Hitting the high notes

Challenge in teaching honors students

Karin Scager

This research was supported by a grant of the Dutch Sirius Program: Excellence in higher education

ISBN: 978-90-393-6060-6

Printed by: Ipskamp Drukkers

Cover design: Gijs Milius

Hitting the high notes

Challenge in teaching honors students

De lat hoger leggen

Uitdaging in honors onderwijs

(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de Universiteit Utrecht op gezag van de rector magnificus, prof. dr. G. J. van der Zwaan, ingevolge het besluit van het college voor promoties in het openbaar te verdedigen op
vrijdag 6 december 2013 des morgens te 10.30 uur
door

Karin Scager

geboren op 27 december 1956 te Bilthoven

Promotoren: Prof. dr. A. Pilot
Prof. dr. T. Wubbels
Co-promotor: Dr. S.F. Akkerman

Contents

1. Introduction and thesis overview

1.1 General introduction

Over the last two decades, honors programs and colleges have mushroomed throughout the Netherlands and elsewhere. Honors programs are intended to challenge the most academically gifted students to develop their full potential. A special issue of the Journal of the National Collegiate Honors Council (2012) reports on the recent expansion of honors programs around the globe, from Latin America to China. Although the first honors program in the United States was already founded in the 1920s, Sederberg's (2008) survey of National Collegiate Honors Council members indicates that over 60% of the honors programs in the U.S. have been established since 1994. In the Netherlands, the first honors program started in the 1990s (Wolffensberger, 2012), followed by many others since.

During the 1960s and through the mid-1970s, higher education showed a period of unprecedented growth, often characterized by the term 'massification'. Accompanying this 'massification' of higher education was an increase in the diversity of students in terms of class, gender, ethnicity, and academic ability and motivation (Biggs, 1999; Coaldrake, 2001). The heterogeneity of students groups has led to creating additional facilities for students with difficulties coping with the demands, such as courses in academic learning, writing, and exam preparation. At class level, most teachers tailor pace and level of difficulty to the average student, which can lead to under-challenge of the group of high-ability students (Heller, 2007). In the last two decades however, the increase of special schools and programs for high-ability students in the Netherlands seems to reflect a subtle cultural change. Equal chances to enter higher education for those who have the ability are still important, but securing the minimum level is no longer the sole focus; creating the opportunity for the top 10% of the students to develop their talents has been added to the ideals (De Boer, Minnaert, & Kamphof, 2013).

Currently, different kinds of services are offered for high-ability students across all educational levels, such as 'schools with gifted profiles', enrichment classes, and pull-out classes (found in primary and secondary education), pre-university colleges, so-called junior colleges (partnership between a university and several secondary schools), and honors programs and colleges (in higher education). Originally, schools and institutes for higher education developed these services on their own initiative and based on their own understanding and knowledge, often derived from experiences in countries abroad such as the United States (De Boer et al., 2013). These services were initially formed to meet the needs of the small number of under-challenged students who experienced boredom in regular classes (Carnicom, 2011; Gallagher, Harradine, & Coleman, 1997), leading to a loss of motivation (Hoekman, McCormick, & Gross, 1999; Lens and Rand 2000), which in turn can lead to underachievement (e.g. Gentry, Gable, & Springer, 2000). The primary goal for creating programs for high-ability students thus was mostly an educational one, grounded in the belief

that high-ability students did not receive the education to which they were entitled, unless instructional strategies were modified (De Boer et al., 2013; Borland, 2003; Heller, 1999).

Only the last couple of years, the attention for education for high-ability students has entered the political agenda in the Netherlands. Domestic economic crisis and the rise of countries on other continents increased the vulnerability of the Netherlands and urged the development of high-ability students (de Boer et al., 2013), adding an economical objective for creating services for high-ability students. In action plans of the Department of Education (Van Bijsterveldt-Vliegenthart & Zijlstra, 2011) specific measures have been announced to focus on top students in primary, secondary, and higher education. In 2008, the Sirius program for the development of excellence in higher education was created by the Dutch Department of Education. Virtually all institutions applied for funding to initiate or extend their honors programs (Reumer & van der Wende, 2010). In the same year the university sector committed itself through an agreement with the Ministry to enroll 10% of students in excellence programs. Since then, growth of honors programs and colleges has expanded even more.

