises are videotaped and scored online (see Mayes, 1997). A further strategy is to design a special course that focuses specifically on the AC and to embed that course within the existing school curriculum. Adding a special course significantly reduces the chances that an AC program, once started, will be terminated by management (see Waldman & Korbar, 2004). In short, institutions can make the AC cost effective while simultaneously adding to the educational experience of students and should consider using exams and AC's alongside one another.



## Chapter 8

# Competencies, grades, intelligence, personality, and motivation as predictors of early career success\*

### Abstract

Undergraduate students participated in an assessment center, once at the end of the first year at college, and a second time at the end of their third year, and were rated on nine competencies. Scores on intelligence, personality, motivation, and GPA were available for each student. Years later, after graduating from college, students were contacted and asked to provide self-ratings of career success. An important purpose of the current research was to assess the predictive validity of competencies as measured by AC exercises for early career success. Not only the direct (based on zero-order correlation) predictive validity, but also the incremental validity, of competency ratings was examined due to the inclusion of GPA and a host of individual differences variables (e.g., intelligence, personality). The main findings of the current study are that competency ratings were the best predictor of early career success, even after holding GPA constant, however the reverse was not true. In fact, competency ratings continued to have incremental predictive validity when other covariates, namely, GPA, intelligence, personality, motivation, and learning style were controlled. In contrast, the incremental contribution of GPA above and beyond the other predictors was marginal. Practical implications of these findings are discussed.

\* Manuscript submitted for publication

Research in educational psychology has shown that prior academic performance typically predicts subsequent academic performance. High school GPA predicts college grades (e.g., Ramist, 1984) and college GPA, in turn, predicts performance in graduate school (Kuncel, Hezlett, & Ones, 2001), business school (Kuncel, Credé, & Thomas, 2004), and law school (Linn & Hastings, 1984). This is as might be expected because the different levels of schooling have similar environments. However, academic learning environments and actual work environments are far less similar, and research has shown that grades are only marginally related to career success (Pfeffer & Fong, 2002; Roth, Bevier, Switzer, & Schippman, 1996; Samson, Graue, Weinstein, & Walberg, 1984).

Researchers have theorized that the relationship between performance in school and performance on the job might strengthen when other predictors such as competencies are considered (Vermeulen-Kerstens, 2006). Donofrio and Davis (1997) have suggested that general competencies such as teamwork and communication skills may be more strongly related to career success than grades, and efforts have been made to decrease the gap between school and work by redesigning the academic learning environment. Learning environments, especially in business schools and undergraduate education, are no longer exclusively based on knowledge acquisition. They now include learning skills and developing competencies. In addition to the traditional lecture format, students receive skills training (e.g., management skills, advising), work in project teams to solve business problems, and have intensive job internships.

The purpose of the current chapter is to consider the link between success in school and subsequent early career success. More specifically, this chapter considers whether competencies, and the requisite learning environments needed to build the competencies, might be a better predictor of career success than is the traditional grade based system. The effects of intelligence, personality, and motivation, which have been shown to be important predictors of both academic and career success, are also explored.

### **8.1 Literature**

#### **8.1.1 *Grades as predictors of career success***

A natural starting point of any analysis of the link between success in school and early career success is to review how well grades predict success at work. Initial meta-analytic studies that have investigated the relationship between college GPA and job performance have reported either no relationship (Bretz, 1989) or a weak relationship (Baird, 1985; Cohen, 1984; Samson et al., 1984). For example, Baird (1985) reported a weak association between college performance and various indices of occupational success, such as scientific output or success in management. A meta-analysis by Samson et al. (1984), based on 209 correlations from 35 studies, reported a mean correlation of .16 between grades and career success. Specifically for undergraduate GPA, based on 132 correlations,

they reported a mean correlation of .15, thus explaining only 2.3% of the variance in career success. Cohen (1984) reported a mean correlation of .18 between grades and overall work success based on a sample of 108 studies. A more recent meta-analysis by Roth et al. (1996), based on 71 studies, examined the relationship between college grades and various job performance measures such as output, supervisor and expert ratings, and reported an overall correlation of .16. Finally, a meta-analysis by Roth and Clarke (1998) investigated how well grades predicted income and reported that college GPA correlated .14 with initial salary, .17 with current salary, and .05 with salary growth.

One argument for the low correlations between performance at school and at work is that academic tasks are different from real world tasks (Sternberg & Wagner, 1993). A second argument is that how students are evaluated at school – often by traditional types of exams – differs from how performance in the workplace is evaluated. In traditional learning environments, exams are based on problems that are highly structured and have a single correct answer (Kuncel, Hezlett, & Ones, 2004). In work situations, actual business problems lack a single correct solution because multiple strategies, all of which may be correct, can be used. Pfeffer and Fong (2002) found that GPA and early job success were unrelated in a sample of MBA students and concluded that there was a poor match between subject matter mastery as represented by grades and the skills and competencies actually necessary to achieve early career success.

#### *8.1.2 Competencies as predictors of career success*

The marginal correlation between GPA and career success indicates that much variance in career success has yet to be explained. Alternative appraisals of student ability, such as 360 degrees feedback, portfolios, and assessment center exercises capture a broader array of skills and competencies and may be a better match with what is actually required for success in the workplace (Boyatzis, Stubbs, & Taylor, 2002). Such alternative measures should help explain variance in career success.

Portfolios are increasingly being cited as a viable alternative to standardized testing in education. A portfolio is a collection of a student's best work or best efforts and student-selected samples of work experiences, and it documents learning and development toward mastery. Portfolio assessments of medical student competencies have shown low to moderate correlations with traditional student examinations (Davis et al., 2001). The authors concluded that the portfolio is a powerful tool for assessing a range of curriculum outcomes that are not easily assessed by other methods. However they did not investigate the predictive validity of the portfolio for career success. Moreover, McMullan et al. (2003, p.291) have noted that there “is much theoretical discussion, but very little empirical research”, concerning the effectiveness of the portfolio.

Assessment centers (AC's), although not as popular in educational settings as portfolios, show promise for assessing competencies. Assessment centers include, a) competencies upon which individuals are assessed, b) exercises or measures to assess the

competencies, and c) assessors to evaluate the candidates. In an academic AC, student behavior is observed and rated by multiple assessors while the students solve actual business problems. Assessment centers allow for the measurement of an array of skills and competencies that cut across all of the courses in a curriculum.

Assessment centers have clearly predicted subsequent career success for employees in business and government organizations (Cook, 2004; Gaugler, Rosenthal, Thornton, & Bentson, 1987; Schmitt, Gooding, Noe, & Kirsch, 1984). Schmitt et al. (1984) meta-analyzed 99 studies and found a mean validity coefficient of .36 between AC ratings and career success. A meta-analysis by Gaugler et al. (1987) three years later reviewed 50 studies concerning the validity of the AC to predict job performance. They reported an overall mean AC validity of .37 for performance, and of .36 between the AC and a criterion measure of career advancement (e.g., level of salary, change in salary over time). Jansen and Stoop (2001) reported that AC ratings were related to career advancement over a 7-year period in a sample of 679 academic graduates. These graduates were assessed as part of the selection procedure at a large Dutch telecommunication company. The validity of the overall assessment rating (OAR) for employees who had been working seven years, corrected for initial differences in starting salaries and restriction of range, was .39. In addition to an OAR, the authors extracted three factors from the AC ratings and reported that the dimension 'firmness' was a stable predictor of job performance across the seven year time span. Similarly, Arthur, Day, McNelly, and Edens (2003) reported that competencies which tapped aspects of feeling and firmness were slightly more predictive of career success than was the OAR (see also Meriac, Hoffman, Woehr, & Fleisher, 2008).

Recent research involving an AC in an academic setting has also suggested that, in comparison to other measures, AC ratings may be more predictive of subsequent job performance (Gniatczyk & Ladd, 2001). In fact, Riggio et al. (2003) reported that students who performed better in an academic AC received higher supervisor ratings and were slightly more successful in their careers two years after graduation. A study by Waldman and Korbar (2004), using a sample of 66 undergraduate business students, showed that AC ratings predicted job satisfaction ( $r = .35, p < .01$ ), number of promotions ( $r = .48, p < .01$ ), and current salary ( $r = .39, p < .01$ ). They also found that the overall AC rating (OAR) had incremental validity in predicting number of promotions ( $\Delta R^2 = .07$ ) and current salary ( $\Delta R^2 = .05$ ), over and above that of grade point average (GPA) and personality traits. Compared to traditional approaches of assessing student capabilities such as multiple choice examinations and short essays, the performance-based approach of the assessment center has more potential. It comes closer to assessing the ability of a student to integrate knowledge, values, attitudes, and skills that are ultimately necessary in real world situations (Gonczi, 1994).



### 8.1.3 *Cognitive ability, personality, and motivation as predictors of career success*

A thorough analysis of the link between success in school and success at work requires an examination of additional variables that are known to be important. There are, in addition to competencies and grades, several other individual differences variables that explain variance in both performance at school as well as on the job, namely, intelligence, personality, and motivation. Each of these factors is considered in turn.

There is a substantial relationship between intelligence and job performance (Schmidt & Hunter, 1998) which generalizes across jobs (Schmidt, 2002) and cultures (Salgado & Anderson, 2001). Intelligence is considered the best predictor of job performance and training proficiency in US samples, with corrected predictive validity coefficients of .55 for job performance and .63 for training proficiency (Schmidt & Hunter, 1998). For European samples, Salgado, Anderson, Moscoso, Bertua, De Fruyt, & Rolland (2003) reported validity coefficients of .62 and .53 for job performance and training proficiency respectively. A meta-analysis by Kuncel, Hezlett, and Ones (2004) showed further that intelligence is a valid predictor of job performance and reported a (corrected) validity coefficient of .41 between intelligence and job performance. Ferris, Witt, and Hochwater (2001) reported that intelligence predicted current salary level.

Intelligence has also predicted academic achievement. The meta-analysis by Kuncel et al. (2004) reported a (corrected) validity coefficient of .39 between intelligence and graduate GPA, and of -.35 between intelligence and time needed to finish a degree. However, Chamorro-Premuzic and Furnham (2005) noted, based on studies performed in the U.S. and the U.K., that the correlation between intelligence scores and academic performance decreased as students became older; declining from .60 to .50, to .40, and to .30, respectively, at the elementary, secondary, university undergraduate, and postgraduate levels. Consistent with the declining validities as students progress through their education, Resing and Drenth (2009) reported a modest relationship between academic achievement and intelligence in a general sample of Dutch university students, and Busato, Prins, Elshout, and Hamaker (2000) reported a correlation coefficient of .13 in a sample of university psychology students.

The Big Five personality traits have been associated with various measures of job performance (Barrick, Mount, & Judge, 2001; Hough, 1992; Hurtz & Donovan, 2000; Mount & Barrick, 1995; Tett, Jackson, & Rothstein, 1991). Barrick and Mount (1991) reviewed 117 studies utilizing 162 samples with 23,994 participants and reported that conscientiousness was consistently associated with all performance criteria across all occupational groups. Salgado (1997) conducted a meta-analysis using European samples and reported that after controlling for intelligence, conscientiousness accounted for 11% of the variance in job performance and neuroticism accounted for 10%. Barrick and Mount (1991) reported that extroversion was a valid predictor for occupations involving social interaction, such as in management and sales positions. Furthermore, extroversion and openness to experience were significantly related to training proficiency criteria. In a

recent meta-analysis, Salgado (2004) reported the following population values for correlations with training success: conscientiousness (.33), neuroticism (.21), agreeableness (.19), extroversion (.10), and openness to experience (.09).

Personality measures have also been found to correlate with job satisfaction. In their meta-analysis, Ng et al. (2005) found personality traits to be related to career satisfaction with overall validity coefficients: conscientiousness (.14), neuroticism (-.36), agreeableness (.11), extroversion (.27), and openness to experience (.12).

Considering the evidence of a strong link between personality and measures of job performance and satisfaction, it is not surprising that personality measures correlate with success in school as well. Of the Big Five personality traits, conscientiousness has the largest and most consistent relationship with academic performance (O'Conner & Paunonen, 2007; Poropat, 2009). In a recent meta-analysis by Poropat (2009), a mean correlation of  $r = .24$  between conscientiousness and academic achievement was reported. Of the other four personality factors, only openness to experience ( $r = .12$ ) and agreeableness ( $r = .07$ ) were related to grade point average (GPA). Kappe and Van der Flier (2010) used five specific measures of academic success instead of the traditional GPA and reported that conscientiousness was consistently related to academic performance across all five performance measures and that neuroticism was positively related to performance when performance was measured in a less stressful learning environment.

Learning style is a particular type of personality variable. Learning styles are often used in management development programs, however evidence of predictive validity for job performance is meagre at best (Coffield, Moseley, Hall, & Ecclestone, 2004). Kappe and Van der Flier (2009) examined the usefulness of learning styles for predicting academic achievement. Although learning styles were matched to correspondingly suitable learning criteria, the learning styles revealed no predictive validity. Nonetheless, learning style typologies are still widely used by organizations and educational institutions for training and learning purposes. One reason may be the face-validity of learning styles. At first glance it seems logical to expect that learners with, for example, a theorist style will outperform those with an activist style when learning new knowledge in a traditional lecture and exam setting.

A fourth predictor of early career success is motivation. Sackett, Gruys, and Ellingson (1998) analysed data for 8,274 Army enlisted personnel and reported a significant association between need for achievement and five components of job performance, with an average correlation of .19. Intrinsic motivation has been associated with academic achievement as well (Lin, McKeachie, & Kim, 2003). Lin et al. (2003) reported that students who were intrinsically motivated achieved better grades.

#### 8.1.4 *Research questions*

An important purpose of the current research was to assess how well competencies would predict success on the job. I sought to determine whether competencies would predict success at work after controlling for grades, intelligence, personality, learning style, and motivation. Although one might expect that competencies would add unique variance when predicting career success, the incremental validity is not expected to be too large because the effects of the covariates are to some degree already reflected in one's competencies.

### 8.2 **Method**

#### 8.2.1 *Participants and Setting*

Students in a professional school of higher education, majoring in human resource management (HRM), participated in a longitudinal study. Students studied such HRM issues as how to reintegrate the long term unemployed into the workforce, help employees plan their current careers, and avoid legal repercussions when employees are terminated. The current HRM program used a variety of learning environments such as classroom tuition, team projects, and skills training. A large amount of time was devoted to internship training during an internship in an actual business setting in the second, third, and fourth years. Students were required to write a thesis in the last six months of their study. Students were generally 18 years old when they entered the HRM program.

Data were collected from 174 students in four cohorts: 2000-2004,  $n = 34$ , 2001-2005,  $n = 44$ , 2002-2006,  $n = 57$ , and 2003-2007,  $n = 39$ . Data were not used from students who dropped out without earning a diploma or who had missing data. The remaining 148 students were sent an alumni questionnaire and 101 students returned the questionnaire for a response rate of 69%. Of the respondents, 26% were male and 74% were female. The survey was sent in 2009, thus, depending on the cohort, students had graduated from two to five years ago.

A power analysis was conducted to determine the probability of finding an expected effect size. With a posited medium effect size of .30 between AC ratings and career success and an alpha level of .05, a sample size of 101 yielded a sufficient power value of .85 (Cohen, 1975).

#### 8.2.2 *Measures*

*Early Career success.* Three measures of objective and three measures of subjective career success were obtained. Objective career success was operationalized as initial salary, current salary, and salary growth. Initial salary is the salary students received when they started their jobs after graduating, current salary was measured in November 2009. Respondents were asked how much they earned per month. Self-reports of income have

been shown to correlate highly with archival company records (Judge, Cable, Boudreau, & Bretz, 1995). For those with part-time jobs, earnings were extrapolated out to the gross monthly income equivalent of full time work. The effect of market conditions on salary should have been minimal because of standardization within cohorts, and because all of the students were applying for jobs in the same field, namely, human resource management. All raw scores were transformed into Z-scores.

Two self-ratings of job performance, and one of job satisfaction, were used to measure subjective career success. The first self-rating was a single item: "How would you rate your own performance over the past two years?" Answers ranged from 1 (*extremely poor*) to 10 (*extremely good*). In the second self-rating of job performance, students rated themselves on nine competencies: analyzing, creativity, advising, teamwork, communicating, reporting, justifying, presenting, and learning goals, using five point scales from 1 (*moderate*) to 5 (*very good*). The nine scores were averaged into a single, overall competency score. Finally, job satisfaction was measured with a single question, "In general, how satisfied are you in your current job", and responses ranged from 1 (*very unsatisfied*) to 5 (*very satisfied*).

*Academic achievement.* Five specific measures of academic achievement were obtained for each student, one each for classroom lectures, skills training, team projects, internship training, and written thesis. The score on each measure was an average of multiple assessments throughout the student's four year curriculum. Overall GPA was also calculated. An extensive description of the specific and overall GPA measures is provided in the methods sections of chapters 3 and 4.

*Competencies.* Students participated in an assessment center in their first and third years at school. Nine competencies were measured in the AC exercises: advising, reporting, presenting, analyzing, creativity, teamwork, communicating, justifying, and learning goals. Exploratory factor analysis extracted three factors from the competency ratings in both AC's and these were labeled: a) Thinking (analyzing, communicating, advising), b) Feeling (creativity, teamwork, justifying), and c) Firmness (presenting, reporting, learning goals). However, a subsequent confirmatory factor analysis revealed a best fit for a model consisting of only one competency factor. Therefore an overall AC score (OAR) was also calculated. Previous factor analytic studies have reported one underlying competency factor (Sackett & Harris, 1988; Turnage & Muchinsky, 1982). In the current study we used both the overall OAR as well as the three dimensions (see also chapter 7).

*Intelligence.* The short version of the Multiculturele Capaciteiten Test voor Hoger Niveau (Multicultural Test of Intellectual Ability for higher education) was chosen to assess intelligence because it has been widely used in Dutch higher education and has been regarded as psychometrically sound by the Dutch Testing Committee (Evers, Van Vliet-Mulder, & Groot, 2000). The following three subtests were used: double word analogies ( $\alpha = .95$ ), exclusion ( $\alpha = .84$ ), and mental speed ( $\alpha = .96$ ). This shorter version of the MCT-H was used because of time constraints, to minimize test fatigue, and because it has been

shown to correlate as high as  $r = .86$  with the long version, whose overall reliability is .97.

*Personality, motivational and learning style variables.* The Dutch version of the NEO-FFI was used to assess the Big Five personality traits. Motivational variables were measured using the “Questionnaire About How One Works” by Akkerman and Van der Linden (2000). Learning styles were measured using the Learning Style Questionnaire (LSQ) of Honey and Mumford (1992). See chapters 4 and 5 for a full description of the questionnaires.

*Time-to-graduation.* On a post hoc basis, time-to-graduation was added as an exploratory predictor variable and was operationalized as the number of months needed to acquire the HRM bachelor degree.

### 8.3 Results

As expected, the objective early career success measures were interrelated. Initial salary correlated .54\*\* with current salary, and current salary in turn correlated .83\*\* with salary growth, however job satisfaction was unrelated to any of the three salary measures.