In the Netherlands as well as internationally, the need to create special programs for high-ability students has been subject of debate. Those against argue that creating these programs creates elitism, removes the best students and teachers from the general classroom, and reduces the status of regular programs, and that differentiation should be organized within the regular programs instead of pulling out the best students (Seifert et al., 2013). Moreover, high-ability students are believed to learn anyway and do not need any additional aid or guidance (De Boer et al., 2013; Wolfensberger, 2012). Beginning research on the effects of honors programs however shows that honors students gain significantly larger first-year growth on cognitive development measures than their non-honors peers (Seifert et al., 2013). Further, honors students value that in these programs, they do not feel the pressure to conform, and find peers with whom they can freely communicate their thoughts and insights (Hébert & McBee, 2007; Rimm, 2002; Robinson, 2003). To conclude, it is widely accepted that instructional strategies should match students‘ differentiated talent and learning needs (Gruber & Mandl, 2000; Heller, 2005; Heller, Mönks, & Passow, 1993; Snow & Swanson, 1986), although insights in how to organize this differ.

Whether or not creating separate honors programs is the most ideal form to differentiate, these programs offer a rich context to investigate how education can be specifically designed for high ability students. The programs are developed by inspired teachers who are allowed freedom to create innovative learning environments, resulting in a diversity of instructional strategies believed to match the needs of high-ability students. However, much is still to be understood about the specific design of education for these students. Given the growth of honors programs, research on how to teach high-ability students in higher education is sorely behind practice (Achterberg, 2005; Clark, 2000; Davis & Montgomery, 2011; Long & Lange, 2002; Rinn & Plucker, 2004). The vast majority of research on the effectiveness of learning environments for high-ability students has been conducted in primary and secondary

education. This research has mainly focused on the curriculum level, specifically on the effect of grouping, acceleration, and enrichment, and on combinations of the three. In this thesis, we focus on the learning environment at the course level in higher education, where high-ability students are already grouped.

Findings in literature on education for high-ability students consistently point to the importance of increasing the level of challenge for these students compared to mainstream students (Clinkenbeard, 1994; Freeman, 1990; Heller, 1999; Lens & Rand, 2000; Lubinski & Benbow, 2000; Marra & Palmer, 2004; Reis & Renzulli, 2010; Sayler, 2009; Snow & Swanson, 1992). It is however not obvious how to create appropriate challenge. The central question of this thesis therefore concerns how and to what extent challenge should be orchestrated for high-ability students at the university level in order to incite them to hit the highest notes.

The remaining part of this introduction focuses on the clarification of basic concepts that are used, respectively ´honors student´ and 'challenge´, in order to help the reader understand the thesis. Next, we discuss the central focus and context of this thesis, and provide an overview of the studies that were conducted.

What is an honors student?

Basically, the term 'honors student' is used to refer to 'gifted' or 'high-ability' students in higher education. The term 'gifted' was originally used chiefly to refer to children who exhibited unusual precocity or, after the development of intelligence tests, to children with high IQ's. Based on this definition, Louis Terman started his famous longitudinal study in 1922 on 1528 children with IQ scores of at least 140 before the age of 12. The research group followed the accomplishments of the group throughout their lifespan. At school, these children did on average very well but not excellent (Terman, 1965). Professionally, the 800 males in the sample developed impressive records in terms of publications, patents, and ratings by experts, more than 10 times as many as the control group. Intelligence thus predicted excellence in professional life. However, the 150 most and the 150 least successful men in the sample had similar IQ scores, but differed in personality traits such as motivation and persistence. The importance of characteristics other than intelligence, also found in subsequent studies (Trost, 2000), led to multi-component definitions of giftedness, such as Renzulli's (1978) well-known three ring model of giftedness that includes the components intelligence, creativity, and motivation.

In the discussion on giftedness, a distinction is made between potential and realized ability (Gagné, 1993; Gruber & Mandl, 2000). Giftedness is understood as a potential for outstanding achievement; the development of giftedness into excellence depends on environmental influences including education (Gruber & Mandl, 2000; Monks and Mason, 2000), and effort (Roe, 1953; Ericsson, 2006; Howe, 1999; Walberg et al., 2003). For our understanding of giftedness we adopted both Renzulli's (1978) three components (intelligence, creativity, and motivation), and Monks and Katzko's (2005) developmental understanding of giftedness. This combination led to the following working definition for this

thesis: giftedness includes a combination of relatively high intelligence, creativity, and motivation, and is an individual potential for exceptional achievements in one or more domains. In daily language, the term gifted is used more often for younger children, while 'high-ability' is more often used for older students. Throughout this thesis, we use the terms gifted and high-ability interchangeably. Additionally, we sometimes simply refer to these students as 'honors students', referring to the specific group of participants studied in this thesis.

The students who are selected for honors programs are assumed to have the potential to excel in their studies, but whether this will lead to excellence in their future professional lives is not evident yet. Universities mostly select honors students based on their grades in secondary education and their level of motivation. In his review of the relationship between grades attained at school or university and adult accomplishment, Hoyt (1966) concluded that grades have hardly any relationship with any measure of professional achievement. Eighteen years later, Cohen (1984) found similar results in his meta-analysis; the mean correlation for the 108 studies reported was r .18. While statistically significant, an effect of this magnitude is considered small. Baird (1985, and Taylor et al. (1985) came to the same conclusion, demonstrating that grades can predict future grades, but that they have low predictive value for professional excellence. Furthermore, not all of the students who are qualified and motivated apply for an honors program or college, and an unknown number of them end up in regular programs. This raises the question about whether and to what extent honors students differ from non-honors students with respect to the qualities that have been found to be essential for exceptional accomplishments in professional life.

What is challenge?

For our conceptualization of challenge, we draw from psychological theories that include the notion of challenge, such as the control-value theory of motivation (Pekrun, 2006), the self-determination theory of motivation (Deci & Ryan, 1985), and the flow theory (Csikszentmyhalyi, 1975). These theories concern the inherent relationship between levels of challenge and ability, indicating that studies of challenge also have to include the notion of ability. In addition, these theories stress the need for balance, with a challenge level matching the ability level for positive effects on motivation and learning. Csikszentmyhalyi (1975) assumes that an appropriate balance between challenge and ability creates a state of flow; accordingly, to maintain flow, tasks should increase in their level of difficulty as the student's level of ability improves. This postulate of optimal challenge in flow theory is consistent with the need for competence in the self-determination theory (Deci & Ryan, 2000): to create a feeling of competence, tasks need to be optimally challenging, implying a balance between challenge and abilities. In the control-value theory of motivation (Pekrun, 2006) it is posited that a lack of balance, caused by demands that are too high or too low, causes negative emotions that can impair intrinsic motivation.

These theories are in line with the findings in gifted literature that high-ability students need higher levels of challenge than mainstream students (Heller, 1999; Lubinski & Benbow, 2000; Snow & Swanson, 1992). High-ability students differ qualitatively from average students in being quicker thinkers, more flexible in their use of strategies, having better memories, and preferring complexity (Freeman, 1990; Shore & Kanevsky, 1993; Wallace, 2000). Given these abilities, challenging high-ability students may not simply be a matter of increasing the level, but may involve also adapting to these qualitative differences. Thus, although optimal balance between challenge and ability levels is needed for all students, it may be that that the way to challenge students is different for high-ability students.