Table 8.1 *Intercorrelations of criterion variables*

	1	2	3	4	5
1 Initial salary	-				
2 Current salary	.54**	-			
3 Salary growth	.05	.83**	-		
4 Job performance	.08	.19†	.23*	-	
5 Competencies	.13	.31**	.37**	.31**	-
6 Job satisfaction	.06	.19†	.17†	.06	.16

\*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$

Bivariate correlations between each predictor and the six criterion measures are presented in Table 8.2. For each (set of) predictor(s), the percentage of variance is also listed. As can be seen in this table the overall competency rating correlated with five of the six criterion measures, the exception being job satisfaction. The highest correlation occurred between the OAR and current salary. However, the three competency factor scores together predicted current salary even better, explaining nearly a quarter of the variance. The separate competency factors of Feeling and Firmness correlated significantly and consistently with the three salary measures.

Also listed in Table 8.2 is how each of the five specific GPA's were related to the six measures of job performance. Lectures correlated much stronger with objective job performance measures. The specific GPA's together explained 12% of the variance in initial salary, 12% in current salary, and 15% in salary growth.

Intelligence explained 8% percent of the variance in current salary and 10% in salary growth. Of the Big Five personality traits, only extroversion and conscientiousness correlated significantly with the career success measures. Of the motivation and learning style variables, intrinsic motivation, anxiety, study motivation, activist and pragmatic learning styles, each correlated significantly with two or more career success measures. In fact, the pragmatist style showed a moderate association with four of the career success measures.

Table 8.2 *Correlations and percentages of variance explained by the predictor variables for six measures of early career success*

	Objective measures			Subjective measures		
	Initial salary	Current salary	Salary growth	Job performance	Competencies	Job satisfaction
<i>Competencies</i>						
Thinking	.09	.13	.17	.28**	.17	-.17
Feeling	.23*	.34**	.31**	.26**	.18	-.16
Firmness	.22*	.44**	.35**	.15	.15	-.02
<i>F</i> (3,97)	2.51	9.92	5.62	3.23	1.17	1.43
	$R^2=.07^\dagger$	$R^2=.24^{**}$	$R^2=.15^{**}$	$R^2=.09^*$	$R^2=.04$	$R^2=.04$
OAR	.21*	.35**	.32**	.29**	.20 <sup>†</sup>	-.14
<i>F</i> (1,99)	4.44	13.87	11.61	8.77	3.53	2.08
	$R^2=.04^*$	$R^2=.12^{**}$	$R^2=.11^{**}$	$R^2=.08^{**}$	$R^2=.04^\dagger$	$R^2=.02$
<i>Academic ach.</i>						
Lectures	.25**	.28**	.19 <sup>†</sup>	.14	.04	-.09
Skills	.16	.31**	.27**	.13	.09	-.12
Projects	-.02	.15	.22*	.18 <sup>†</sup>	.06	-.04
Internship	-.07	.17	.25*	.18 <sup>†</sup>	.07	.01
Thesis	-.01	.22*	.31**	.14	.08	-.11
<i>F</i> (5,95)	2.56	2.68	3.45	1.08	.26	.68
	$R^2=.12^*$	$R^2=.12^*$	$R^2=.15^{**}$	$R^2=.05$	$R^2=.02$	$R^2=.04$
GPA	.04	.27**	.32**	.23*	.10	-.07
<i>F</i> (1,99)	.17	8.03	11.39	5.28	.88	.47
	$R^2=.00$	$R^2=.08^{**}$	$R^2=.10^{**}$	$R^2=.05^*$	$R^2=.01$	$R^2=.01$
Intelligence	.05	.28**	.32**	-.06	.03	-.11
<i>F</i> (1,99)	.26	8.07	11.31	.26	.10	1.23
	$R^2=.00$	$R^2=.08^{**}$	$R^2=.10^{**}$	$R^2=.00$	$R^2=.00$	$R^2=.01$
<i>Personality</i>						
Neuroticism	-.05	.11	.11	-.01	-.01	-.04
Extroversion	.25*	.13	.05	-.02	.09	.15
Openness	.04	.02	.01	.12	.02	-.13
Agreeableness	-.04	-.06	-.01	.01	.03	.12
Conscientiousness	.07	.26**	.25*	.16	.10	-.08
<i>F</i> (5,90)	1.47	3.40	2.16	.71	.31	1.37
	$R^2=.08$	$R^2=.16^{**}$	$R^2=.11^\dagger$	$R^2=.04$	$R^2=.04$	$R^2=.07$
Intrinsic mot.	.07	.25*	.22*	.16 <sup>†</sup>	-.07	-.05
Anxiety	-.19 <sup>†</sup>	-.20*	-.10	-.09	-.39**	-.14
Need for pressure	.00	-.01	-.01	-.10	-.05	.00
Need for status	.10	.18 <sup>†</sup>	.10	-.01	-.08	-.16 <sup>†</sup>
Study motivation	.11	.29**	.26**	.17*	-.12	-.09
<i>F</i> (5,95)	1.32	3.43	1.66	.91	3.17	.81
	$R^2=.07$	$R^2=.16^{**}$	$R^2=.08$	$R^2=.05$	$R^2=.17^{**}$	$R^2=.04$
<i>Learning styles</i>						
Activist	.17*	.16 <sup>†</sup>	.11	.03	.19*	-.14 <sup>†</sup>
Reflector	-.09	-.07	-.07	.07	-.10	.02
Theorist	-.04	.04	.06	.10	.07	-.03
Pragmatist	.19*	.30**	.25**	.12	.31**	-.10
<i>F</i> (4,93)	1.92	3.30	1.78	.70	2.78	.98
	$R^2=.08$	$R^2=.12^*$	$R^2=.07$	$R^2=.03$	$R^2=.12^*$	$R^2=.04$

\*\*  $p < .01$ , \*  $p < .05$ , <sup>†</sup>  $p < .10$

*Incremental validity*

Hierarchical multiple regressions were conducted on each criterion to determine the incremental validities of certain predictors over and above other predictors. The variables in steps one through three were the same for all criteria and were entered in the same order, namely, GPA, OAR, and intelligence, whereas the variables in step four varied as a function of whether those variables and the criterion in question showed significant bivariate correlations with one another.

GPA and OAR were used because these overall measures explained about as much variance as the adjusted portion of variance explained by the five specific GPA's and by the three AC factors. For example the adjusted percentages of variance of the specific GPA's, that is the percentage of variance corrected for chance because of the inclusion of multiple predictors, were as follows: initial salary (7%), current salary (8%), and salary growth (11%). These adjusted percentages are largely in line with the percentages of variance explained by GPA. This meant that the number of variables in the equations could be minimized. It should however be noted that the procedure of entering only significant predictors in step four capitalizes on chance. In other words, the variables in step four have more chance of being incrementally significant. The results are presented individually for each of the six criteria in Tables 8.3 to 8.7.

Table 8.3 *Initial salary regressed on predictors*

Predictor	Initial salary			
	$\beta$	$\beta$	$\beta$	$\beta$
Step 1: GPA	.06	-.01	-.02	.00
Step 2: OAR		.17 <sup>+</sup>	.17 <sup>+</sup>	.14
Step 3: Intelligence			.08	.09
<i>Step 4: other variables</i>				
Extroversion				.27*
Activist				.02
Pragmatist				.11
Multiple R	.06	.17	.18	.35
F	.32	1.32 <sup>+</sup>	1.00 <sup>+</sup>	2.11 <sup>+</sup>
( $df_1, df_2$ )	(1,93)	(2,92)	(3,91)	(4,90)
$R^2$	.00	.03	.03	.13
$R^2_{\text{change}}$		.03	.00	.09
$F_{\text{change}}$		.13	.55	.03

\*\*  $p < .01$ , \*  $p < .05$ , <sup>+</sup>  $p < .10$

Although the zero-order correlation between the OAR and initial salary was significant ( $r = .21$ ), the change in  $R^2$  in initial salary was only 3% after controlling for GPA (see Table 8.3). The incremental  $R^2$  change for intelligence was .00%, however, even after controlling for GPA, OAR, and intelligence, extroversion explained 9% of the variance in initial salary. Apparently organizations initially value undergraduate employees with a social and



energetic nature. All variables together explained 13% of the variance in initial salary when  $\alpha = .10$ .

Table 8.4 *Current salary regressed on predictors*

Predictor	$\beta$	Current salary		
		$\beta$	$\beta$	$\beta$
Step 1: GPA	.29**	.19†	.15	.09
Step 2: OAR		.25*	.23*	.16
Step 3: Intelligence			.24*	.24*
<i>Step 4: other variables</i>				
Conscientiousness				.09
Intrinsic motivation				.08
Anxiety				-.12
Study motivation				.01
Pragmatist				.23*
Multiple R	.29	.37	.44	.54
F	8.70	7.39	7.26	4.44
( $df_1, df_2$ )	(1,93)**	(2,92)**	(3,91)**	(8,86)**
$R^2$	.08	.14	.19	.29
$R^2_{\text{change}}$		.05*	.06*	.10*
$F_{\text{change}}$		.02	.02	.04

\*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$

GPA explained 8% of the variance in current salary and the OAR had an incremental contribution of 5% over and above GPA. A considerable additional amount of variance was explained when intelligence, conscientiousness, and the motivational variables were entered in the regression analysis. Intelligence explained an additional 6% over and above GPA and the OAR. When the OAR and intelligence were entered into the regression equation the contribution of GPA was no longer significant. The personality variables, especially the pragmatist style, explained another 9% of the variance in current salary over and above GPA, OAR, and intelligence. Beyond the initial entry and adjustment phase of a new job, intelligence and pragmatism appear to be key factors.

Table 8.5 *Salary growth regressed on predictors*

Predictor	$\beta$	Salary growth		
		$\beta$	$\beta$	$\beta$
Step 1: GPA	.33*	.24*	.19†	.15
Step 2: OAR		.22*	.20*	.16
Step 3: Intelligence			.26**	.27**
Step 4: other variables				
Conscientiousness				.12
Intrinsic motivation				.03
Study motivation				-.03
Pragmatist				.21*
Multiple R	.33	.39	.46	.53
F	11.25	8.04	8.25	4.76
( $df_1, df_2$ )	(1,93)**	(2,93)**	(3,91)**	(7,87)**
$R^2$	.11	.15	.22	.28
$R^2_{\text{change}}$		.04*	.07**	.06
$F_{\text{change}}$		.03	.01	.11

\*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$

GPA explained 11% of the variance in salary growth and the OAR had an incremental contribution of 4% over and above GPA. Intelligence explained an additional 7% of the variance over and above the combined effects of GPA and the OAR. When the OAR and intelligence were entered into the regression equation, the contribution of GPA was no longer significant at the .05 significance level. In total, 28% of the variance in salary growth could be explained, with significant contributions stemming from intelligence and a pragmatic learning style. Apparently for HR professionals to advance in salary, intelligence and a pragmatic style are prominent factors.

Table 8.6 *Self-rating of job performance regressed on predictors*

Predictor	$\beta$	Self-rated job performance		
		$\beta$	$\beta$	$\beta$
Step 1: GPA	.22*	.11	.14	.12
Step 2: OAR		.25*	.26*	.24*
Step 3: Intelligence			-.11	-.11
Step 4: other variables				
Study motivation				.07
Multiple R	.22	.32	.33	.34
F	4.84	5.28	3.92	3.01
( $df_1, df_2$ )	(1,97)*	(2,96)**	(3,95)*	(4,94)*
$R^2$	.05	.10	.11	.11
$R^2_{\text{change}}$		.05*	.01	.00
$F_{\text{change}}$		.02	.28	.55

\*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$

The statistics in Table 8.6 show that GPA explained 5% of the variance in the single-item, self-rating of job performance and that the OAR had an incremental contribution of 5% over and above GPA. When the OAR was entered into the regression, the contribution of GPA was no longer significant. In a model with all variables included, only the OAR had a significant contribution. Young HR professionals appear to attribute their self-rated success mainly to their previous academic competencies. Table 8.7 shows that all variables together explained 18% of the variance in self-rating of the nine competencies, with significant contributions coming from anxiety and the pragmatist style. High anxiety apparently leads to a lower estimate of one's self perceived level of competence.

Table 8.7 *Self-rating of nine competencies regressed on predictors*

Predictor	$\beta$	Self-rated competencies		
		$\beta$	$\beta$	$\beta$
Step 1: GPA	.07	.02	.02	.02
Step 2: OAR		.18	.18	.07
Step 3: Intelligence			-.02	.01
Step 4: other variables				
Anxiety				-.32**
Activist				.03
Pragmatist				.23*
Multiple R	.07	.18	.19	.46
F	.37	1.45	1.00	3.47**
( $df_1, df_2$ )	(1,81)	(2,80)	(3,79)	(6,76)
$R^2$	.00	.04	.04	.22
$R^2_{\text{change}}$		.04	.00	.18**
$F_{\text{change}}$		.03	.86	.00

\*\*  $p < .01$ , \*  $p < .05$ , †  $p < .10$

A further set of regressions tested whether competencies predicted career success when intelligence, personality, motivation and learning style were controlled for, compared to how well grades did so. In sum, the results showed that competencies outperformed GPA as a predictor of career success. For current salary, the OAR had a marginal incremental contribution, over and above intelligence and five personality variables,  $F(7,87) = 4.99$ ,  $\Delta R^2 = .03$ ,  $p = .06$ , and GPA had no incremental contribution,  $F(7,87) = 4.65$ ,  $\Delta R^2 = .02$ ,  $p = .19$ . For salary growth, the OAR had an incremental contribution, over and above intelligence and four personality variables,  $F(6,88) = 5.21$ ,  $\Delta R^2 = .04$ ,  $p = .04$ , and GPA had a marginal incremental contribution,  $F(6,88) = 5.06$ ,  $\Delta R^2 = .02$ ,  $p = .06$ . For the self-rating of job performance, the OAR had an incremental contribution, over and above IQ and study motivation,  $F(3,95) = 3.64$ ,  $\Delta R^2 = .07$ ,  $p = .01$ , and GPA had no significant incremental contribution,  $F(3,95) = 2.25$ ,  $\Delta R^2 = .03$ ,  $p = .08$ . For initial salary and the self-rating of the nine competencies, neither the OAR nor GPA had an incremental contribution over intelligence and personality.

Time-to-graduation correlated inversely and significantly with current salary ( $r = -.20$ ,  $p < .05$ ) and inversely and marginally with salary growth ( $r = -.17$ ,  $p < .10$ ). However, because students within a cohort typically differed in the number of years that they have been employed, the measures of current salary and salary growth were standardized by taking the years of employment since graduation into account. When the analyses were repeated, the correlation between time-to-graduation and current salary ( $r = -.14$ ,  $p = .19$ ) was no longer significant, and the correlation between time-to-graduation and salary growth was marginally significant ( $r = -.16$ ,  $p < .10$ ). The results seem to indicate that students who need more time to graduate progress less well in the early stages of their professional career.

#### **8.4 Discussion**

A novel aspect of the current research is the use of competencies, in addition to GPA and a host of individual differences variables, to predict early career success. The main findings of the current study are, a) AC competency ratings are related to early career success and show incremental validity over and above GPA, b) AC competency ratings have direct predictive validity and incremental predictive validity when intelligence, personality, motivation, and learning style are controlled, and c) in contrast to the effect of competencies, the incremental contribution of GPA above and beyond the other predictors is marginal.

Higher education in The Netherlands is experimenting with competence-based learning and with assessment centers to measure and provide feedback for competencies. Students develop competencies when they take part in various learning modules. For example, team projects about labor relations and reintegration into the work force emphasize not only learning about the labor market and equal employment opportunity legislation, but also help students acquire interpersonal and communications skills, and support problem-solving and decision-making skills. Internships provide the opportunity to practice competencies in an actual place of work. However, no single learning environment, whether it be internship, lecture, or team project, can assess the entire domain of competencies in a standardized manner, as is possible with an assessment center.

The AC competency ratings show strong correlations with current salary, salary growth and self-rated performance. This result confirms the hypotheses of Donofrio and Davis (1997) that competencies correlate better with job performance than does GPA. This finding also coincides with a study by Waldman and Korbar (2004) who report that AC ratings correlate .39 with current salary and account for 5% of current salary variance over and above GPA and personality traits. Although the competencies in this research correlate .41 with GPA, they still explain additional variance beyond GPA in current salary (5%),

salary growth (5%), and the self-rating of job performance (5%). These percentages appear small at first glance, however it should be noted that GPA in the current research consists of grades from learning modules that themselves are quite competency oriented (e.g., the internship, team projects).

Additional hierarchical regression analyses show that, when intelligence and personality are treated as covariates, the overall assessment center rating continues to be significantly and meaningfully related to career success, whereas GPA in contrast is only marginally related. In fact, of all the variables, only the OAR is related to the self-rating of job performance, which is in line with findings from earlier studies (Gaugler et al., 1987). Waldman and Korbar (2004) report that GPA, personality, and OAR are unrelated to job satisfaction and such nonfindings are confirmed in the current research.

GPA correlates .27 with current salary, .32 with salary growth, and .22 with the self-rating of job performance. These three coefficients are in the top of the range compared to results of previous meta-analyses. Samson et al. (1985) report validity coefficients in the range of .15 to .20, and Roth et al. (1996) report an uncorrected validity coefficient of .16. Roth and Clarke (1998) report correlation coefficients of .14, .17, and .05 between GPA and initial salary, current salary, and salary growth, respectively. There are two potential explanations for why the correlations in the current study are higher than those found in the literature. First, the GPA values in this study come in large measure from nontraditional, practice-oriented learning environments such as internships, whereas most GPA values in the literature derive from grades that are based on traditional classroom test scores (e.g., O'Connor & Paunonen, 2007). The nontraditional learning environments in the current research simulate more closely those of a normal work environment.

Second, tertiary education in The Netherlands is divided into the traditional university and the university for applied science, the latter also being known as Hoger Beroeps Onderwijs (HBO - higher professional education) and called college throughout this dissertation. Whereas traditional universities provide a conceptual learning environment and a broad based liberal arts education, colleges are more vocational in nature and educate students for a particular profession. Thus, the knowledge-based learning in the latter will more closely resemble knowledge needed to do well at work.

In line with meta-analytic findings, intelligence predicts career success. Moreover, intelligence explains unique variance over and above that of GPA and of OAR in current salary (6%) and salary growth (6%). These results show that intelligence has a unique role over other predictors for two important measures of career success. In line with previous studies, intelligence is an important predictor of career advancement (Schmidt & Hunter, 1998).

Two unexpected associations between personality and early career success are noteworthy - the correlations between extroversion and initial salary, and between a pragmatic learning style and current salary. One may conclude that employers initially value an extrovert HRM undergraduate over an introvert HRM undergraduate. However

one may also reason that extroverts managed to negotiate a higher salary and/or finds jobs with higher starting salaries.