1.2 The current thesis

Literature on education for high-ability students in the field of primary and secondary education recommend a compilation of methods, including promoting higher level thinking skills and creative problem-solving (Gallagher et al., 1997; Kanevsky & Keighly, 2003; Rogers, 2007; Shavinina, 2009; Sternberg & Lubart, 1993; Tassel-Baska & MacFarlane, 2009), and offering student-centered teaching methods such as inquiry learning, problem-based learning, or learning in authentic or contextualised settings (Gruber & Mandl, 2000; Heller, 1999; Hertberg-Davis & Callahan, 2008; Mönks & Katzko, 2005; Shore & Kanevsky, 1993; Snow & Swanson, 1992). As explained in the previous section, the concept of challenge seems to underpin many suggestions for gifted education, though only few of these studies rely on empirical research to support their claims regarding challenging learning environments. Hence, it is relevant to conceptualize and operationalize challenge in order to strengthen practice and theory more systematically in education for high-ability students at the university level. It is also necessary to consider the level of challenge in relation to the level of ability, since it is this relation that has been theorized as affecting motivation as well as learning. Further, we aimed in this thesis to understand the factors in the course environment that challenge high-ability students. The above mentioned suggestions are still quite wide-ranging. It is for example not clear how inquiry learning should be designed to challenge students to the full, and what the role of the teacher should be. This calls for more extensive research on challenge, including manifestations of challenging elements in the design and implementation of courses. Since the role of the teachers has been found to be crucial with respect to the effectiveness of teaching (Darling-Hammond, 1999; Hattie, 2007), it seemed important to include the experiences of teachers when employing challenge. In challenging courses, tensions may arise when some students feel that they cannot cope with the demands, and subsequently lose their motivation. Maximizing challenge after all is not the only responsibility teachers have; they also pursue non-academic objectives, such as caring for the well-being of individual students and developing a good learning climate (Butler, 2012; Oser, 1993). Therefore, not surprisingly, we found that teachers encountered dilemmas when providing high levels of challenge. Yet, which dilemmas occur, and what considerations are

involved in the choice whether or not to intensify challenge in the learning environment has not yet been examined.

Given the growth of honors programs in Europe and the scarcity of research into appropriate teaching methods (Wolffensberger, 2012), empirical research on challenge will be particularly informative for teachers of honors students. Insight in how to challenge students, what dilemmas can emerge when increasing the challenge and how experienced teachers reflect on these dilemmas can be informative for teachers of high-ability students in higher education.

Research questions

The main research aim, examining challenging instructional strategies for honors students, was based on the assumption that honors students differ significantly from their non-honors peers, and need instructional strategies that match their abilities. As explained above, honors students are selected on the basis of school grades and motivation, which does not ensure a significant difference with non-honors students with respect to their potential for professional excellence. Therefore, we first wanted to examine whether the students selected for honors programs indeed differ from non-honors students with respect to the three dimensions of high-ability: intelligence, motivation, and creativity. Our further research questions focused on both the appropriate level of challenge for these students and the factors that constituted the challenge. To complete the research, we explored the experience of teachers, seeking their reflection on evoking challenge in their courses. These aims led to three main research questions:

1. Do honors students differ from non-honors students with respect to the qualities that have been found to be essential for exceptional accomplishments in professional life (intelligence, creativity, and motivation)?
2. How can high-ability students in higher education be challenged?
3. What dilemmas do teachers experience when challenging high-ability students?

The first study explored differences between honors and non-honors students at Utrecht University, for which a quantitative approach by means of a questionnaire method seemed suitable. The search for challenging instructional strategies was based on the premise that good practice for honors education already exists. There are plenty of good and creative teachers, and when allowed the freedom to design their own courses, they provide an important data source for educational research. What the researcher needs to do is finding these examples, and, as once stated by Dewey (1929, p.11), use methods which enable us to make an analysis of what excellent teachers do intuitively, so that something accruing from their work can be communicated to others. Therefore, we chose for case studies, using mixed methods including student interviews, observation, and content analysis. For our last study, focusing on the experience of teachers on the creation of challenge, open ended interviews seemed the best method to use.

Context of this study

The studies were all carried out at Utrecht University. Founded in 1636, Utrecht University is one of the oldest and largest research universities in the Netherlands, with approximately 30.000 students. Utrecht University aims to encourage students to develop their talents to the best of their abilities, and in the last two decades additional honors programs and colleges[1] (about 15 to date) have been created for the most talented and motivated students. Utrecht University offers four honors undergraduate colleges: University College Utrecht, University College Roosevelt, College of Pharmaceutical Sciences, and Utrecht Law College. Further, all seven faculties offer selective honors programs. Most of these programs are taken in addition to the regular programs (about 15 EC per year), and offer activities that cover more in-depth study, interdisciplinary study, research, and/or applying theory into practice. In addition, social activities are organized to advance an honors community. The University College Utrecht (UCU) was chosen as setting for the largest part of the research. Founded in 1998, UCU is an international undergraduate honors college and offers a Liberal Arts and Sciences program (leading to a BA or BSc degree) through small-scale and intensive education. UCU has its own residential campus in the city of Utrecht, housing about 650 students from about 68 countries. UCU offers a selective academic program, and has been rated as the number one undergraduate degree in the Netherlands for several consecutive years. Among the special characteristics of the college are the international composition of its student body, the freedom for students to compose their own curriculum, its small classes and individual attention. The admissions process is selective. Students are admitted on the basis of grades and motivation, but an international focus, social engagement and extracurricular pursuits are of equal importance.