To advance in an HRM career, a pragmatic (learning) style is important. That a pragmatic (learning) style is valued in Dutch HR professionals can be explained by the nature of the early career tasks such as giving realistic and legally correct advice to managers and employees concerning hiring decisions, contract termination, and business decisions, to name but a few.

One feature of the research design in the current study is particularly relevant to an ongoing discussion in the assessment center literature, namely, that of criterion contamination. Two types of criterion contamination, direct and subtle, have been distinguished (Jansen & Stoop, 2001). Direct contamination occurs in validity studies when criterion and predictor ratings are not independent of one another. This occurs in practice when a manager, who appraises an employee and makes promotion decisions, is aware of how the employee was rated during the assessment center exercises. Subtle criterion contamination occurs when managers and assessors share the same stereotype of a 'good manager' in a particular organization and when promotions and assessor ratings are jointly influenced by that stereotype. Klimoski and Brickner (1987) and Jansen and Vinkenburg (2006) have suggested that predictor/criterion validity correlations from assessment centers might be artificially inflated because of contamination. Because the assessors who rated the HRM students could not have known what type of jobs and organizations the students would move into after graduation, and because it is not likely that the current supervisors of those same students are aware of the original AC ratings, the correlations between predictor and criterion in this study can be considered to be free of contamination. An exception is the correlation between the AC competency ratings and the two self-ratings of job performance.

Another issue in AC research concerns the incremental validity of cognitive ability and personality over AC outcomes. Research of the incremental validity is complicated because in a lot of studies it is unclear if in the OAR test results of cognitive ability or personality are incorporated. Thornton and Gibbons (2009) detected that of the all studies used in the meta-analysis by Gaugler et al. (1987) in only four studies it was reported how test scores were used in the AC. In the present study the scores of intelligence and personality were not incorporated in the OAR and raters, but also participants, were unaware of these test scores. Furthermore, the results indicate that the predictive validity of the OAR is not due to inclusion of scores of cognitive ability or personality.

The current study has several limitations. One limitation concerns the sample – all students are HRM professionals. There might be less variation in career and salary advancement of HRM professionals compared to professionals in other economic sectors. Any type of restriction of range in the variables will attenuate correlations; thus the findings in this study may actually be on the conservative side. That 75% of the sample are female might potentially affect the results because the starting salary, as well as salary

advancement, is generally lower for females. However, no gender effects occur in the data. A third limitation concerns the use of student assessors. Although the students receive 20 hours of training and although their assessment ratings are as reliable as those of faculty (Riggio, Mayes, & Schleicher, 2003) and business assessors, findings may differ when outside faculty, psychologists, or business professionals are the assessors. A final limitation concerns the fact that it is not possible to correct for attenuation due to unreliability in the objective criterion measurements. For example, the participants are employed at a number of different companies who more than likely use different salary and promotion systems.

### *Practical implications*

The results of this study suggest that when students learn competencies, as measured by the exercises in an assessment center, they progress faster in their careers. This is useful for students who seek responsibility and promotions, and for employers who seek talent. Institutes of higher education should consider integrating competencies into their curricula and provide students with feedback concerning how well they are doing on each of the competencies. It may also be possible to train teachers to use behavioral anchors, similar to the BARs used in this study (see also chapter 7), to provide students with behavioral feedback in class. Findings show that extroversion is an important trait for a job interview, therefore shy students might be trained to exhibit more extrovert behavior during the interview. Students could similarly be trained to use the competencies as they prepare for job interviews. By describing one's competencies in detail, students are in a good position to answer that most dreaded of interview questions, "So tell me, why do you think that you would be good for this job".

The results also carry implications for organizations. Competencies correlate with salary growth, which is important because employees who receive the largest salary increases generally make the biggest contributions to organizational goals and are therefore the most valuable employees. Organizations should use assessment centers to select prospective employees. Doing so will add a significant amount of valuable information, especially when only GPA information is available about the candidate, and will increase the chance of hiring the most successful candidate. If an assessment center is considered too costly, then organizations should rely more on cognitive ability tests rather than on GPA. Finally, a battery of personality tests should be considered because such a battery is cheap and easy to administer and because it too will add valuable information for selection and development purposes.





## Chapter 9

### Summary of findings and recommendations

#### 9.1 The aim of the study

Knowledge-based economies require highly educated citizens, and insight into the factors that affect success in higher education is more important than ever. Enrolment numbers in higher education are increasing annually, yet graduation rates are decreasing (CBS, 2010). Although the Dutch government has set new targets to increase graduation rates, they are not likely to be met.

The findings from this dissertation provide insight into the relationships between intelligence, personality, motivation, and competencies on the one hand, and academic achievement and early career success on the other. The findings can help promote policies that increase graduation rates, decrease time-to-graduation, and help students make more informed decisions. Knowing how student characteristics and academic achievement are related to early career success is important for individuals and business organizations alike. Individuals can work on their developmental needs in the safety of the college environment and businesses can make better selection decisions. The findings are useful to several stakeholders: science, business organizations, educational institutions, and students.

Three questions drove the research in this dissertation:

1. *How well do student characteristics predict grades?*
2. *How well do student characteristics predict time-to-graduation?*
3. *How well do student characteristics and academic achievement predict early career success?*

An overview of all variables in the current study is presented in Figure 9.1, the research model.

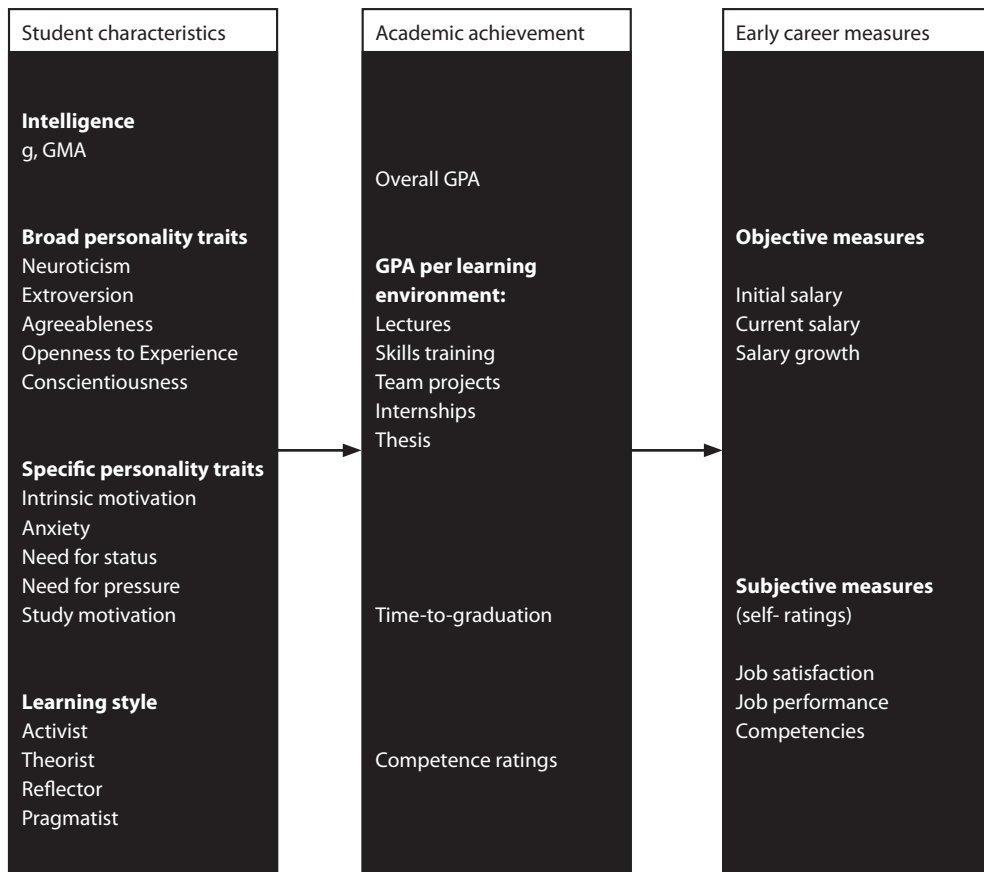


Figure 9.1 Research model

## 9.2 Summary of the main findings

The set of student characteristics in the current study are intelligence, the Big Five personality traits, motivation, four specific personality traits, and learning styles. Academic achievement and early career success are the two major sets of dependent variables. Overall GPA, grades in five specific learning environments, and time-to-graduation measure academic success, whereas salary and self-reports measure early career success. In the main, the different studies in this dissertation focus on how well various predictors account for variance in the dependent measures.

### 1. *How well do student characteristics predict grades?*

#### *Intelligence*

The findings suggest a modest and stable positive correlation between intelligence and grades throughout college (chapters 3 and 4). Intelligence is related to first year GPA as well as to overall GPA (chapter 6). This result is in line with other research findings concerning the predictive validity of intelligence for GPA in university samples, both within The Netherlands (Busato et al., 2000; Resing & Drenth, 2009; Wolting, 2006) and outside it (Farsides & Woodfield, 2003; Furnham & Chamorro-Premuzic, 2004). It is often argued that the lack of a strong correlation is surely an artefact due to restriction of range in the intelligence variable (Drenth, 2004).

#### *Big Five Personality Traits*

The main finding is that the personality trait of conscientiousness correlates more highly with overall GPA than does intelligence (chapter 3). Not only does conscientiousness, after controlling for intelligence, account for 22% of the variance in GPA, it is also consistently related to each of the five specific academic achievement criteria. In fact, conscientiousness explains five times as much variance in academic achievement as does intelligence (chapter 4). An additional set of studies within other academic disciplines confirms the importance of conscientiousness for first year grades. There were comparable validity coefficients in a second HRM sample as well as in a Sports, Health & Movement sample (chapter 6). Many studies report that conscientiousness is related to academic achievement (Bratko et al., 2006; Gilles & Bailleux, 2001; Nofle & Robins, 2007; Poropat, 2009), or that conscientiousness has even more predictive power than intelligence (Conard, 2006; Di Fabio & Busoni, 2007; Furnham & Chamorro-Premuzic, 2004; Furnham, et al., 2003; Petrides et al., 2005). Conscientious individuals perform better because they persevere longer and are more organized than their counterparts.

Relationships with grades for the other personality traits occur only sporadically: neurotic students perform better in the skills training learning environment, extroverts

perform less well in skills training, and open minded students perform less well in team projects (chapter 3).

### *Intrinsic motivation*

The findings show that intrinsic motivation is moderately correlated with overall GPA as well as with first year GPA (chapters 4 and 6). Intrinsic motivation explains more than twice as much variance in GPA as does intelligence, and intrinsic motivation has incremental validity over and above intelligence. The magnitude of this effect coincides with meta-analytic findings (Robbins et al., 2004; Vallerand et al., 1993). However, intrinsic motivation has no incremental validity above and beyond conscientiousness. As noted in chapter 4, motivation and conscientiousness are strongly related and an item content analysis shows that both tap 'sustained effort'. Therefore, the lack of incremental validity, once conscientiousness is held constant, can create the misperception that intrinsic motivation does not predict academic achievement. Of course, nothing is further from the truth.

### *Specific Personality Traits*

The effects of anxiety, need for pressure, need for status, and study motivation were explored. The findings show that anxiety is unrelated to GPA. Environmental pressure correlates negatively, and study motivation correlates positively, with GPA (chapter 4). Study motivation is of importance across the curriculum - that is, study motivation correlates with achievement in all learning environments. Environmental pressure showed a negative correlation with GPA and achievement across the learning environments – of note significantly with thesis. Students who need additional external pressure achieve lower grades in general and especially for their thesis.

### *Learning Style*

The findings reported in chapter 5 show that the learning styles of the Learning Style Questionnaire by Honey and Mumford (1992), namely Activists, Theorists, Pragmatists, and Reflectors are neither related to overall GPA, nor to grades from any of the five specific learning environments. This casts doubt on the importance of learning styles for academic achievement, contrary to the believe of many teachers and scholars.

### *Summary*

Student characteristics that are related to academic achievement are, in order of importance; conscientiousness, motivation and intelligence (chapter 3-6). When combined, the student characteristics explain 33% of the variance in GPA, with conscientiousness and intelligence contributing the most variance. This means that intelligent, eager, disciplined, and hardworking students obtain the best grades.

## *2. How well do student characteristics predict time-to-graduation?*

### *Intelligence*

The results presented in chapter 4 report a meagre correlation between intelligence and timely graduation. Intelligent students seem to graduate only slightly faster than less intelligent students.

### *Big Five Personality Traits*

By far, the most important predictor of time-to-graduation is, again, conscientiousness, which explains about 17% of the variance. This is in line with results from meta-analytic studies by Nofhle and Robins (2007), O'Conner and Paunonen (2007), and Poropat (2009) and also with results from studies in the Netherlands by Van Bragt (2010) and Busato (2000) who reported that conscientiousness is a key determinant for study progress in the first year of college (Van Bragt, 2010) and the first year of university (Busato, 2000). Neuroticism and openness to experience are somewhat related to time-to-graduation (chapters 3 and 4). This means that disciplined and hard working students need less time to graduate and open, curious as well as emotionally less stable students need more time to graduate.

### *Intrinsic motivation*

Intrinsic motivation explains twice as much variance, as does intelligence, in time-to-graduation, and has incremental validity over and above intelligence (chapter 4). However it should be noted that intrinsic motivation and conscientiousness are highly intercorrelated.

### *Specific Personality Traits*

None of the specific personality traits, namely, anxiety, need for pressure, need for status, and study motivation, is related to time-to-graduation.

### *Learning Style*

The findings reported in chapter 5 show no significant relationships between the four learning styles of the LSQ and time-to-graduation.

### *Summary*

Student characteristics that are related to time-to-graduation are, in order of importance; conscientiousness, openness to experience, intrinsic motivation, neuroticism and intelligence (chapter 3-6). This implies that eager, disciplined, motivated and hardworking students need less time to obtain their diploma and open, curious as well as emotional less stable students generally need more time. When combined, the student characteristics explain 30% of the variance in time-to-graduation, with conscientiousness

and openness to experience being the most important predictors. This means that disciplined, hard working students who are not easily distracted and are able to focus on the study goals are those who graduate on the timeliest manner.

### *3. How well do student characteristics and academic achievement predict early career success?*

Factors related to success in early careers are examined in chapter 8. Objective measures of career success include initial salary, current salary, salary growth; and subjective measures include job satisfaction, self-rated performance, and self-rated competency. Intelligence, broad and specific personality traits, intrinsic motivation, and learning styles are used as predictor variables, as are the academic achievement measures of overall GPA, specific learning achievement criteria, competencies, and time-to-graduation.

#### *Student Characteristics*

Intelligence explains 8% of the variance in current salary and 10% in salary growth, but no variance in initial salary, self-rated job performance and competencies, or job satisfaction. This is in line with research findings concerning the predictive validity of intelligence tests for work settings (e.g., Hunter & Hunter, 1984; Kunzel et al., 2004; Salgado et al., 2003; Schmidt & Hunter, 1998). Intelligence also shows incremental validity, over and above GPA and an overall assessment rating (OAR), for current salary and salary growth. Of the Big Five personality traits, only extroversion and conscientiousness correlate significantly with career success – and then only with the objective measures. Conscientiousness predicts salary growth on its own, but not after controlling for GPA and the OAR. In other words, there is less chance of an incremental contribution over and above GPA because GPA and conscientiousness are so strongly intercorrelated.

Of the other predictor variables, intrinsic motivation, anxiety, study motivation, activist and pragmatic learning styles, each correlate significantly with two objective career success measures. The pragmatic learning style predicts three career success variables (chapter 8) after controlling for GPA, the OAR, and intelligence. That the pragmatic learning style had an incremental validity was unexpected because this style is not related to academic achievement in the HRM program (chapter 5) and because a review by Coffield et al. (2004) indicates little evidence that learning styles have predictive validity at all. Nevertheless, pragmatism predicts current salary and a self-rating of competency, over and above the other predictor variables.

### *Academic Achievement*

*Overall GPA.* Overall GPA explains 8% of the variance in current salary and 12% in salary growth. The percentages are of the same magnitude as those explained by intelligence.

*Specific Learning Environments.* The findings show that the GPA's based on achievement in the specific learning environments together explain 12% of the variance in initial salary, 12% of the variance in current salary, and 15% of the variance in salary growth. Initial salary is best predicted by lecture grades, current salary is best predicted by lecture and skills grades, and salary growth is best predicted by grades in all five learning environments.

*Time-to-graduation.* Time-to-graduation correlates inversely and significantly with current salary and inversely and marginally with salary growth.

*Competencies.* The overall competency rating (OAR), which is an average of nine competencies ratings based on two large scale assessment centers (chapter 7), is related to the three objective measures and explains significant portions of the variance in initial salary (4%), current salary (12%), and salary growth (11%). Further analyses show that the OAR has an incremental contribution over and above GPA on four of the six career success measures.

Besides an OAR, three competency dimensions were distinguished, namely thinking, feeling and firmness. The competency dimensions of feeling and firmness are significantly related to all three objective measures of career success. The competency ratings show stronger correlations with initial salary, current salary, salary growth, and self-ratings of performance, than does GPA. Although the competency ratings correlate moderately with GPA, they still explain unique variance in career success beyond GPA, a result in line with a study by Waldman and Korbar (2004). In general, competencies outperform GPA as a predictor of career success when the student characteristic variables are partialled. This result supports the hypothesis by Donofrio and Davis (1997) that competencies correlate better with job performance than does GPA. It appears that competencies capture additional important job related aspects, notably behavior. That AC competency ratings are more effective than GPA as predictors of early career success can be explained by the goal of the AC's, cutting across the curriculum, whereas tests and exams are designed to measure a specific skill or knowledge base.

### *Summary*

The set of predictor variables explain up to 29% of the variance in certain early career success variables such as current salary. The predictor variables explain more variance in objective than in subjective measures of career success. Competency ratings show incremental validity for early career success measures, after controlling for the other

predictors. This finding provides clear support to the value of competence-based educational methods.

### 9.3 Recommendations

One of the goals, as enumerated in chapter 1 of this dissertation, is to provide practical recommendations to the various stakeholders who might benefit from educational research: business organizations, educational institutions, and students. This goal answers the proverbial “so what” question and bridges the gap between knowledge (*kennis*) and application (*kunnen*).

Companies should, where possible, factor student competencies, intelligence, and conscientiousness into their hiring process. Research shows that interviewers give higher ratings to job applicants who exert extrovert behavior (Barrick, Dustin, Giluk, Stewart, Shaffer, & Swider, in press). That extroverts do better in interviews is not unusual because the interview is a social situation and extroverts generally do better in social situations. The results from chapter 8 show that extroversion is correlated with initial salary. However, salary growth is a better measure of employee capability than is initial salary and growth is predicted not by extroversion, but rather by competencies, intelligence, and conscientiousness. Because extroversion is unrelated to conscientiousness,  $r = .05$  (chapter 3) and unrelated to competencies,  $r = .03$  (chapter 7), companies are losing many good prospective employees by focusing too much on extroversion and not enough on competencies and conscientiousness.

Institutes of higher education that use admission tests may consider including tests of personality and/or motivation. Some caution is, based on the percentages of explained variance, in place. Not only is much variance in academic success unaccounted for, one should also be aware that the students in this study filled out the questionnaires on a voluntary basis, which is very different from using tests for selection. Institutions that have an open admission policy, and are therefore unwilling or unable to be selective, may consider administering a personality or motivation test during a student's first coaching session. Students who score low on conscientiousness are at risk of not graduating on time and such students could be 'signaled' so that proper and timely interventions can be developed. In the main, there are two types of strategies that can be used to help students.

The first strategy develops conscientiousness related skills, such as planning, time-management, and being more organized in general. There are three possible interventions with this strategy. A course on time-management skills during the first semester might increase the students' first year grade point average as well as their study pace. A second intervention would be to teach teachers how to act upon the strengths and weaknesses of



students in their classes. Teachers could make use of formative assessment rather than just a summative assessment at the end of their academic course in order to stimulate active learning and avoid procrastination. The use of short assignments can also be helpful to actively engage students in learning during a course. Formulating clear learning objectives and translating them into sub-goals may also help students to structure their learning. Setting clear, well defined and also challenging learning objectives may also trigger students' motivation. Another intervention is to create study associations in which students share notes, discuss difficult topics, collect and disseminate lecture notes and book summaries, organize presentations and workshops by (foreign) guest speakers, but also study jointly on difficult courses. These study groups could be effective for students who have motivational problems and therefore need additional pressure to start studying or for students who have low self-discipline and who therefore have difficulty studying alone.

A second strategy, currently being widely discussed in the Netherlands, is for institutions to make changes in the learning environment or (elements of) the educational concept. Currently there is an emphasis, starting on day 1, on self-regulated learning. Therefore skills such as being organized and being able to reflect on one's actions and to plan learning activities are important. Students often have less than 20 hours of instruction and coaching, which, assuming a study load of 40 hours per week, leaves a full 20 hours for self-study activities. Moreover, in later years, students often have 15 hours or less of instruction and coaching. This educational approach is counteractive to students who lack conscientiousness. Institutions can help students by placing less emphasis on self-regulation at the start of a program so that students have the opportunity to develop self-regulation skills gradually. For example, curricula could be very structured in the first year, somewhat less so in the second year, and students should be prepared to self-regulate more during the third and fourth academic years. A second curriculum based intervention may be altering the amount of lectures, skills training, internship training or team projects. The study reported in chapter 3 shows that conscientiousness has the lowest correlation with team projects. Based on this finding one may suggest more complex team projects with longer time spans. Team projects clearly have certain advantages like the opportunity for students to develop communication and leadership skills. However, performance on team projects is negatively correlated with openness to experience. When curious, open minded students underperform in team projects one wonders if the project deadlines tend to become the main objective of students rather than discussing and learning with and from each other. The clear deadlines may be interfering with deep learning. So implementing more complex team projects will be a challenge for curriculum developers because what is an advantage for one student may be a disadvantage for another student. A major policy change may be needed, namely more individually tailored education. In such an approach, students have the same learning objectives, however the paths along which they reach these learning goals can differ.

Students who score high on neuroticism could be assessed at the end of the year with a portfolio whereas other students, for example those who need additional pressure because they are less conscientiousness, could be frequently assessed with small, individual, or group assignments, and traditional exams. Critics may argue that such an approach is too costly and results from aptitude-treatment interactions (Kanfer & Ackerman, 1989) and personality-treatment interaction (Trown & Leith, 1975) are not encouraging. This is not an argument for individualized higher education, but there is a clear need for customized learning packages for groups of individuals with similar characteristics. Such an approach could benefit large institutions that have large numbers of students who come from diverse backgrounds and who suffer from low graduation rates.

The intercorrelations between the different GPA's are reported in chapter 2. The medium sized correlations between the GPA's were tested and revealed that the HRM curriculum is heterogenous in nature. Heterogeneity in types of learning environments is one of the basic principles of competence-based education. However, of note is the correlation between thesis and the other learning environments. The thesis is often considered to be the cumulative end product of student learning and should reflect the student's overall capability or level of knowledge. The modest correlations between thesis and the other learning environments suggests that the thesis is not the cumulative end product but should be regarded as a separate measure. This is probably caused by the individualistic, active nature of writing a thesis versus the more passive, collective and instruction driven nature of traditional learning environments and exams. If the thesis is to remain as the final exam of student learning, then curriculums should be designed so skills that are relevant to completing a thesis are learned in the earlier parts of a student's education. For example students, as individual or as a group, should be trained to write research proposals, short research articles or advisory reports, and even a paper that starts to approximate the basic thesis. As in universities, students are asked to write a bachelor thesis to acquire their bachelor diploma and to write a masters thesis as part of their master's degree. There is currently a debate in the Netherlands about whether to start testing student knowledge with a standardized exam at the end of the college education. Such an exam would ensure that standards at institutions of higher education are being maintained on an equal basis throughout the entire country. As reported in chapter 2, the intercorrelations between lectures and the other specific performance measures are above .40 (except for thesis), and lectures has the highest reliability estimate. From this perspective, a knowledge exam may be a better criterion than is a written thesis. However, from the social constructivist view of learning (chapter 1) implementing a knowledge exam at the end of a bachelor program is not desirable. After all, tests are known to control learning (Biggs, 2003) and the purpose of competence-based education is to learn how to apply knowledge in complex work situations rather

than merely possess knowledge. From this perspective a knowledge-based exam is undesirable.

Clearly, we need to know more about the construct validity of conscientiousness. The correlations between the NEO-FFI personality questionnaire and the intrinsic motivation scale of the Work Style Inventory is  $r = .45$ . A simple reason for this finding is that a number of items between the two scales of conscientiousness and intrinsic motivation have similar content. A factor analysis of the conscientiousness and intrinsic motivation items yields a three factor solution and showed that the overlapping items load on the same factor. As reported, adding motivational factors in a prediction model with personality traits can lead to the inaccurate perception that motivation is less important. The use of the full version, the NEO-PI-R, which measures five facets of conscientiousness, may prove to be a fruitful approach to unravel the overlap between conscientiousness and motivation. This is important for practice because teacher and student counsellors need to know how conscientiousness and motivation affect academic achievement so that they can develop and implement suitable interventions.

Recent research suggests that many young adolescents may not be fully equipped to act as self-regulated learners. The prefrontal cortex, which makes reflection and planning possible, is not fully developed until the age of 22 to 25 (Paus, 2005). So, if students need self-discipline to be successful in competence-based education and students are not fully equipped to plan and organize their own learning activities right from the start, then they have less chance to succeed. Research shows that the prefrontal cortex in men matures later than in women, so men might be more at risk during the early stages of their higher educational careers. Students should be made aware of the strengths and weaknesses in their own learning capabilities, style and personality, and they should act on that knowledge. However, simply informing students about their strengths and weaknesses may not be enough because they are not yet equipped to fully understand how these factors affect academic achievement. What could be helpful in this respect is peer-coaching by senior students in addition to the evidence-based information.

The average time students need to graduate is increasing which, in light of the current budget cuts, is an immediate problem for the government and a potential problem for students. Although a prolonged stay in college is not necessarily a waste, targets have been set to reduce study duration. In 2012, students who need more than five years to finish will be charged an extra 3,000 euro by the government. As is shown in chapter 8, openness to experience is related to a prolonged study duration. Why and how openness exerts a negative effect on study pace deserves further examination. A starting point may be that 50% of the students in higher education testify that they are not challenged enough by the content of the curriculum nor by their teachers, and therefore spend only 30 hours a week on school activities instead of the required 40 hours. Open and curious students may then be tempted to seek challenge outside the classroom which

may have a negative effect on their academic performance and cause them to stay in school longer. These students may also benefit from time-management courses to strengthen their self-discipline and to learn how to focus on their (learning) goal(s), or alternatively, may decide to choose a more challenging program and additional courses to help them stay focussed.

#### **9.4 Limitations, strengths, and suggestions for future research**

There are several limitations as well as strengths in the current research that warrant discussion. One limitation is that data are collected from a naturally occurring field study. Thus, no interventions or controlled experiments with random samples were possible. For example, it would have been interesting to compare directly the effects of traditional and competence-based curricula on measures of early career success and job performance. It also meant that existing data (grades) were used of which the independence of the measurements can not explicitly be guaranteed.

A second issue that limits how far the results may be generalized concerns how competencies are measured. Competency ratings in the current study stem from two large scale assessment centers that were specially designed for the HRM program by a commercial assessment firm. However competency ratings in higher education are also often derived by assessing student portfolios or through observation in the classroom or during coaching. Because of the differences between assessment center, portfolio, and classroom observation, the results presented in chapters 7 and 8 cannot be generalized to other assessment formats – at least not until the intercorrelations among the various competency measures are investigated. Research on the validity of portfolio competency ratings is recommended because the portfolio is widely used in the Netherlands and validity evidence is lacking (McMullan et al., 2003). A third issue is the collective nature of the grades students received for their team projects. Such group grades do not only reflect the individual's capabilities which limits the conclusions concerning the team projects criterion.

There are additional relevant issues that cannot be addressed in the current research. For example, there is no measure of class attendance, a variable that accounts for significant variance in GPA beyond that of intelligence. However, conscientiousness may be used as a surrogate for attendance because conscientious students will most likely attend more classes than their less conscientious colleagues. A minor limitation is that only pass grades rounded up to the nearest whole number were available. Clearly, GPA's based on pass grades that range from 6 to 10 have smaller standard deviations than GPA's based on first try grades that range from 1 to 10. Grades with decimal points (e.g., 6.5) might also be more meaningful because they contain less error.

Students who dropped-out of school are not included in the current research. Drop-out rates in higher education are high and dropping out has financial as well as psychological consequences for a student. The highest drop-out rates in the Netherlands occur during the first six to twelve months and individual differences measures (chapter 3) were administered during the assessment center (chapter 7) at the end of the first year of college.

Whether the HRM curriculum in this research is fully competence-based is an open question. The characteristics that are associated with competence-based education include: a) the use of authentic, open problems and learning materials that have personal meaning for students and are presented in a variety of formats, b) teaching methods that arouse interest, activate prior knowledge, clarify meanings, and model appropriate learning strategies and reflective processes, c) small group learning that facilitates competencies like teamwork, debating and reflection, and d) changes in the goal and use of assessments like integrating instruction, learning and assessment (e.g., Entwistle, 2000). The HRM curriculum uses different learning environments, including small group learning, and also promotes competency development through the student counselling program, the internships, and in the assessment centers. Moreover, competencies, along with learning goals, are described in the student manuals of each course. Based on these characteristics, the HRM program is in the early stage of a competence-based program. Although meaningful relationships between the learning environments (e.g., internships) and the competency ratings are reported, these correlations might be higher in a program that is fully competence-based. To that end, effect sizes in this dissertation may be regarded as being conservative.

A known problem with assessment center research is that competency ratings by assessors are often contaminated with bias, which causes correlations with criterion variables to be artificially inflated (chapter 8). There can be no rating contamination in the current research, except for the self-ratings of job performance, because supervisors of the HRM graduates would not have known how graduates performed in the assessment center while they were still students. That the current findings show meaningful relationships between AC competency ratings and performance appraisal scores at work strengthens the belief that competencies predict important real life outcomes.

Although gender differences are not pivotal in this dissertation, the findings show that female students in general outperform male students. Therefore, future research should consider the direct effect, and the incremental effect of gender on academic achievement.

Most empirical studies in the Netherlands use university samples, despite the fact that there are twice as many college, as university, students. Findings from the current study are based on a college sample and parallel those of studies conducted with university students.

### **9.5 Concluding remarks**

Businesses are in constant need of talent to stay competitive. Institutes of higher education try to provide individuals with the right qualifications to be successful in the labour market. Therefore there is a need for knowing what makes individuals successful in their careers. There is also a need to know what makes a student successful in college. This dissertation examines both the determinants of early career success and of academic achievement. Results show that intelligence, conscientiousness and intrinsic motivation are important predictors of academic and early career success. Academic achievement and especially competency ratings are also related to early career success. Knowing what key determinants of success are and acting upon that knowledge may improve the absolute number of students who graduate as well as time-to-graduation rates. Students, counsellors, teachers and managers need to be educated on what factors are related to academic and career success. All involved can use this knowledge to increase their chances of success.

# Samenvatting

## Doel en aanleiding van het onderzoek

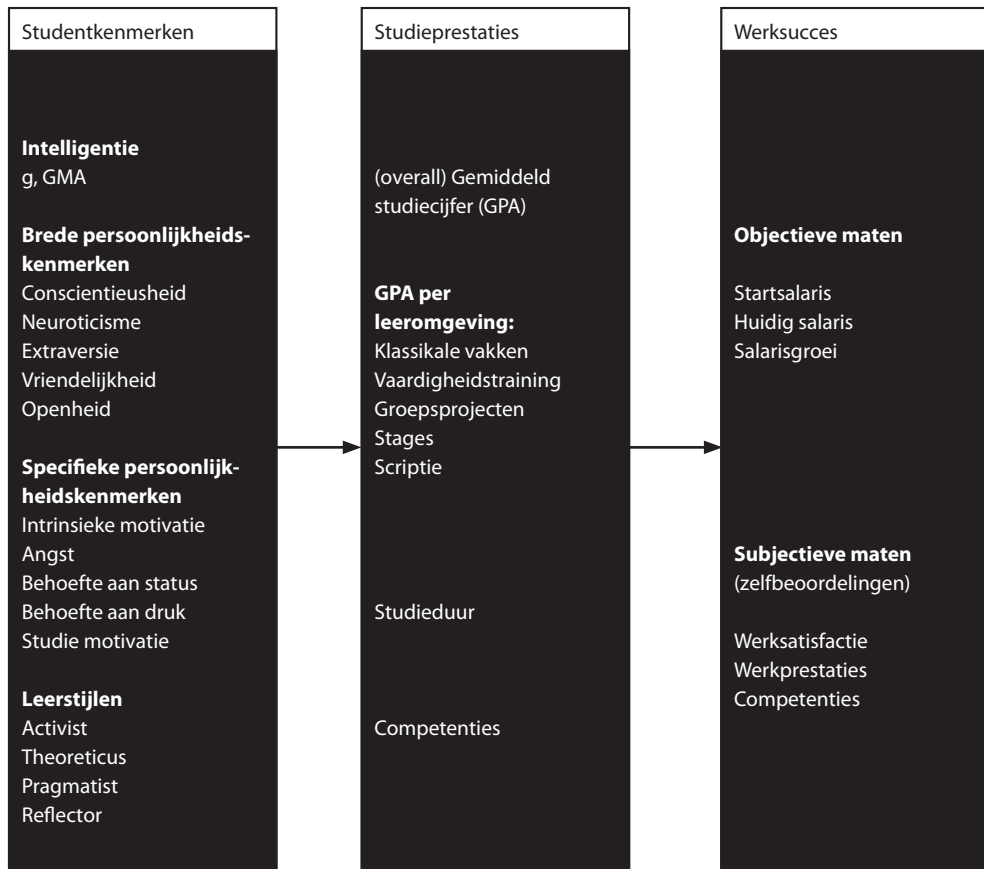
Kenniseconomieën hebben hoog opgeleide burgers nodig. Inzicht in de factoren die van invloed zijn op succes in het hoger onderwijs is daarmee belangrijker dan ooit. Overheden binnen de Europese Unie hebben zich tot doel gesteld om een toename te realiseren van het percentage hoger opgeleiden binnen hun beroepsbevolking. Hoewel er een positieve trend is betreffende het aantal inschrijvingen in het hoger onderwijs nemen gelijktijdig de slagingspercentages af (CBS, 2010). Daarmee wordt het minder waarschijnlijk dat de streefpercentages die de Nederlandse overheid heeft gesteld gehaald worden. De bevindingen uit dit proefschrift geven inzicht in die studentkenmerken die gerelateerd zijn aan het succesvol afronden van een hogere beroepsopleiding op HRM gebied. Tevens wordt onderzocht welke kenmerken gerelateerd zijn aan het latere werksucces van studenten en welke rol opleidingsprestaties daarbij spelen. Specifiek wordt de relatie tussen intelligentie, persoonlijkheid, motivatie, leerstijlen en competenties aan de ene kant, en academische prestaties en succes in de eerste fase van de carrière aan de andere kant, onderzocht.

De bevindingen zijn nuttig voor verschillende stakeholders: de wetenschap, het bedrijfsleven, onderwijsinstellingen en studenten. De bevindingen kunnen helpen bij het formuleren van beleid om het aantal afstudeerders te verhogen en hun studieduur te verlagen. Daarbij kan inzicht in hun prestaties en studieduur helpen om beter onderbouwde beslissingen te nemen over maatregelen, bijvoorbeeld op het gebied van interventies. Kennis over hoe student kenmerken en academische prestaties zijn gerelateerd aan een succesvolle carrièrestart is van belang voor studenten en organisaties binnen het bedrijfsleven. Studenten kunnen werken aan hun ontwikkelingsbehoeften in de veiligheid van de onderwijssetting en bedrijven kunnen betere selectie beslissingen nemen.

Drie vragen staan in het onderzoek, beschreven in dit proefschrift, centraal:

1. Hoe goed voorspellen studentkenmerken studiecijfers?
2. Hoe goed voorspellen studentkenmerken de tijd die studenten over hun studie doen?
3. Hoe goed voorspellen studentkenmerken en studieprestaties een succesvolle start van de carrière?

Een overzicht van alle variabelen in de huidige studie wordt gegeven in Figuur 1, het onderzoeksmodel.



Figuur 1 Onderzoeksmodel



## Samenvatting van de belangrijkste bevindingen

De onderzochte studentkenmerken in de huidige studie zijn intelligentie, de “Big Five” persoonlijkheidskenmerken, motivatie, specifieke persoonlijkheidskenmerken en leerstijlen. Academische prestaties en werksucces in de eerste jaren van de carrière zijn de twee centrale afhankelijke variabelen. Het gemiddelde van alle studiecijfers (GPA), cijfers binnen vijf specifieke leeromgevingen en studieduur tot diplomering, zijn de maten van academisch succes, terwijl salaris en zelfrapportages als maten voor carrièresucces fungeren. De verschillende studies in dit proefschrift richten zich op de vraag hoe goed verschillende voorspellers variantie verklaren in de afhankelijke variabelen.

### 1. *Hoe goed voorspellen studentkenmerken studiecijfers?*

Deze vraagstelling is erop gericht de relatie tussen studentkenmerken en studiecijfers vast te stellen. Onderzocht is de rol van intelligentie, de “Big Five” persoonlijkheidskenmerken (conscientieusheid, neuroticisme, extraversie, vriendelijkheid en openheid), motivatie, angst, behoefte aan druk, behoefte aan status, studiemotivatie en leerstijlen (activist, theoreticus, pragmatist, reflector).