1.3 Overview of and relation between the studies

*Differences between honors and non-honors students (**chapter 2**)*

The objective of honors programs is to develop the abilities to help students on their way to excellence, and develop their abilities to the full. The selection process for these programs often relies on the students' prior achievements in school. Research has shown, however, that school grades do not sufficiently predict future achievement. According to Renzulli's (1986) three-ring model, the overlap and interaction between intelligence, motivation and creativity predicts excellent achievements in professional life. In this first study, we therefore firstly investigated whether honors students differ from non-honors students in terms of these three characteristics. Secondly, we explored which of these characteristics primarily differentiated between honors and non-honors students. For more than 1,100 honors and non-honors students at Utrecht University their intelligence, creativity, and motivation were measured using a self-report questionnaire. When it would appear that honors students differ significantly from non-honors students in terms of the characteristics which predict excellent

[1] An honors program is an additional enrichment program, while an honors college is an entire bachelor program for honors students.

professional achievement, it then would seem justifiable to provide special attention for the education of these students. After all, the effectiveness of education depends to a large extent on the fit between the learning environment and the abilities, interests and motivation of the students (McKeachie 1986; Snow 1986; Pascarella and Terenzini 1991).

*Challenging high-ability students (**chapter 3 and 4**)*

Challenging students to the full implies an appropriate balance between challenge and ability (Csikszentmyhalyi, 1975; Deci & Ryan, 2000). To study how students experienced the balance between challenge and ability, and the relation of this balance with students' perceived learning and motivation, we chose a course that was known to be quite extreme in challenging students to the full (chapter 3). A second aim was, to determine what factors constituted challenge. In this course, Advanced Cell Biology, the students developed a research program according to national scientific standards, which they did successfully, according to an external jury of experts in the field. Methods included interviews with teachers and students, analysis of course materials, and observation of class meetings. The student interviews included a storyline method to retrospectively discuss students' experiences of course activities.

In the next study (chapter 4), we explored generalizability of the findings of this extreme case. We examined six different honors courses to discover which factors in the learning environment specifically designed for high-ability students challenged these students, and determined how such challenges were manifested in the design and implementation of these courses. Perceived challenges (factors and intensity) in the different courses were identified from focus group interviews with students. These interviews were also guided by a storyline method to retrospectively discuss students' experiences of course activities. In addition, course materials, observation notes, and audio-recorded classes were analyzed to describe how challenge was established.

*Teacher dilemmas (**chapter 5**)*

Courses of the third study differed with respect to the extent of challenge, which left us wondering why experienced teachers in similar settings made different choices with respect to the level of challenge they provided students. Therefore, we interviewed the teachers of the study in chapter 4 and six more teachers, to analyze their considerations regarding the choice whether or not to challenge students. This study addresses dilemmas teachers encountered when challenging high ability students, revealing how teachers can arrive at different courses of action, based on the same considerations, depending on how they value the consequences. Knowledge of the considerations of teachers relating to dilemmas in practice could provide valuable insight into the complexities of teaching. Additionally, a framework for analysis of teacher dilemmas was developed.