#### *Intelligentie*

De bevindingen suggereren een bescheiden maar stabiele positieve correlatie tussen intelligentie en studiecijfers tijdens de studie (hoofdstuk 3 en 4). Intelligentie is gerelateerd aan het eerste jaars GPA evenals het overall GPA (hoofdstuk 6). Dit resultaat is in lijn met andere onderzoeksresultaten met betrekking tot de predictieve validiteit van intelligentie voor GPA in universitaire populaties, zowel binnen Nederland (Busato et al., 2000; Resing & Drenth, 2009; Wolting, 2006) als daarbuiten (Farsides & Woodfield, 2003; Furnham & Chamorro-Premuzic, 2004). Er wordt vaak beweerd dat het ontbreken van een sterke correlatie een artefact is en te wijten aan “restriction of range” (Drenth, 2004) dat wil zeggen een beperking van de variantie in de intelligentie variabele doordat hoofdzakelijk slimme studenten weten door te dringen tot het hoger onderwijs.

#### *De “Big Five” persoonlijkheidskenmerken*

De belangrijkste bevinding is dat de persoonlijkheidseigenschap conscientieusheid hoger correleert met het gemiddelde van alle studiecijfers (GPA) dan intelligentie (hoofdstuk 3). Conscientieusheid blijkt, na correctie voor intelligentie, 22% van de variantie in GPA te verklaren en is tevens consistent gerelateerd met elk van de vijf specifieke academische prestatiecriteria. In feite, verklaart conscientieusheid vijf keer zoveel variantie in academische prestaties als intelligentie (hoofdstuk 4). Een extra set van studies in andere academische disciplines bevestigt het belang van conscientieusheid voor studieprestaties in het eerste jaar. Zo zijn er vergelijkbare validiteitscoëfficiënten in een

tweede HRM steekproef en in een Sport & Bewegen steekproef gevonden (hoofdstuk 6). Veel studies melden dat consciëntieusheid is gerelateerd aan academische prestaties (Bratko et al., 2006; Gilles & Bailleux, 2001; Nofle & Robins, 2007; Poropat, 2009), of dat consciëntieusheid een grotere voorspellende waarde heeft dan intelligentie (Conard, 2006; Di Fabio & Busoni, 2007; Furnham & Chamorro-Premuzic, 2004; Furnham, et al., 2003; Petrides et al., 2005). Consciëntieuze studenten presteren beter, omdat ze langer volharden en beter georganiseerd zijn dan hun tegenhangers. Voor de andere persoonlijkheidskenmerken wordt sporadisch een relatie met studiecijfers gevonden: neurotische studenten presteren beter in de vaardigheidstraining, extraverten presteren minder goed in vaardigheidstraining en "open minded" studenten presteren minder goed in groepsprojecten (hoofdstuk 3).

### *Intrinsieke motivatie*

Intrinsieke motivatie correleert matig met GPA en met eerste jaar GPA (hoofdstukken 4 en 6). Intrinsieke motivatie verklaart meer dan twee keer zo veel variantie in GPA als intelligentie en heeft incrementele validiteit boven intelligentie. De omvang van het effect komt overeen met meta-analytische bevindingen (Robbins et al., 2004; Vallerand et al., 1993). Echter, intrinsieke motivatie heeft geen incrementele validiteit boven consciëntieusheid. Zoals opgemerkt in hoofdstuk 4, zijn motivatie en consciëntieusheid sterk aan elkaar gerelateerd en uit een analyse van de inhoud van de items blijkt dat beide 'volhardendheid' meten. Door het ontbreken van incrementele validiteit, zodra consciëntieusheid constant wordt gehouden, kan de misvatting ontstaan dat intrinsieke motivatie studieprestaties niet voorspelt. Natuurlijk is niets minder waar.

### *Specifieke persoonlijkheidskenmerken*

Het effect van angst, behoefte aan druk, behoefte aan status, en studiemotivatie op studieprestaties is onderzocht. De bevindingen laten zien dat angst niet gerelateerd is aan GPA. Behoeft aan druk correleert negatief, en studiemotivatie correleert positief, met GPA (hoofdstuk 4). Studiemotivatie blijkt voorts van belang voor prestaties in alle typen leeromgevingen. Behoeft aan druk toont een negatieve correlatie met GPA en prestaties binnen de verschillende typen leeromgevingen en in het bijzonder de scriptie. Studenten die behoefte hebben aan extra druk van buitenaf behalen in het algemeen lagere cijfers voor hun scriptie.

### *Leerstijlen*

De bevindingen gerapporteerd in hoofdstuk 5 laten zien dat de leerstijlen van de Learning Style Questionnaire van Honey en Mumford (1992), namelijk activist, theoreticus, pragmatist en reflector niet zijn gerelateerd aan GPA, noch aan prestaties binnen de vijf specifieke leeromgevingen. Dit wekt twijfel aan het belang van leerstijlen voor academische prestaties, in tegenstelling tot de overtuiging van vele docenten.

Samengevat zijn de studentkenmerken die gerelateerd zijn aan academische prestaties in volgorde van belangrijkheid: consciëntieusheid, motivatie en intelligentie (hoofdstuk 3-6). Gecombineerd verklaren de studentkenmerken 33% van de variantie in GPA, waarbij consciëntieusheid en intelligentie het meest bijdragen. Dit betekent dat intelligente, gedreven, gedisciplineerde en hard werkende studenten de beste studieprestaties leveren in termen van het halen van hoge cijfers.

## *2. Hoe goed voorspellen studentkenmerken studieduur?*

Deze vraagstelling is erop gericht te achterhalen wat de relatie is tussen studentkenmerken en de tijd die studenten over hun studie doen (studieduur). Onderzocht is de rol van intelligentie, de “Big Five” persoonlijkheidskenmerken, motivatie, angst, behoefte aan druk, behoefte aan status, studiemotivatie en leerstijlen.

### *Intelligentie*

De resultaten die gepresenteerd worden in hoofdstuk 4 laten een beperkte relatie zien tussen intelligentie en studieduur. Intelligente studenten lijken gemiddeld maar iets korter over hun studie te doen dan minder intelligente studenten.

### *De “Big Five” persoonlijkheidskenmerken*

Verreweg de belangrijkste voorspeller van studieduur is, wederom, consciëntieusheid dat ongeveer 17% van de variantie verklaart. Dit is in lijn met de resultaten van meta-analytisch onderzoek door Nofle en Robins (2007), O’Conner en Paunonen (2007) en Poropat (2009) en ook met de resultaten van studies in Nederland door Van Bragt (2010) en Busato (2000), die aangeven dat consciëntieusheid een bepalende factor voor de studievoortgang is in het eerste jaar binnen het HBO (Van Bragt, 2010) en het eerste jaar binnen de universiteit (Busato, 2000). Neuroticisme en openheid zijn enigszins gerelateerd aan studieduur (hoofdstukken 3 en 4). Dit betekent dat gedisciplineerde en hard werkende studenten minder tijd nodig hebben om af te studeren en dat open, nieuwsgierige maar ook emotioneel minder stabiele studenten meer tijd nodig hebben om af te studeren.

### *Intrinsieke motivatie*

Intrinsieke motivatie verklaart twee keer zoveel variantie als intelligentie in de tijd die studenten nodig hebben tot diplomering. Intrinsieke motivatie heeft tevens incrementele validiteit boven intelligentie (hoofdstuk 4). Daarbij moet wel worden opgemerkt dat intrinsieke motivatie en consciëntieusheid sterk met elkaar samenhangen.

### *Specifieke persoonlijkheidskenmerken*

Geen van de overige specifieke persoonlijkheidskenmerken, te weten angst, behoefte aan druk, behoefte aan status en studiemotivatie is gerelateerd aan studieduur.

### *Leerstijlen*

De bevindingen uit hoofdstuk 5 laten zien dat er geen significante relatie is tussen de vier leerstijlen van de LSQ en studieduur.

Samengevat zijn de studentkenmerken die gerelateerd zijn met studieduur in volgorde van belangrijkheid: consciëntieusheid, openheid, intrinsieke motivatie, neuroticisme en intelligentie (hoofdstukken 3-6). Dit houdt in dat gedreven, gedisciplineerde, gemotiveerde en hard werkende studenten minder tijd nodig hebben om hun diploma te halen en dat open, nieuwsgierige en emotioneel minder stabiele studenten over het algemeen meer tijd nodig hebben tot diplomering. Gecombineerd verklaren de studentkenmerken 30% van de variantie in studieduur, waarbij consciëntieusheid en openheid het meest bijdragen.

### *3. Hoe goed voorspellen studentkenmerken en studieprestaties een succesvolle start van de carrière?*

Factoren die samenhangen met succes in het begin van de carrière zijn onderzocht en uitkomsten worden gerapporteerd in hoofdstuk 8. Objectieve maten van carrièresucces zijn het startsalaris, het huidige salaris en de salariscroei. Subjectieve maten zijn voldoening in het werk, een zelfbeoordeling van de werkprestaties en een zelfbeoordeling van het competentieniveau. Intelligentie, brede en specifieke persoonlijkheidskenmerken en leerstijlen zijn ook hier gebruikt als voorspellende variabelen en daar zijn de academische succescriteria van GPA, competenties en studieduur aan toegevoegd.

### *Studentkenmerken*

Intelligentie verklaart 8% van de variantie in huidig salaris en 10% in salariscroei, maar geen variantie in startsalaris, de twee zelfbeoordelingen (werkprestaties en competenties) of voldoening in het werk. Dit is in lijn met de onderzoeksresultaten betreffende de voorspellende waarde van intelligentietests voor werksituaties (o.a. Hunter & Hunter, 1984; Künzle et al., 2004; Salgado et al., 2003; Schmidt & Hunter, 1998). Intelligentie heeft incrementele validiteit, boven GPA en een overall score van competenties (OAR), voor huidig salaris en salariscroei. Van de "Big Five" persoonlijkheidskenmerken correleren alleen extraversie en consciëntieusheid significant met werksucces en dan alleen met de objectieve maten. Consciëntieusheid voorspelt salariscroei, maar niet nadat gecontroleerd is voor GPA en de OAR. Met andere woorden, er is minder kans op een incrementele bijdrage boven GPA omdat GPA en consciëntieusheid sterk zijn gecorreleerd.

Van de andere voorspellende variabelen correleren de volgende elk significant met twee objectieve carrièresucces maten: intrinsieke motivatie, angst, studiemotivatie, de activistische leerstijl en de pragmatische leerstijl. De pragmatische leerstijl voorspelt drie

werksucces variabelen (hoofdstuk 8) na gecontroleerd te hebben voor GPA, de OAR en intelligentie. Dat de pragmatische leerstijl incrementele validiteit had was onverwacht omdat deze stijl niet gerelateerd is aan academische prestaties in het HRM-programma (hoofdstuk 5) en omdat een review door Coffield et al. (2004) weinig aanwijzingen biedt dat leerstijlen predictieve validiteit hebben voor werkprestaties. Desalniettemin voorspelt pragmatisme het huidige salaris en de zelfbeoordeling van competenties boven de andere voorspellende variabelen.

### *Academische prestaties*

#### *GPA*

Het gemiddelde van alle studiecijfers (GPA) verklaart 8% van de variantie in huidig salaris en 12% in salarisgroei. De percentages zijn van dezelfde grootte als de percentages die verklaard worden door intelligentie.

#### *Prestaties in specifieke leeromgevingen*

De bevindingen laten zien dat de GPA's die gebaseerd zijn op prestaties in de specifieke leeromgevingen samen 12% van de variantie verklaren in startsalaris, 12% van de variantie in het huidige salaris en 15% van de variantie in salarisgroei. Startsalaris wordt het best voorspeld door cijfers die behaald zijn voor klassikale vakken en het huidige salaris wordt het best voorspeld door cijfers voor klassikale vakken en vaardigheidstrainingen. Voor salarisgroei zijn de cijfers die zijn behaald in alle vijf de specifieke leeromgevingen relevant. Echter, de incrementele validiteit van GPA boven de andere variabelen is zeer beperkt.

#### *Studieduur*

Studieduur correleert significant negatief met huidig salaris en marginaal negatief met salarisgroei.

#### *Competenties.*

De overall competentiebeoordeling (OAR), die een gemiddelde is van negen beoordeelde competenties op basis van twee grootschalige assessment centers (hoofdstuk 7), is gerelateerd aan de drie objectieve maten van werksucces en verklaart variantie in het startsalaris (4%), het huidige salaris (12%) en de salarisgroei (11%). Verdere analyses tonen aan dat de OAR een incrementele bijdrage heeft boven GPA op vier van de zes maten van werksucces. De competentiebeoordelingen laten een sterkere correlatie zien met de werksucces variabelen dan GPA. Hoewel de overall competentiebeoordeling enigszins correleert met GPA, verklaart ze nog altijd unieke variantie in werksucces boven GPA. Het resultaat komt overeen met een van de weinige studies naar de incrementele validiteit van competenties boven GPA door Waldman en Korbar (2004). In het algemeen overtreffen

competenties dus GPA als een voorspeller van werksucces, wanneer gecontroleerd is voor studentkenmerken. Dit resultaat ondersteunt de hypothese van Donofrio en Davis (1997) dat competenties beter correleren met prestaties op het werk dan studiecijfers.

Gezamenlijk verklaart de set van voorspellende variabelen tot 29% van de variantie in bepaalde werksucces variabelen, zoals huidig salaris. De voorspellende variabelen verklaren meer variantie in de objectieve dan in de subjectieve maten van werksucces. Competenties hebben incrementele validiteit voor de gebruikte maten van werksucces, na correctie voor de andere voorspellers. Deze bevinding geeft een duidelijke steun aan de waarde van competentie gerelateerde onderwijsmethoden.

### Aanbevelingen

Een van de doelen van dit proefschrift, zoals opgesomd in hoofdstuk 1, is om praktische aanbevelingen te doen voor de verschillende partijen die kunnen profiteren van dit onderwijs-onderzoek: organisaties binnen het bedrijfsleven, onderwijsinstellingen en studenten. Hiermee wordt gepoogd de kloof tussen kennis en toepassing te overbruggen. Bedrijven moeten, waar mogelijk, competenties, intelligentie en consciëntieusheid als aspecten opnemen in hun selectieprocedure. Uit onderzoek blijkt dat interviewers hogere beoordelingen geven aan sollicitanten die extravert gedrag vertonen (Barrick, Dustin, Giluk, Stewart, Shaffer, & Swider, in druk). Dat extraverten het beter doen in interviews is niet ongevoel, omdat het interview een sociale situatie is en extraverten het over het algemeen beter doen in sociale situaties. De resultaten uit hoofdstuk 8 laten zien dat extraversie gecorreleerd is met het startsalaris. Echter, salarismet is een betere maatstaf voor capaciteiten van een medewerker dan het startsalaris en de salarismet wordt niet voorspeld door extraversie maar door competenties, intelligentie en consciëntieusheid. Omdat extraversie niet gerelateerd is aan consciëntieusheid,  $r = .05$  (hoofdstuk 3) en niet gerelateerd is aan competenties,  $r = .03$  (hoofdstuk 7), verliezen bedrijven mogelijk veel potentieel goede werknemers door te veel te focussen op extraversie en te weinig op competenties en consciëntieusheid.

Instellingen voor hoger onderwijs die gebruik maken van toelatingstests kunnen overwegen om een persoonlijkheidstest of motivatievragenlijst te gebruiken. Enige terughoudendheid is dan, op basis van de percentages verklaarde variantie, wel op zijn plaats. Niet alleen blijft veel variantie in studiesucces onverklaard, tevens dient in acht genomen te worden dat de studenten in dit onderzoek de vragenlijsten belangeloos invulden wat in geval van gebruik ten behoeve van selectie heel anders ligt. Instellingen die een open toelatingsbeleid hebben en daarom niet bereid of in staat zijn om te selecteren, kunnen overwegen om dergelijke tests af te nemen in hun reguliere studieloopbaan

programma's. Studenten die laag scoren op consciëntieusheid lopen het risico om niet op tijd af te studeren en die studenten kunnen op deze manier worden 'gesignaleerd', zodat de juiste interventies tijdig kunnen worden ingezet. In het algemeen zijn er twee soorten strategieën die gebruikt kunnen worden om studenten te begeleiden. De eerste strategie is erop gericht de aan consciëntieusheid gerelateerde vaardigheden te ontwikkelen, zoals plannen en timemanagement. Een tweede strategie, die momenteel veel besproken wordt in Nederland is aanpassingen te maken in de leeromgeving of (elementen van) het onderwijsconcept. Binnen deze strategie past het wijzigen van het aantal klassikale vakken, stages of projecten, het stapsgewijs invoeren van zelfregulerend leren en meer op maat samengesteld onderwijs.

De correlaties tussen de verschillende GPA's zijn beschreven in hoofdstuk 2. De correlaties tussen de GPA's zijn getoetst en daaruit bleek dat het HRM-curriculum samengesteld is uit diverse typen leeromgevingen. Deze heterogeniteit in typen leeromgevingen is één van de basisprincipes van competentiegericht onderwijs en biedt studenten afwisseling, variatie en uitdaging en biedt hen de mogelijkheid om hun talenten te ontwikkelen. Echter één punt vraagt extra aandacht en dat is de zwakke correlatie tussen het cijfer van de scriptie en de overige cijfers. De scriptie wordt beschouwd als het cumulatieve eindproduct van het leren en zou daarmee de capaciteiten, kwaliteiten en het totale kennisniveau van de student moeten weerspiegelen, echter uit dit onderzoek blijkt dat de scriptie beschouwd moet worden als een afzonderlijk assessment. Aanbevolen wordt om onderwijsprogramma's zo te ontwerpen dat de competenties, kennis en vaardigheden die relevant zijn voor de voltooiing van een scriptie geleerd zijn in de eerdere fase van de opleiding van een student. Er is momenteel een debat in Nederland over de vraag of er in het hoger beroepsonderwijs aan het einde van de opleiding een landelijk examen moet worden ingevoerd dat (delen van) de kennisbasis van elke student toetst. Zo'n examen zou ervoor moeten zorgen dat de normen aan instellingen voor hoger onderwijs worden gehandhaafd op een gelijke basis in het hele land. Zoals beschreven in hoofdstuk 2 zijn de correlaties tussen de cijfers die behaald zijn in de klassikale vakken die op kennisoverdracht gericht zijn en prestaties in de andere leeromgevingen boven .40 (met uitzondering van de scriptie), en heeft dit GPA de hoogste betrouwbaarheid. Vanuit dit perspectief is een kennisexamen een beter criterium dan een scriptie. Echter, vanuit de sociaal constructivistische visie op leren (hoofdstuk 1) is de implementatie van een kennisexamen aan het einde van een bacheloropleiding onwenselijk. Immers, van tests is bekend dat zij het leren sturen (Biggs, 2003) en het doel van competentiegericht onderwijs is om kennis toe te passen in complexe werksituaties en niet louter kennis (re)productie. Vanuit dit perspectief is een op kennis gebaseerd landelijk examen minder gewenst.

De gemiddelde tijd die studenten nodig hebben tot diplomering stijgt en dat is in het licht van de huidige bezuinigingen een direct probleem voor de overheid en een potentieel probleem voor studenten. Hoewel een verlengd verblijf op de universiteit niet per definitie

een verspilling is, zijn er doelen gesteld om te studieduur te verminderen. In de nabije toekomst zal aan studenten die behoefte hebben aan meer dan vijf jaar een lang-studeerboete in rekening gebracht worden door de overheid. Zoals te lezen is in hoofdstuk 8 is dat nadelig voor open, nieuwsgierige en neurotische studenten en studenten die moeite hebben met plannen en organiseren.

### Beperkingen

Er zijn een aantal beperkingen in het huidige onderzoek aan te wijzen. Eén beperking is dat gegevens zijn verzameld binnen de reguliere onderwijsuitvoering. Zo is er ondermeer gebruik gemaakt van bestaande gegevens (cijfers) waarbij de onafhankelijkheid van de metingen niet expliciet gegarandeerd is. Tevens kon er alleen met voldoende (>6) gewerkt worden en tevens met cijfers die afgerond waren. Daarnaast waren er geen ingrepen of gecontroleerde experimenten met aselechte steekproeven mogelijk. Een tweede kwestie die beperkend is voor de mate waarin de resultaten kunnen worden gegeneraliseerd is hoe competenties zijn gemeten. Competentiebeoordelingen in de huidige studie zijn gebaseerd op twee grootschalige assessment centers die speciaal zijn ontworpen voor het HRM-programma, terwijl competenties ook vaak beoordeeld worden met behulp van een student's portfolio. Een derde punt is het collectieve karakter van de cijfers die studenten kregen voor hun groepsprojecten. Een dergelijk groeps cijfer reflecteert niet alleen maar de individuele capaciteiten van het individu. Er zijn nog meer relevante onderwerpen die niet konden worden opgepakt in het huidige onderzoek. Bijvoorbeeld, er is geen meting gebruikt betreffende collegebezoek terwijl bekend is dat collegebezoek net zoveel variantie in cijfers kan verklaren als intelligentie. Gesteld kan echter worden dat we een surrogaat gebruikt hebben voor het aanwezig zijn omdat consciëntieuze studenten waarschijnlijk meer lessen volgen dan hun minder consciëntieuze medestudenten. Studenten die zijn uitgevallen, hoofdzakelijk in het eerste studiejaar, zijn niet opgenomen in het huidige onderzoek omdat dergelijk onderzoek al verricht is en dit onderzoek zich specifiek richt op aspecten die gerelateerd zijn aan succes. Daarbij zijn de gegevens betreffende de studentkenmerken verzameld aan het eind van het eerste studiejaar. Hoewel de verschillen tussen de geslachten niet centraal staan in dit proefschrift, lijken de bevindingen erop te wijzen dat vrouwelijke studenten in het algemeen beter presteren dan mannelijke studenten. Daarom zou toekomstig onderzoek rekening moeten houden met de directe en de indirecte effecten van geslacht op academische- en werkprestaties.



**Tot slot**

Organisaties binnen het bedrijfsleven hebben een constante behoefte aan talent om competitief te blijven en het is aan onderwijsinstellingen om daaraan bij te dragen door talenten tot ontwikkeling te laten komen. Studenten beschikken idealiter na hun hogere beroepsopleiding over de juiste competenties om succesvol te kunnen zijn op de arbeidsmarkt. Daarom is er een behoefte om kennis te verzamelen over wat mensen succesvol maakt in hun (studie)carrière. Dit proefschrift onderzoekt zowel de determinanten van academisch succes als werksucces en heeft aangetoond dat intelligentie, consciëntieusheid en intrinsieke motivatie belangrijke determinanten voor succes zijn. Voor werksucces blijkt dat competenties belangrijker zijn dan studiecijfers. Kennis van de belangrijkste determinanten van succes en op basis daarvan handelen kan leiden tot een stijging van het absolute aantal studenten dat afstudeert alsmede een afname van de tijd die zij over hun studie doen. Studenten, studiebegeleiders, docenten en managers kunnen daar hun voordeel mee doen.



## Appendices

### Appendix A

Table A *Intercorrelations of specific performance criteria corrected for attenuation*

	Lectures	Training	Projects	Internship	Thesis
Lectures	-				
Skills training	.94	-			
Projects	.62	.47	-		
Internship	.63	.78	.66	-	
Thesis	.42	.44	.46	.61	-

Note. All correlations are significant at  $p < .01$

Appendix B

Table B *Results of Lisrel analysis*

	$\chi^2$	df	p	Fit indices			Path with smallest T-value
				SRMR	RMSEA	CFI	
Full model	0	0	1.00		Perfect fit		Agreeableness-GPA (.01)
Model A	.0002	1	1.00		Perfect fit		Need-for-status-TTG (.08)
Model B	.007	2	1.00	.0005	.00	1.00	Need-for-status-GPA (-.10)
Model C	.002	1	.97		Perfect fit		Anxiety-TTG (-.14)
Model D	.02	2	.99	.001	.00	1.00	Pressure-TTG (.26)
Model E	.01	3	.99	.002	.00	1.00	Anxiety-GPA (-.35)
Model F	.07	2	.97	.002	.00	1.00	Agreeableness -TTG (.08)
Model G	.03	1	.87	.001	.00	1.00	Intrinsic-GPA (.55)
Model H	.35	2	.84	.004	.00	1.00	Openness-GPA (-.86)
Model I	1.14	3	.77	.009	.00	1.00	Neuroticism-GPA (.93)
Model J	2.07	4	.72	.01	.00	1.00	Extroversion-TTG (-.95)
Model K	3.02	5	.70	.01	.00	1.00	Extroversion-GPA (.89)
Model L	1.29	4	.86	.01	.00	1.00	Study motivation-TTG (1.29)
Model M	3.04	5	.69	.01	.00	1.00	Intrinsic-TTG (-1.43)
Model N	.83	4	.93	.009	.00	1.00	Study motivation-GPA (-1.51)
Model O	1.23	3	.75	.01	.00	1.00	All T > 2.00

Note. SRMR = Standardized root-mean-square residual; RMSEA = Root-mean-square error of approximation; CFI = Comparative fit index.

The degrees of freedom in Table B fluctuate due to the fact that subsequent models are nested within each other. Where the degrees of freedom decrease, there is a loss of a variable from the model, in that case the variable has no significant contribution for both GPA and TTG. For example in model A and model B the need-for-status variable had the smallest T-value and was therefore completely dropped, leading to a decrease in the degrees of freedom in model C.

Appendix C

Table C *Confirmatory factor analysis results for AC1 and AC3*

	Model	Configuration	$\chi^2$	<i>df</i>	$\chi^2 / df$	SRMR	RMSEA	CFI
AC1	Model 1	1 factor 4 exercises	74.08	43	1.73	.13	.12	.81
	Model 2	3 factors 4 exercises	69.32	40	1.73	.18	.13	.85
	Model 3	1 factor 0 exercises	209.39	55	3.81	.29	.32	.23
	Model 4	3 factors 0 exercises	167.73	52	3.23	.25	.24	.42
	Model 5	0 factors 4 exercises	108.96	56	1.95	.23	.13	.73
AC3	Model 1	1 factor 4 exercises	59.12	76	.78	.07	.00	1.00
	Model 2	3 factors 4 exercises	56.63	73	.78	.11	.00	1.00
	Model 3	1 factor 0 exercises	393.08	91	4.32	.27	.29	.39
	Model 4	3 factors 0 exercises	288.55	88	3.28	.25	.22	.60
	Model 5	0 factors 4 exercises	107.00	91	1.18	.18	.04	.97

Note. SRMR = Standardized root-mean-square residual; RMSEA = Root-mean-square error of approximation; CFI = Comparative fit index.

In both AC's the models 1 and 2 have the lowest chi-square indicating the best fit for a model with 4 exercises and 1 or 3 factors. When model 1 is compared to model 2 the results indicate that adding additional factors does not lead to a significant improvement of the model . Table A further shows that the exercises are important however when model 1, with one factor, is compared to model 5, with no factors, the result indicates that a 1 factor, 4 exercise model fits the data best.



## List of abbreviations

AC	Assessment Center
CBE	Competence Based Education
EC	European Credits
ECTS	European Credit Transfer System
GMA	General Mental Ability
GPA	Grade Point Average
HE	Higher Education
HBO	Hoger BeroepsOnderwijs ( <i>Higher Professional Education</i> )
HEI	Higher Education Institute
HRM	Human Resource Management
IQ	Intelligence Quotient
LSQ	Learning Style Questionnaire
MBO	Middelbaar BeroepsOnderwijs ( <i>Secondary Vocational Education</i> )
MCT-H	Multicultural Test of Intellectual Ability for Higher education
OAR	Overall Assessment Center Rating
TTG	Time-To-Graduation





## References

- Ackerman, P. L., Bowen, K. R., Beier, M. B., & Kanfer, R. (2001). Determinants of individual differences and gender differences in knowledge. *Journal of Educational Psychology, 93*, 797-825.
- Ackerman, P. L., & Heggstad, E. D. (1997). Intelligence, personality, and interests: Evidence for overlapping traits. *Psychological Bulletin, 121*, 219-245.
- Akkerman, A. E., & Van der Linden, M. C. (2000). *Handleiding WerkStijlenVragenlijst* [Manual: Questionnaire About How One Works]. Leiden: Pits.
- Allinson, C. W., & Hayes, J. (1988). The Learning Styles Questionnaire: an alternative to Kolb's Inventory. *Journal of Management Studies, 25*, 269-281.
- Arcuri, A., Daly, W., & Mercado, P. (1982). Academic boot camp. In K. V. Lauridgen & C. Myers (Eds.), *Summer programs for underprepared freshman: New directions for college learning assistance* (p. 13-20). San Francisco: Jossey-Bass.
- Arguelles, A., & Gonczy, A. (Eds.). (2000). *Competency based education and training: a world perspective*. Balderas, Mexico: Editorial Limusa S.A. de C.V. Grupo Noriego Editores.
- Arthur, W., Day, E. A., McNelly, T. L., & Edens, P. S. (2003). A meta-analysis of the criterion-related validity of assessment center dimensions. *Personnel Psychology, 56*, 125-154.
- Baartman, L. K. J., Bastiaans, T. J., & Kirschner, P. A. (2004, June). *Requirements for Competency Assessment Programmes*. Paper presented at the Onderwijs Research Dagen 2004, Utrecht, The Netherlands.
- Bacon, D. R., & Bean, B. (2006). GPA in research studies: an invaluable but neglected opportunity. *Journal of Marketing Education, 28*, 35-42.
- Baird, L. L. (1985). Do grades and tests predict adult accomplishment? *Research in Higher Education, 23*, 3-85.
- Barrick, M. R., Dustin, S. L., Giluk, T., Stewart, G. L., Shaffer, J. A., & Swider, B. W. (in press). Candidate characteristics driving initial impressions during rapport building: Implications for employment interview validity. *Journal of Occupational and Organizational Psychology*.
- Barrick, M. R., & Mount, M. K. (1991). The Big Five personality dimensions and job performance: A meta-analysis. *Personnel Psychology, 44*, 1-26.
- Barrick, M. R., Mount, M. K., & Judge, T. (2001). Personality and performance at the beginning of the new millennium: What do we know and where do we go next? *International Journal of Selection and Assessment, 9*, 9-30.
- Barritt, L. S. (1966). Note: The consistency of first-semester college grade point average. *Journal of Educational Measurement, 3*, 261-262.

- Bartels, L. K., Bommer, W. H., & Rubin, R. S. (2000). Student performance: assessment centers versus traditional classroom evaluation techniques. *Journal of Education for Business*, 75, 198-201.
- Bean, J. P. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12, 155-187.
- Bean, J. P., & Metzner, B. S. (1985). A conceptual model of non-traditional undergraduate student attrition. *Review of Educational Research*, 55, 485-540.
- Bidjerano, T., & Dai, D. Y. (2007). The relationship between the big-five model of personality and self-regulated learning strategies. *Learning and Individual Differences*, 17, 69 - 81.
- Biemans, H., Nieuwenhuis, L., Poell, R., Mulder, M., & Wesselink, R. (2004). Competence-based VET in The Netherlands: backgrounds and pitfalls. *Journal of Vocational Education and Training*, 56, 523-538.
- Biggs, J. B. (2003). *Teaching for quality learning at university* (2nd ed.). Buckingham: Society for Research into Higher Education and Open University Press.
- Bleichrodt, N., & Van den Berg, R. H. (1999). *Handleiding MCT-H* [Manual: MCT-H]. Amsterdam: NOA/VU.
- Boekaerts, M. (1995). The interface between intelligence and personality as determinants of classroom learning. In D.H. Saklofske & M. Zeidner (Eds.), *Handbook of Personality and Intelligence* (p. 161-183). New York: Plenum.
- Boyatzis, R. E., Stubbs, E. C., & Taylor, S. N. (2002). Learning cognitive and emotional intelligence competencies through graduate management education. *Academy of Management Learning and Education*, 2, 150-162.
- Bratko, D., Chamorro-Premuzic, T., & Saks, Z. (2006). Personality and school performance: incremental validity of self- and peer-ratings over intelligence. *Personality and Individual Differences*, 41, 131-142.
- Bretz, R. D. (1989). College grade point average as a predictor of adult success. *Public Personnel Management*, 18, 11-22.
- Bruinsma, M. (2003). *Effectiveness of higher education: Factors that determine outcomes of university education*. Doctoral dissertation. Groningen: Rijksuniversiteit Groningen.
- Busato, V. V. (1998). *Leerstijlen nader geanalyseerd* [Learning styles explored]. Doctoral dissertation. Amsterdam: University of Amsterdam.
- Busato, V. V., Prins, F. J., Elshout J. J., & Hamaker, C. (2000). Intellectual ability, learning style, personality, achievement motivation and academic success of psychology students in higher education. *Personality and Individual Differences*, 29, 1057-1068.
- Campbell, D. T., & Fiske, D. W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cattell, R. B. (1965). *The Scientific Analysis of Personality*. Penguin: Harmondsworth.
- CBS. (2010). *Jaarboek onderwijs in cijfers 2010* [Yearbook education in numbers 2010]. Den Haag/Heerlen: Centraal Bureau voor de Statistiek.

- Chamorro-Premuzic, T. (2007). *Personality and Individual Differences*. Malden: Blackwell Publishers.
- Chamorro-Premuzic, T., & Furnham, A. (2003a). Personality predicts academic performance: Evidence from two longitudinal university samples. *Journal of Research in Personality, 37*, 319-338.
- Chamorro-Premuzic, T., & Furnham, A. (2003b). Personality traits and academic examination performance. *European Journal of Personality, 17*, 237-250.
- Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. Mahwah, NJ: Lawrence Erlbaum.
- Coffield, F., Moseley, D., Hall, E., & Ecclestone, K. (2004). *Learning Styles and Pedagogy in Post-16 Learning a Systematic and Critical Review*. Trowbridge, Wiltshire: Learning and Skills Research Centre.
- Cohen, J. (1975). *Applied multiple regression/correlation analysis for the behavioral sciences*. New Jersey: Lawrence Earlbaum Associates.
- Cohen, P.A. (1984). College grades and adult achievement: a research synthesis. *Research in Higher Education, 20*, 281-294.
- Collins, J. M., Schmitt, F. L., Sanchez-Ku, M., Thomas, L., McDaniel, M. A., & Le, H. (2003). Can basic individual differences shed light on the construct meaning of assessment center evaluations? *International Journal of Selection and Assessment, 11*, 17-29.
- Conard, M.A. (2006). Aptitude is not enough: How personality and behavior predict academic performance. *Journal of Research in Personality, 40*, 339-346.
- Cook, M. (2004). *Personnel Selection: adding value through people*. Sussex: Wiley.
- Costa, P., & McCrae, R. (1992). *Revised NEO Personality Inventory (NEO-PI-R) and NEO Five-factor Inventory (NEO-FFI): Professional Manual*. Odessa: Psychological Assessment Resources Inc..
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika, 16*, 297-334.
- Davis, M. H., Friedman, Bd. M., Harden, R. M., Howie, P., Ker, J., McGhee, C., Pippard, M. J., & Snadden, D. (2001). Portfolio assessment in medical students' final examinations. *Medical Teacher, 23*, 357-366.
- De Groot, A. D. (1978). *Vijven en Zessen* [Five and Six]. Groningen: Wolters-Noordhoff.
- De Raad, B., & Schouwenburg, H. C. (1996). Personality in learning and education: a review. *European Journal of Personality, 10*, 303-336.
- De Vries, A., De Vries, R. E., & Born, M. Ph. (2011). Broad versus narrow traits: Conscientiousness and honesty-humility as predictors of academic criteria. *European Journal of Personality, 25*, 336-348.
- Deary, I. J., Whiteman, M. C., Starr, J. M., Whalley, L. J., & Fox, H. C. (2004). The impact of childhood intelligence on later life: Following up the Scottish Mental Surveys of 1932 and 1947. *Journal of Personality and Social Psychology, 86*, 130-147.