2. Do honors students have more potential for excellence in their professional lives?[2]

Abstract

Universities in many countries increasingly value talent, and do so by developing special honors programs for their top students. The selection process for these programs often relies on the students' prior achievements in school. Research has shown however, that school grades do not sufficiently predict academic success. According to Renzulli's (1986) three-ring model, student characteristics relating to intelligence, motivation and creativity are the most important predictors of excellent achievements in professional life. In this paper, we will investigate whether honors students differ from non-honors students in terms of these characteristics. By means of a questionnaire, more than 1,100 honors and non-honors students at Utrecht University were asked to assess themselves on six characteristics: intelligence, creative thinking, openness to experience, the desire to learn, persistence, and the drive to excel. The results showed that the honors students differed significantly from the non-honors students in terms of the combined variables as well as for the separate variables, with the exception of 'persistence'. The strongest distinguishing factors between honors and non-honors students appeared to be the desire to learn, the drive to excel and creativity, whilst there was little difference in terms of intelligence and persistence. However, the profiles of these differences varied according to the study program. While Law and Humanities honors students differed from their non-honors peers in terms of their drive to excel, Physics honors students were primarily more eager to learn than their non-honors peers, while the LA&S honors students scored higher on creative thinking than non-honors students.

[2] This chapter was published as Scager, K., Akkerman, S. F., Keesen, F., Mainhard, M.T., Pilot, A., & Wubbels, T. (2012). Do honors students have more potential for excellence in their professional lives? *Higher Education,* 64, 19-39

2.1 Introduction

Universities in many countries increasingly value talent, and do so by developing special honors programs for their top students. The objective of these programs is to provide opportunities for students to develop their talents to the full, enabling them to make significant contributions to science and society. Honors students are assumed to have the potential to excel in their future professional lives. It is, however, unclear whether and to what extent these honors students do indeed have this potential in comparison to non-honors students. In contrast with the huge body of research on giftedness in primary and secondary education, empirical research on talent in higher education is surprisingly scarce (Achterberg, 2005; Clark, 2000; Long & Lange, 2002; Rinn & Plucker, 2004). This is remarkable given the growth of programs specifically designed for groups of students who are assumed to be academically talented. Universities often select honors students based on their grades in secondary education and their level of motivation. In his review of the relationship between grades attained at school or at university and adult accomplishment, Hoyt (1966) concluded that grades have hardly any relationship with any measure of future achievement. Twenty years later, Cohen (1984) found similar results in his meta-analysis, indicating that the predictive value of grades for professional success is small. Taylor et al. (1985) came to the same conclusion, demonstrating that grades and standardized tests can predict future grades, but that they do not predict professional excellence. High school grade point averages (GPAs) are used as predictors, as well as scores on scholastic aptitude tests (SATs) and letters of recommendation (Rinn & Plucker, 2004). The area of the selection process, which is based on motivation, can be either active, by means of letters and interviews, or passive, relying on the self-selection of the students.

However, it is questionable whether it is safe to trust these selection methods. We do not know whether these selection methods supply honors programs with students who are significantly more motivated than non-honors students are. Furthermore, not all of the students who are qualified apply for the honors program or college, and a number of them end up in regular programs. Thus, even if application forms, letters, and interviews are used as evidence of motivation, we cannot know whether honors students are more motivated than their non-honors peers.

This study investigates whether honors students differ from non-honors students with respect to qualities that have been found to be essential for exceptional accomplishments in professional life. Further, we examine which of these qualities primarily differentiate between honors and non-honors students. Empirical evidence regarding the specific qualities of honors students is needed if honors programs are to be able to judge whether they have selected the right students, to match their programs to the qualities of the students, and to justify their existence. In order to comprehend what the important factors are, we directed our attention to literature regarding excellent professionals. As honors programs are designed to encourage potential innovative professionals to bloom, it is reasonable to search for aptitudes and personality traits that are required for making creative contributions to a

professional domain in adult life. The body of research derived from retrospective and longitudinal studies on eminence is especially informative in this respect.

In the following section, we will first discuss the important predictors of excellence as found in the literature on excellent professionals. Next, we will describe previous studies that have provided evidence for the differences between honors and non-honors students, leading to the central question of this study.

Predictors of excellence

The concept of 'excellent professionals' has not been clearly defined; studies in the field generally speak of persons with outstanding achievement, or eminence (Trost, 2000), or, like Simonton (1999) does, of 'people who made a name for themselves', referring to their reputation. The theoretical perspective of excellence usually defines how researchers try to assess the construct. According to Simonton (1999), the most common assessment criteria are nominations by experts in the field; occupations of special positions, like political leaders; awards, such as Nobel prizes; and biographical entries in encyclopedias or similar sources. Criteria also depend on the domain. For scientific excellence for example, research productivity, citation ratings, and peer ratings, or combinations of these methods, are commonly used. One major focus of interest in studies of excellence in the last century has been on the conditions and characteristics related or contributing to excellence (Albert, 1969; Friedman-Nimz & Skyba, 2009). In a review of the research on characteristics that predict excellence in adult life, Trost concluded that a combination of characteristics is necessary for outstanding accomplishments in later life; however, no combination of predictors could explain more than 50% of the variance in adult achievement (Trost, 2000: 332).

Several models have been designed in order to represent the combination of factors necessary for excellent performance in work, and one of the first and most renowned models was Renzulli's (1986, 2003) 'three ring conception of giftedness'. Renzulli's model comprises a combination of three interacting basic clusters of human traits: 'above average ability', 'creativity' and 'task commitment', covering the traits of people with the potential to become creative and productive professionals (Renzulli, 1986). Each of the model's three components is necessary for creative performance, and no one component is sufficient in itself. Since 1986, a lot of research has been done in the field of giftedness and several variations, adjustments and alternatives to the three ring model have been created (for instance, Gagne´, 1995; Mönks & Katzko, 2005; Sternberg 2003). Most models, like the Munich model of giftedness (Heller et al., 2005) and Gagné's model (1995), are developmental models, and include the environmental factors that are necessary to allow potential talent to bloom. As we were particularly interested in the personal characteristics which are necessary to excel in work, we used Renzulli's (1986, 2003) multidimensional conception of giftedness. Starting from Renzulli's three rings, we chose to use only general intelligence instead of Renzulli's 'above average ability', which also includes domain specific abilities. General intelligence is likely to be involved as a central characteristic of giftedness (Detterman & Ruthsatz, 1999; Sternberg,

2005; Thompson et al., 2010). Further, 'task commitment' was replaced by the more commonly used concept of 'motivation', depicting task commitment as well as a love for learning. Renzulli (2003) also included 'self-confidence', and 'the ability to identify significant problems within specialized areas' under the heading of task commitment. We left these two out, because of their conceptual overlap with the concept of intelligence.

A brief review of the theory and empirical evidence for each of these three components as predictors of excellence in professional life follows, using a combination of retrospective and prospective longitudinal studies. Retrospective analysis of excellent professionals allows us to determine the adolescent antecedents of their achievements, such as their environment, developmental characteristics and traits. For this study, we were mainly interested in the traits that eminent people showed in their youth. Prospective longitudinal studies indicate whether traits, measured in youth predict their level of achievement in their professional life. According to Simonton (1994), the findings obtained using each method corroborate rather than contradict each other.

Intelligence

Regarding the concept of intelligence, there is broad consensus that cognitive abilities are organized hierarchically, with Spearman's general intelligence topping over cognitive traits such as memory, and verbal and spatial abilities (Carroll, 1993; McGrew, 2009; Lubinski, 2004; Schweizer et al., 2011). General intelligence is defined as 'a very general mental capability that, amongst other things, involves the ability to reason, plan, solve problems, think abstractly, comprehend complex ideas, learn quickly and learn from experience' (Baumert et al., 2009; Gottfredson, 1997: 13).

Traditionally, social scientists have assumed that intelligence is the sole predictor of excellence in later life. Accordingly, the dominant operational definition of giftedness has long been based solely on measures of IQ (Callahan, 2000; Tannenbaum, 1996). Several longitudinal studies have proven that high intelligence is indeed a valid predictor of exceptional performance in professional life (Gottfredson, 1997; Kuncel et al., 2004; Lubinski et al., 2006). In their longitudinal study of a group of students who had achieved exceptional SATs scores before the age of 13, Lubinski et al. (2006) found that the same students, only 10 years later, had obtained an impressive list of achievements, including numerous scientific publications, inventions, and original contributions to literature and the arts. Kuncel et al. (2004), in their meta-analysis, concluded that intelligence tests, specifically the Miller Analogies test, predict performance in graduate studies as well as job performance. In their recent review study, Kuncel et al. (2010) convincingly confirm earlier findings, finding strong correlations between intelligence and job performance, and conclude that 'validities of cognitive ability tests are substantial and useful across industries, job families, and even cultures' (p.333).