- Descy, P., & Tessaring, M. (2001). *Training and Learning for Competence. Second report on vocational training research in Europe : executive summary*. Luxembourg : Office for Official Publications of the European Communities.
- Di Fabio, A., & Busoni, L. (2007). Fluid intelligence, personality traits and scholastic success: Empirical evidence in a sample of Italian high school students. *Personality and Individual Differences*, 43, 2095-2104.
- Diaz, R. J., Glass, C. R., Arnkoff, D. B., & Tanofsky-Kraff, M. (2001). Cognition, anxiety, and prediction of performance in 1st-Year law students. *Journal of Educational Psychology*, 93, 420-429.
- Dierick, S., Van de Watering, G., & Muijtjens, A. (2002). De actuele kwaliteit van assessment: ontwikkelingen in de edumetrie [The quality of assessment]. In F. Dochy, L. Heylen, & H. van de Mosselaer (Eds.), *Assessment in onderwijs*. (p.91-122). Utrecht: Lemma.
- Dochy, F. J. R. C., & De Rijke, T. R. (Eds.) (1995). *Assessment Centers: Nieuwe toepassingen in opleiding, onderwijs en HRM* [Assessment Centers: New applications in training, education and HRM. Utrecht: Lemma.
- Dochy, F. J. R. C, Segers, M., & De Rijdt, C. (2002). Nieuwe ontwikkelingen: de assessment cultuur [New developments: the assessment culture]. In F. Dochy, L. Heylen, & H. Van de Mosselaer (Eds.), *Assessment in onderwijs*. Utrecht: Lemma.
- Donofrio, H. H. & Davis, K. (1997). *Oral communication across discipline: Adding value to academic pursuit and marketability*. Paper presented at the Annual Meeting of the Southern States Communication Association, Savannah, GA.
- Drenth, P. J. D. (2004). Selectie aan de poort van het hoger onderwijs [selection at entry of higher education]. *Thema*, 4, 48-51.
- Duff, A. (2000). Learning styles of UK higher education students. Four studies of the reliability and replicability of the Learning Styles Questionnaire (LSQ). *Bristol Business School Teaching and Research Review*, 3. Retrieved August 25, 2007, from [http://www.uwe.ac.uk/bbs/trr/Issue3/ls3-1\\_4.htm](http://www.uwe.ac.uk/bbs/trr/Issue3/ls3-1_4.htm)
- Duff, A., & Duffy, T. (2002). Psychometric properties of Honey & Mumford Learning Styles Questionnaire. *Personality and Individual Differences*, 33, 147-163.
- Ebert, J. (2009). *Projectmanagement Assessment Center*. Zwitserland: Universität Zürich.
- Eggen, L. (2011). *The student X-factor, social and psychological determinants of students' attainment in higher education*. Doctoral dissertation. Groningen: Rijksuniversiteit Groningen.
- Entwistle, N. (2000, November). *Promoting deep learning through teaching and assessment: conceptual frameworks and educational contexts*. Paper presented at the TLRP Conference, Leicester.
- Evans, M. (2000). *Planning for the transition to tertiary study: A literature review*. Retrieved Sept 23, 2008, from <http://www.monash.edu.au/transition/research/evans.html>

- Evers, A., Van Vliet-Mulder, J. C., & Groot, C. (2000). *Documentatie van test en testresearch in Nederland 2000* [Documentation of tests and research on tests in The Netherlands 2000]. Assen: Van Groenou.
- Farsides, T., & Woodfield, R. (2003). Individual differences and undergraduate academic success: the roles of personality, intelligence, and application. *Personality and Individual Differences, 34*, 1225-1243.
- Ferris, G. R., Witt, L. A., & Hochwarter, W. A. (2001). Interaction of social skill and general mental ability on job performance and salary. *Journal of Applied Psychology, 86*, 1075-1082.
- Ford, N., & Y Chen, S. (2001). Matching/mismatching revisited: an empirical study of learning and learning styles. *British Journal of Educational Technology, 32*, 5-22.
- Friedlander, P. (1996). Competency-driven, component-based curriculum architecture. *Performance & Instruction, 35*, 14-21.
- Furnham, A., & Chamorro-Premuzic, T. (2004). Personality and intelligence as predictors of statistics examination grades. *Personality and Individual Differences, 37*, 1013-1022.
- Furnham, A., Chamorro-Premuzic, T., & McDougall, F. (2003). Personality, cognitive ability, and beliefs about intelligence as predictors of academic performance. *Learning and Individual Differences, 14*, 47-64.
- Furnham, A., & Medhurst, S. (1995). Personality correlates of academic seminar behavior: A study of four instruments. *Personality and Individual Differences, 19*, 197-208.
- Gaugler, B. B., Rosenthal, D. B., Thornton, G. C., III, & Bentson, C. (1987). Meta-analysis of assessment center validity. *Journal of Applied Psychology, 72*, 493-511.
- Geen, R.G. (1985). Test anxiety and visual vigilance. *Journal of Personality and Social Psychology, 49*, 963-970.
- Gilles, P. Y., & Bailleux, C. (2001). Personality traits and abilities as predictors of academic achievement. *European Journal of Psychology of Education, 16*, 3-15.
- Gniatczyk, L. A., & Ladd, R. T. (2001, April). *Assessment centre validity in a downsizing organisation*. Paper presented at the Conference of the Society for Industrial and Organisational Psychology. San Diego, CA.
- Goldberg, M. D., & Cornell, D. G. (1998). The influence of intrinsic motivation and self-concept on academic achievement in second- and third-grade students. *Journal of Education of the Gifted, 21*, 179-205.
- Goncz, A. (1994). Competency based assessment in the professions in Australia. *Assessment in Education, 1*, 27-44.
- Gottfredson, L. S. (2002). Where and why g matters: Not a mystery. *Human Performance, 15*, 25-46.
- Gottfried, A. E. (1990). Academic Intrinsic Motivation in Young Elementary School Children. *Journal of Educational Psychology, 82*, 525-538.

- Grant, G., Elbow, P., Ewens, T., Gamson, Z., Kohli, W., Neumann, W., Olesen, V., & Riesman, D. (1979). *On Competence. A critical analysis of competence-based reforms in higher education*. San Francisco : Jossey-Bass.
- Guion, R.M. (1998). *Assessment, Measurement, and Prediction for Personnel Decisions*. New Jersey: Lawrence Erlbaum.
- Harris, D. (1940). Factors affecting college grades: a review of the literature, 1930-1937. *Psychological Bulletin*, 37, 125-166.
- HBO-raad. (2008). *Fact sheet*. Den Haag: HBO-raad.
- HBO-raad. (2009). *Fact sheet*. Den Haag: HBO-raad.
- HBO-raad. (2010). *Feiten en cijfers* [Facts and figures]. Den Haag: HBO-raad.
- Hembree, R. (1988). Correlates, causes, and treatment of test anxiety. *Review of Educational Research*, 58, 47-77.
- Hendriks, P., & Schoonman, W. (2006). *Handboek Assessment deel I, gedragsproeven* [Handbook Assessment part I]. Assen: Van Gorcum.
- Hermans, H. J. M. (1967). *Motivatie en prestatie* [Motivation and Achievement]. Doctoral dissertation. Amsterdam: Swets & Zeitlinger.
- Hermans, H. J. M. (1976). *Handleiding bij de Prestatie Motivatie Test (PMT)* [Manual Achievement Motivation Test]. Amsterdam: Swets & Zeitlinger.
- Hoekstra, H. A., Ormel, J., & de Fruyt, F. (1996). *Handleiding NEO persoonlijkheids-vragenlijsten NEO-PI-R en NEO-FFI* [Manual NEO-PI-R en NEO-FFI]. Lisse, Swets Test Services.
- Honey, P., & Mumford, A. (1986, 1992). *The Manual of Learning Styles*. Berkshire: Honey, Ardingly House.
- Honey, P., & Mumford, A. (2000). *The learning styles helper's guide*. Maidenhead: Peter Honey Publications Ltd.
- Hough, L. M. (1992). The "Big Five" personality variables-construct confusion: Description versus prediction. *Human Performance*, 5, 139-155.
- Huitt, W. (1994). A transactional model of the teaching/learning process. Retrieved June 7, 2007, from <http://chiron.valdosta.edu/whuitt/materials/tchlrmnd.html>.
- Humphreys, L. G. (1968). The fleeting nature of college academic success. *Journal of Educational Psychology*, 59, 375-380.
- Hunter, J. E., & Hunter, R. F. (1984). Validity and utility of alternate predictors of job performance. *Psychological Bulletin*, 96, 72-98.
- Hurtz, G. M., & Donovan, J. J. (2000). Personality and job performance: The Big Five revisited. *Journal of Applied Psychology*, 85, 869-879.
- IBG, (Informatie Beheer Groep)(2008). *Alles over studiefinanciering voor studenten en afgestudeerden*. [Everything about study funding for students and graduates]. Groningen: Informatie Beheer Groep.
- Jackson, C. J., & Lawty-Jones, M. (1996). Explaining the overlap between personality and learning style. *Personality and Individual Differences*, 20, 293-300.

- Jansen, P. G. W., & Stoop, L. A. M. (2001). The dynamics of assessment center validity: Results of a seven year study. *Journal of Applied Psychology, 86*, 741-753.
- Jansen, P. G. W., & Vinkenburger, C. J. (2006). Predicting management career success from assessment center data: A longitudinal study. *Journal of Vocational Behavior, 68*, 253-266.
- Jenewein, K., Knauth, P., & Zülch, G. (2002). *Kompetenzentwicklung in Unternehmensprozessen* [Competence development and processes]. Aachen: Shaker.
- Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.
- Jöreskog, K. G., & Sörbom, D. (1988). *LISREL 7: A Guide to the Program and Applications*. Chicago: SPSS Inc..
- Judge, T. A., Cable, D. M., Boudreau, J. W., & Bretz, R. D. (1995). An empirical investigation of predictors of executive career success. *Personnel Psychology, 3*, 485-519.
- Kanfer, R., & Ackerman, P. L. (1989). Motivation and cognitive abilities: An integrative/aptitude-treatment interaction approach to skill acquisition. *Journal of Applied Psychology, 74*, 657-690.
- Kappe, F. R. (2009, May). *Quantitative versus Qualitative criteria for assessing student success*. Poster session presented at the annual ORD dagen [Educational Research Days], Leuven, Belgium.
- Kappe, F. R., Boekholt, L., Rooyen, den C., & Van der Flier, H. (2009). A predictive validity study of the Learning Style Questionnaire (LSQ) using multiple, specific learning criteria. *Learning and Individual Differences, 19*, 464-467.
- Kappe, F. R., & Pluijter, M. M. (2009, May). *Student characteristics related to first year withdrawal and student success across four disciplines*. Paper presented at the European First Year Experience (EFYE) conference "Enhancing the first-year experience: Theory, research, practice", Groningen, the Netherlands.
- Kappe, F. R., & Van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic achievement. *Journal of Research in Personality, 44*, 142-145.
- Kappe, F. R., & Van der Flier, H. (2011). Predicting academic achievement in higher education: What's more important than being smart? *Manuscript submitted for publication*.
- Kayzel, R. (2005). De belofte van het nieuwe leren [The promise of new learning]. *Tijdschrift Onderzoek van Onderwijs, 35*, 5-9.
- Kenny, D. A., & Kashy, D. A. (1992). Analysis of multitrait-multimethod matrix by confirmatory factor analysis. *Psychological Bulletin, 112*, 165-172.
- Kleinmann, M., & Köller, O. (1997). Construct validity of assessment centers: Appropriate use of confirmatory factor analysis and suitable construction principles. In R.E. Riggio & B.T. Mayes, (Eds.), Perspectives on assessment centers [Special issue]. *Journal of Social Behavior and Personality, 12*, 65-84.



- Klimoski, R. J., & Brickner, M. (1987). Why do assessment centers work? The puzzle of assessment center validity. *Personnel Psychology*, 40, 243-260.
- Kolb, D. A. (1984). *Experiential Learning: Experience as the source of learning and development*. New Jersey: Prentice-Hall.
- Kolk, N. J. (2001). *Assessment Centers: Understanding and improving construct-related validity*. Doctoral dissertation. Enschede: Kurt Lewin Instituut Dissertatiereeks.
- Komarraju, M., Karau, S. J., Schmeck, R. R., & Avdic, A. (2011). The Big Five personality traits, learning styles, and academic achievement. *Personality and Individual Differences*, 51, 472-477.
- Korthagen, F. (2004). Zin en onzin van competentiegericht leren [Issues concerning competence based education]. *VELON Tijdschrift voor Lerarenopleiders*, 25, 13-23.
- Kottke, J. L., & Shultz, K. S. (1997). Using an assessment center as a developmental tool for graduate students: A demonstration. In R.E. Riggio & B.T. Mayes, (Eds.), Perspectives on assessment centers [Special issue]. *Journal of Social Behavior and Personality*, 12, 289-302.
- Kuncel, N. R., Crede, M., & Thomas, L. L. (2004, August). *The validity of the GMAT: A meta-analysis*. Paper presented at the annual conference of the Academy of Management, New Orleans.
- 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.
- Kuncel, N. R., Hezlett, S. A., & Ones, D. S. (2004). Academic performance, career potential, creativity, and job performance: Can one construct predict them all? *Journal of Personality and Social Psychology*, 86, 148-161.
- Linn, R. L., & Hastings, C. N. (1984). A meta-analysis of the validity of predictors of performance in law school. *Journal of Educational Measurement*, 21, 245-259.
- Linn, Y. G., McKeachie, W. J., & Kim, Y. C. (2003). College student intrinsic and/or extrinsic motivation and learning. *Learning and Individual Differences*, 13, 251-258.
- Mackintosh, N. J. (1998). *IQ and human intelligence*. Oxford, UK: Oxford University Press.
- Mayes, B. T. (1997). Insights into the history and future of assessment centers: An interview with Dr. Douglas W. Bray and Dr. William Byham. In R.E. Riggio & B.T. Mayes, (Eds.), Perspectives on assessment centers [Special issue]. *Journal of Social Behavior and Personality*, 12, 3-12.
- McCrae, R., & Costa, P. (1994). Conceptions and correlates of openness to experience. Openness to Experience: Expanding the boundaries of Factor V. *European Journal of Personality*, 8, 251-272.
- McMullan, M., Endacott, R., Gray, M. A., Jasper, M., Miller, C. M. L., Scholes, J., & Webb, C. (2003). Portfolios and assessment of competence: a review of the literature. *Journal of Advanced Nursing* 41, 283-294.



- Meriac, J. P., Hoffman, B. J., Woehr, D. J., & Fleisher, M. S. (2008). Further evidence for the validity of assessment center dimensions: A meta-analysis of the incremental criterion-related validity of dimension ratings. *Journal of Applied Psychology*, 93, 1042-1052.
- Messick, S. (1994). The interplay of evidence and consequences in the validation of performance assessments. *Educational Researcher*, 23, 13-23.
- Messick, S. (1995). Validity of psychological assessment. Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. *American Psychologist*, 50, 741-749.
- Mitchell, J. V. (1992). Interrelationships and predictive efficacy for indices of intrinsic, extrinsic, and self-assessed motivation for learning. *Journal of Research and Development in Education*, 25, 149-155.
- Mount, M. K., & Barrick, M. R. (1995). The Big Five personality dimensions: Implications for research and practice in human resources management. In K. M. Rowland & G. Ferris (Eds.), *Research in personnel and human resources management* (p. 153-200). Greenwich, CT: JAI Press.
- Mount, M. K., Barrick, M. R., & Stewart, G. L. (1998). Five-factor model of personality and performance in jobs involving interpersonal interactions. *Human Performance*, 11, 145-165.
- Mouw, J. T., & Khanna, R. K. (1993). Prediction of academic success: A review of the literature and some recommendations. *College Student Journal*, 27, 328-336.
- Mulder, M. (2003). Ontwikkelingen in het competentiedenken en competentiegericht beroepsonderwijs. In M. Mulder, Wesselink, R., Biemans, H., Nieuwenhuis, L., & Poell, R. (Eds.), *Competentiegericht beroepsonderwijs. Gediplomeerd, maar ook bekwaam?* [Competence based education] Houten: Wolters-Noordhoff.
- Mulder, M. (2004). *Education, competence and performance: on training and development in the agri-food complex*. Inaugural address. Wageningen: Wageningen University.
- Neisser, U., Boodoo, G., Bouchard, T., Boykin, A., Brody, N., Ceci, S., Halpern, D., Loehlin, J., Perloff, R., Sternberg, R., & Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist*, 51, 77-101.
- Ng, T. W. H., Eby, L. T., Sorensen, K. L., & Feldman, D. C. (2005). Predictors of objective and subjective career success: A meta-analysis. *Personnel Psychology*, 58, 367-408.
- Noftle, E., & Robins, R. (2007). Personality predictors of academic outcomes: Big Five correlates of GPA and SAT scores. *Journal of Personality and Social Psychology*, 93, 116-130.
- Nunnally, J. C. (1978). *Psychometric theory*. New York: McGraw-Hill.
- O'Conner, M. C., & Paunonen, S. V. (2007). Big Five personality predictors of post-secondary academic performance. *Personality and Individual Differences*, 43, 971-990.

- OECD. (2010). *Education at a glance 2010: OECD indicators* (Organisation for Economic Cooperation and Development). Paris: OECD Publishing.
- Onderwijsraad, (2008). *Een succesvolle start in het hoger onderwijs* [A successful start in Higher Education]. Den Haag: Onderwijsraad.
- Parry, S. B. (1998). Just what is a competency? *Training*, 35, 58-64.
- Pascarella, E. T., & Terenzini, P. T. (1991). *How college affects students: Findings from twenty years of research*. San Francisco: Jossey-Bass.
- Paus, T. (2005). Mapping brain maturation and cognitive development during adolescence. *Trends in Cognitive Sciences*, 9, 60-68.
- Petrides, K. V., Chamorro-Premuzic, T., Frederickson, N., & Furnham, A. (2005). Explaining individual differences in scholastic behavior and achievement. *British Journal of Educational Psychology*, 17, 239-255.
- Pfeffer, J., & Fong, C. T. (2002). The end of business schools? Less success than meets the eye. *Academy of Management Learning and Education*, 1, 78-95.
- Pheiffer, G., Holley, D., & Andrew, D. (2005). Developing thoughtful students: using learning styles in an HE context. *Education & Training*, 47, 422-431.
- Poropat, A. E. (2009). A meta-analysis of the Five-Factor model of personality and academic performance. *Psychological Bulletin*, 2, 322-338.
- Price, L., & Richardson, J. T. E. (2003). Meeting the challenge of diversity: a cautionary tale about learning styles. In C. Rust (Eds). *Improving student learning theory and practice 10 years on*, Oxford: Oxford Centre for Staff and Learning Development, Oxford Brookes University.
- Ramist, L. (1984). Predictive validity of the ATP tests. In T. F. Donlon (Eds.). *The College Board technical handbook for the Scholastic Aptitude Test and Achievement Tests* (p. 141-170), New York: College Entrance Examination Board.
- Reilly, R. R., & Warech, M. A. (1993). The validity and fairness of alternatives to cognitive tests. In L.C. Wing & B. R. Gifford (Eds.), *Policy issues in employment testing* (p. 131-224). Boston: Kluwer.
- Resing, W. C. M., & Drenth, P. J. D. (2009). *Intelligentie: weten en meten* [Intelligence] (2nd edition). Amsterdam: Nieuwezijds.
- Riding, R. J. (2005). Individual differences and educational performance. *Educational Psychology*, 25, 659-672.
- Riggio, R. E., Aguirra, M., Nayas, B. T. Belloli, C., & Kubiak, C. (1997). The use of assessment center methods for student outcome assessment. *Journal of Social Behavior and Personality*, 12, 273-288.
- Riggio, R. E., Mayes, B. T., & Schleicher, D. J. (2003). Using assessment center methods for measuring undergraduate business student outcomes. *Journal of Management Inquiry*, 12, 68-78.

- Robbins, S. B., Lauver, K., Le, H., Davis, D., Langley, R., & Carlstrom, A. (2004). Do psychosocial and study skill factors predict college outcomes? A meta-analysis. *Psychological Bulletin*, 130, 271-288.
- Rolfhus, E., & Ackerman, P. L. (1999). Assessing individual differences in knowledge: knowledge, intelligence, and related traits. *Journal of Educational Psychology*, 91, 511-526.
- Roth, P. L., Be Vier, C. A., Switzer, F. S., & Schippman, J. S. (1996). Meta-analyzing the relationship between grades and job performance. *Journal of Applied Psychology*, 81, 548-556.
- Roth, P. L., & Clarke, R. L. (1998). Meta-analyzing the relation between grades and salary. *Journal of Vocational Behavior*, 53, 386-400.
- Rothstein, M., Paunonen, S., Rush, J., & King, G. (1994). Personality and cognitive ability predictors of performance in graduate business school. *Journal of Educational Psychology*, 86, 516-530.
- Ruble, T. L., & Stout, D. E. (1993). Comments on the use of the LSI in research on student performance in accounting education. *Accounting Educators' Journal*, 5, 35-45.
- Rupp, D. E., Thornton, G. C., III, & Gibbons, A. M. (2008). The construct validity of the assessment center method and usefulness of dimensions as focal constructs. *Industrial and organizational Psychology: Perspectives on Science and Practice*, 1, 116-120.
- Sackett, P. R., & Dreher, G. F. (1982). Constructs and assessment center dimensions: Some troubling empirical findings. *Journal of Applied Psychology*, 67, 401-410.
- Sackett, P. R., Gruys, M. L., & Ellingson, J. E. (1998). Ability-Personality Interactions When Predicting Job Performance. *Journal of Applied Psychology*, 80, 545-556.
- Sackett, P. R., & Harris, M. M. (1988). A further examination of the constructs underlying assessment center ratings. *Journal of Business and Psychology*, 3, 214-229.
- Salgado, J. F. (1997). The five factor model of personality and job performance in the European Community. *Journal of Applied Psychology*, 82, 30-43.
- Salgado, J. F., & Anderson, N. (2001). Cognitive and GMA testing in the European Community: issues and evidence. *Human Performance*, 15, 75-96.
- Salgado, J. F., Anderson, N., Moscoso, S., Bertua, C., De Fruyt, F., & Rolland, J. P. (2003). Meta-analytic study of GMA validity for different occupations in the European Community. *Journal of Applied Psychology*, 88, 1068-1081.
- Samson, G. E., Graue, M. E., Weinstein, T., & Walberg, H. J. (1984). Academic and occupational performance: A quantitative synthesis. *American Education Research Journal*, 21, 311-321.
- Schleicher, D. J., Day, D. V., Mayes, B., & Riggio, R. E. (1999, April). *A new frame for frame of reference training: Enhancing the construct validity of assessment centers*. Paper presented at meeting of the Society for Industrial and Organizational Psychology, Atlanta, GA.

- Schmidt, F. L. (2002). The role of general cognitive ability and job performance: Why there cannot be a debate. *Human Performance*, 15, 187-210.
- 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.
- Schmitt, N. (1977). Interrater agreement in dimensionality and combination of assessment center judgments. *Journal of Applied Psychology*, 63, 171-176.
- Schmitt, N., Gooding, R., Noe, R., & Kirsch, M. (1984). Meta-analysis of validity studies published between 1964 and 1982 and the investigation of study characteristics. *Personnel Psychology*, 37, 407-422.
- Schmitt, N., Oswald, F. L., Kim, B. H., Imus, A., Drzakowski, S., Friede, A., & Shivpuri, S. (2007). The use of background and ability profiles to predict college student outcomes. *Journal of Applied Psychology*, 92, 165-179.
- Segers, M., Dochy, F. J. R. C., & Cascallar, E. (Eds.). (2003). *Optimising new modes of assessment: In search of qualities and standards*. Boston/Dordrecht: Kluwer Academic Publishers.
- Sijsma, K. (2010). Over misverstanden rond Cronbach's alfa en de wenselijkheid van alternatieven [Misunderstandings about Cronbach's alpha]. *De Psycholoog*, 11, 561-567.
- Sims, R. R., Veres, J. G., & Shake, L. G. (1989). An exploratory examination of the convergence between the Learning Styles Questionnaire and the Learning Style Inventory II. *Educational and Psychological Measurement*, 49, 227-233.
- Skaalvik, E. M., & Rankin, R. J. (1994). Gender differences in mathematics and verbal achievement, self-perception and motivation. *British Journal of Educational Psychology*, 64, 419-428.
- Smith, W., Sekar, S., & Townsend, K. (2002). The impact of surface and reflective teaching and learning on student academic success. In M. Valcke, & D. Gombeir (Eds.), *Learning styles: Reliability and validity. Proceedings of the 7th annual learning style conference, Belgium*, 407-418.
- Steiger, J. H. (1980). Tests for Comparing Elements of a Correlation Matrix. *Psychological Bulletin*, 87, 245-251.
- Stellwagen, J. B. (2001). A challenge to the learning style advocates. *The Clearing House*, 75, 265-268.
- Sternberg, R. J., & Wagner, R. K. (1993). The g-centric view of intelligence and job performance is wrong. *Current Directions in Psychological Science*, 2, 1-5.
- Swales, S., & Senior, B. (1999). The dimensionality of Honey and Mumford's Learning Style Questionnaire. *International Journal of Selection and Assessment*, 7, 1-11.
- Tett, R. P., Jackson, D. N., & Rothstein, M. (1991). Personality measures as predictors of job performance: A meta-analytic review. *Personnel Psychology*, 44, 703-742.

- Tett, R. P., Jackson, D. N., Rothstein, M., & Reddon, J. R. (1999). Meta-analysis of bidirectional relations in personality-job performance research. *Human Performance*, 12, 1-29.
- Thomas, S. L. (2000). Deferred costs and economic returns to college major, quality, and performance. *Research in Higher Education*, 41, 281-313.
- Thornton, G. C. III, & Gibbons, A. M. (2009). Validity of assessment centers for personnel selection. *Human Resource Management Review*, 19, 169-187.
- Tinto, V. (1987). *Leaving College: rethinking causes and cures of student attrition*. Chicago: The university of Chicago.
- Trown, A. E., & Leith, G. O. (1975). Decision rules for teaching strategies in primary schools: Personality-treatment interactions. *British Journal of Educational Psychology*, 45, 130-140.
- Turnage, J. J., & Muchinsky, P. M. (1982) Transituational variability in human performance with assessment centers. *Organizational Behavior and Human Performance*, 30, 174-200.
- U.S. Department of Education, National Center for Educational Statistics. (2002). *Defining and assessing learning: exploring competency-based initiatives* (NCES 2002-159). Washington, DC.: U.S. Department of Education.
- Vallerand, R. J., Pelletier, L. G., Blais, M. R., Briere, N. M., Senecal, C., & Vallieres, E. F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: evidence on the concurrent and construct validity of the academic motivation scale. *Educational and Psychological Measurement*, 53, 159-172.
- Van Bragt, C. A. C. (2010). *Students' educational careers in higher education: a search into key factors regarding study outcome*. Doctoral dissertation. Eindhoven: Fontys Hogescholen.
- Van de Griend, P. C. (1964). Mogelijkheden en grenzen van een brugjaar [Possibilities and limits of a bridge year]. Amsterdam: Universiteit van Amsterdam.
- Van den Brink, W. P., & Mellenbergh, G. J. (2006). Testleer en testconstructie [Test knowledge and test construction]. Amsterdam: Boom.
- Van der Sanden, J. M. M., De Bruijn, E., & Mulder, R. H. (2003). Ontwikkelingen in het onderzoek op het terrein van het beroepsonderwijs. In M. Mulder, R. Wesselink, H. Biemans, L. Nieuwenhuis, & R. Poell (Eds.). *Competentiegericht beroepsonderwijs. Gediplomeerd, maar ook bekwaam?* [Competence based education]. Houten: Wolters-Noordhoff.
- Van der Vleuten, C. P. M., & Driessen, E. W. (2000). *Toetsen in probleem gestuurd onderwijs* [Exams in problem based education]. Groningen: Wolters-Noordhoff.
- Van Merriënboer, J. J. G., Van der Klink, M. R., & Hendriks, M. (2002). *Competenties: van complicaties tot compromis. Over schuifjes en begrenzers* [Competencies: from complications to compromise]. Den Haag: Onderwijsraad.

- Veres, J. G., Sims, R. R., & Locklear, T. S. (1991). Improving the reliability of Kolb's revised LSI. *Educational and Psychological Measurement*, 51, 143-150.
- Vermeulen-Kerstens, L. (2006). *The study of alumni*. Doctoral dissertation. Lemma: Utrecht.
- Vermunt, J. D. (1992). *Learning styles and regulation of learning processes in higher education: Towards process-oriented instruction in autonomous thinking*. Amsterdam: Swets & Zeitlinger.
- Vickers, J. M. (2000). Justice and truth in grades and their averages. *Research in Higher Education*, 41, 141-164.
- VSNU. (2010). *Feiten en cijfers* [Facts and figures]. The Hague: VSNU.
- Waldman, D. A., & Korbar, T. (2004). Student assessment center performance in the prediction of early career success. *Academy of Management Learning and Education*, 3, 151-167.
- Wiersma, U. J. (1992). The effects of extrinsic rewards in intrinsic motivation: A meta-analysis. *Journal of Occupational and Organizational Psychology*, 65, 101-114.
- Wolting, D. (2006). *Studentenselectie: zijn de selectie instrumenten goede voorspellers van success* [Student selection: Are selection instruments good predictors of success]. Unpublished master's thesis, Free University, Amsterdam, The Netherlands.
- Zeidner, M. (1998). *Test anxiety: The state of the art*. New York: Plenum Press.
- Zwanenberg, V. N., Wilkinson, L. J., & Anderson, A. (2000). Felder and Silverman's Index of Learning Styles and Honey and Mumford's Learning Styles Questionnaire: How do they compare and do they predict academic achievement? *Educational Psychology*, 20, 365-380.
- Zyphur, M. J., Bradley, J. C., Landis, R. S., & Thoresen, C. J. (2008). The effects of cognitive ability and conscientiousness on performance over time: A censored latent growth model. *Human Performance*, 21, 1-27.

# Dankwoord

Met succes een opleiding zoals een PhD afronden is zondermeer een kwestie van doorzettingsvermogen en hard werken, maar ook van geluk. Daarmee bedoel ik het geluk om die mensen om je heen te hebben die je inspireren, coachen, liefdevol toespreken, corrigeren, helpen, luisteren en laten lachen. Zo ben ik dankbaar voor de vele ontmoetingen, gesprekken, aanmoedigingen en support van collega's, vrienden, vreemden en familie.

In het bijzonder wil ik Henk van der Flier bedanken voor zijn support. Dank, dat je met emeritaat bent gegaan waardoor je alle tijd voor mij had, met uitzondering natuurlijk van de donderdagen. Ik vind het een eer dat ik onder jouw vleugels deze proeve van bekwaamheid heb mogen afleggen. Ook wil ik Jeroen Onstenk bedanken voor zijn begeleiding en opname in zijn kenniskring. Jouw vrije wijze van begeleiding gaf mij veel vertrouwen en de mogelijkheid om mijn eigen weg te vinden. Als begeleidingsteam vulden jullie elkaar mooi aan waardoor het proefschrift zowel een academisch als toegepast karakter heeft gekregen.

Dank aan alle studenten die, als onderwerp van onderzoek, alle vragenlijsten, zelfs jaren na afstuderen, hebben willen invullen. Ik hoop dat ik jullie, als HR professionals, met dit proefschrift van bruikbare adviezen heb kunnen voorzien. Ik denk dat jullie bijdrage gevolgen kan hebben voor hoe het HRM onderwijs er in de toekomst uit zal zien. Dank aan de Vrije Universiteit voor het bieden van gastvrijheid en aan alle collega's bij de afdeling Sociale en Organisationspsychologie die mij een zeer welkom gevoel hebben gegeven als ik er kwam werken. Peter Dekker in het bijzonder dank voor je hulp bij de statistische analyses.

Dank aan alle oud collega's van het NOA en LTP en alle huidige collega's binnen de hogeschool die onaflatend geïnteresseerd waren, vooral in of het proefschrift al af was. Dank aan mijn collega promovendi voor de nuttige sessies waarin we op onze ervaringen als buitenpromovendi konden reflecteren en aan Theo Douma die als toenmalig directeur van de dienst Onderwijs, Kwaliteit & Research zijn vertrouwen uitsprak in mijn promotievoorstel waardoor ik met mijn promotieonderzoek kon starten. Dank aan het bestuur van de Stichting LTP voor hun vertrouwen en financiële steun en in het bijzonder Arjan Eleveld. Uco Wiersma wil ik bedanken voor onze inspirerende en productieve "oversea's discussions".

Naast doorzetten en geluk is ongetwijfeld het thuisfront van cruciaal belang en wat is er veel gebeurd om dankbaar voor te zijn met als hoogtepunten de geboorten van Morris en Bente. Over hoe je iemand moet bedanken die er altijd voor je is, je aanmoedigt, je vertrouwen en liefde geeft, daar ga ik nu wat voor bedenken.

Het is een mooie tijd geweest en het is mooi dat het nu voorbij is.

Rutger

Velserbroek, 5 oktober 2011



## Curriculum Vitae

Rutger Kappe was born on the 14th of August, 1971, in Dordrecht, the Netherlands. He studied Industrial & Organizational Psychology at the Free University in Amsterdam. After graduation he worked at a Dutch research and development center, NOA, which specializes in developing culture free tests and assessments for businesses and education. He also worked as assessment psychologist and quality manager assessment at LTP a large, renowned HRM consultant firm in the Netherlands. He also worked part-time as advisor and lecturer at the Inholland college for professional education. Since 2004 he is fulltime employed at Inholland as an advisor, researcher, and trainer at the office of Education, Quality, Research & Development. In 2006 he started in part-time with his PhD research at the Free University. His main areas of research interest include peer-, intake-, formative- and summative assessment, quality issues in assessment, learning and instruction, competence based education and foremost student success.



# List of publications

## *Articles in peer-reviewed journals*

- Kappe, F. R., Boekholt, L., Den Rooyen, C., & Van der Flier, H. (2009). A predictive validity study of the Learning Style Questionnaire (LSQ) using multiple, specific learning criteria. *Learning and Individual Differences, 19*, 464-467.
- Kappe, F. R., & Van der Flier, H. (2010). Using multiple and specific criteria to assess the predictive validity of the Big Five personality factors on academic achievement. *Journal of Research in Personality, 44*, 142-145.
- Kappe, F. R., & Van der Flier, H. (2011). Predicting academic achievement in higher education: What's more important than being smart? *Conditionally accepted*.
- Kappe, F. R., & Van der Flier, H. (2011). The predictive validity of academic assessment centers for career success: Using uncontaminated criteria. *Manuscript submitted for publication*.

## *Other scientific articles, conference papers and posters*

- De Volder, M., & Kappe, F. R. (2007, January). *Espace: A New web-tool for peer assessment with in-built feedback quality system*. Paper presented at The Conference on teaching and learning, Hawaii, United States.
- De Volder, M., Rutjens, M., Slootmaker, A., Kurvers, H., Bitter, M., Kappe, F. R., Roossink, H., De Goeijen, J., & Reitzema, H. (2007). Espace: A student peer-feedback tool enhances the educational quality of assignments without overburdening faculty overburdening faculty or adding administrative costs. *Educational Quarterly, 3*, 63-66.
- Kappe, F. R. (2008, November). *Educational careers in Higher Education: Identifying student characteristics in student success*. Poster session presented at the 'day for research' of the Inholland University of applied sciences, Haarlem, The Netherlands.
- Kappe, F. R. (2008, November). *Student perceptions on formative and summative peer-assessment (center)*. Poster session presented at the EARLI PBPR Conference, Bergen, Norway.
- Kappe, F. R. (2009, May). *Quantitative versus Qualitative criteria for assessing student success*. Poster session presented at the annual Educational Research Days, Leuven, Belgium.
- Kappe, F. R. (2009, May). *The role of intelligence, personality, motivation and learning style in students' academic achievement in the first year of college*. Poster session presented at the European First Year Experience (EFYE) conference "Enhancing the first-year experience: Theory, research, practice", Groningen, the Netherlands.
- Kappe, F. R., Boekholt, L., & Den Rooyen, C. (2008, November). *Does matching and mismatching using learning styles have an effect on learning and learning outcomes?* Paper presented at the EARLI PBPR Conference, Bergen, Norway.

Kappe, F. R., & Pluijter, M. M. (2009, May). *Student characteristics related to first year withdrawal and student success across four disciplines*. Paper presented at the European First Year Experience (EFYE) conference “Enhancing the first-year experience: Theory, research, practice”, Groningen, the Netherlands.

*National (refereed) book chapters and journal articles*

Kappe, F. R. (2005). *Assessment bij de opleiding HRM [Assessment at the study HRM]*. In M. Reichert & P. Rottenveel. *Competentiegericht toetsen: Best Practices*. Utrecht: Digitale Universiteit.

Kappe, F. R. (2007). *Achtergronden van het competentie denken [History of competence based education]*. Backbone II. The Hague: Inholland.

Kappe, F. R. (2008). Hoe betrouwbaar is peer-assessment [The reliability of peer-assessment]. *Tijdschrift voor Hoger Onderwijs*, 26, 93-102.

Kappe, F. R. (2008). Voor- en nadelen van diverse typen assessoren [Pro's and con's of different types of assessors]. *Onderwijs voor Onderzoek*, 37, 23-26.

Kappe, F. R., & Pluijter, M. M. (2009). Intake-assessment: uitval verlagen en rendement verhogen [Intake: Decreasing drop-out and increasing graduation rates]. *Hoger Onderwijs actueel*, 1, 18-20.

Kappe, F. R., & ScholtenLinde, J. J. (2005). Het assessment center beoordeeld: face- en consequentiële validiteit [An assessment of the Assessment Center]. *Tijdschrift voor Hoger Onderwijs*, 23, 170-182.

*National conference presentations*

Jansen, F., Van Hees, J., & Kappe, F. R. (2009, February). *Studiekeuzegesprekken [Talking about choice of study]*. Presentation given at the yearly Higher Education conference, Utrecht, The Netherlands.

Kappe, F. R. (2007, May). *Monitoring competence development using assessment centers*. Workshop given at the yearly National Education Congress, Amsterdam, the Netherlands.

Kappe, F. R. (2008, April). *Predicting college success*. Workshop given at the yearly conference of the higher education counsel (HBO raad), Rotterdam, the Netherlands.

Kappe, F. R., & Pluijter, M. M. (2008, October). *Intake assessment*. Workshop given at the yearly conference of school deans NVS-NVL, Rotterdam, The Netherlands.

*Earlier publications (non-refereed)*

- Kappe, F. R. (1999). Dringen op de arbeidsmarkt: interventies voor organisaties [It's busy on the labor market: Interventions for businesses]. *Nederlandse Vereniging voor Personeelsbeleid*, 35(1), 20-23.
- Kappe, F. R. (1999). Geachte heer, mevrouw: Selectie op basis van brieven [Dear Sir, Madame: Selection based on letters]. *Nederlandse Vereniging voor Personeelsbeleid*, 35(3), 39-41.
- Kappe, F. R. (1999). Een goed gesprek [A good interview]. *Nederlandse Vereniging voor Personeelsbeleid*, 35 (9), 71-75.
- Kappe, F. R., & Van den Berg, R. H. (1997). Intake-assessments. *Loopbaan*, 2, 8-10.
- Kappe, F. R., & Van den Berg, R. H. (1998). *Succesvol omscholen in de ICT* [Research report on ICT careers]. Amsterdam: Vrije Universiteit/Stichting PION.