Creativity

Traits of creative persons are considered to be a result of interacting cognitive and attitudinal dimensions (Batey & Furnham 2006; Hennessey & Amabile 2010), including creative thinking, openness to experience, intelligence and motivation. Since the latter two are positioned in the two other (overlapping) Renzulli rings, we here focus on creative thinking and openness. The cognitive dimension (creative thinking) is defined in the Dictionary of the Psychological Association (Vandenbos, 2007) as mental processes leading to a new invention, solution, or synthesis in any area. Openness to experience is one of five major domains, which are used to describe the human personality, namely an active imagination, aesthetic sensitivity, attentiveness to inner feelings, preference for variety and intellectual curiosity (Costa & McCrae 1992). Lists of the attitudinal characteristics of creative people are plentiful (Barron & Harrington, 1981; Feist, 1999; McCrea, 1987; Selby et al., 2005; Simonton, 1997; Treffinger et al., 2002). These lists overlap, and their contents make up openness to experience.

Creativity has been proven to be an important quality which enables people to develop into adults who are able to make a significant contribution to their domain (Renzulli, 1986; Simonton, 2000; Trost, 2000). Creativity, measured during adolescence, is a valid predictor of excellent performance in adult life (Feist, 1999; Milgram & Hong, 1993; Simonton, 1988; Tannenbaum, 1983), and some researchers have found creativity to be an even better predictor of future achievements than intelligence. Milgram and Hong (1993), for example, studied a group of high school students over a period of 18 years, and found that creativity and creative performance were better predictors of achievements in adult life than intelligence or school grades. Park et al. (2008) found that intelligence itself predicts creativity. In a longitudinal study over the course of 25 years, the authors selected a large sample of students who were the top 1% of the SAT math test at the age of 13 and tracked their creative productiveness in their adult lives in terms of publications and patents earned. Their findings suggest that quantitative reasoning ability at an early age predicts scientific creativity and innovative accomplishment. These erratic findings regarding the relative importance of creativity and intelligence could be due to the (partial) overlap of the concepts of intelligence and creativity; one can consider divergence and fluidity of thinking as elements of intelligence as well as being the main components of creativity. Conversely, creativity also involves analytical and critical reasoning, as novel ideas must be reflected upon once they are produced. Despite the intertwined nature of intelligence and creativity, the literature on creativity indicates that creativity is more than simply intelligence.

Motivation

Motivation is a multifaceted concept (Pintrich, 1999) and includes notions such as persistence, task commitment, intrinsic or extrinsic interest, the desire to learn and the drive to succeed (Friedman-Nimz & Skyba, 2009). In our study, we narrowed the concept of motivation down to ‘a desire to learn’, ‘persistence’, and ‘the drive to excel’. The desire to learn

is defined as the enjoyment of learning, characterized by an orientation towards mastery, curiosity, and the learning of challenging, difficult, and novel tasks (Gottfried et al., 2005). Adults with an academic interest actively search for cognitive stimulation and insights, seek out in-depth approaches to learning and find enjoyment in engaging in cognitive activities (Schick & Phillipson, 2009; Biggs, 1989). One may expect that the desire to learn and to master the material may lead people to put more time and effort into a task (Lens & Rand, 2000). Accordingly, this aspect relates to another aspect of motivation: 'persistence', which has also been labeled as task commitment, diligence, determination, or perseverance regardless of setbacks and difficulties. The third motivational dimension, which we will call 'the drive to excel', refers to the desire to achieve good scores on external indicators of success, such as grades, as opposed to the mastery orientation, which resembles the 'desire to learn' dimension as described above (Dweck & Leggett, 1988). This motivational factor has also been described as 'external motivation' and 'performance motivation'. Harackiewicz et al. (2000) suggested that the optimal motivational pattern for college students includes goals based on both mastery (the desire to learn) and performance (the drive to excel).

Various studies have shown that all three factors are indispensable for intellectual and creative achievement. There is a consensus amongst researchers in the field that the 'desire to learn' is an important factor for academic achievement in one's school, life and career (Collins & Amabile, 1999; Gottfried et al., 2005; Schick & Phillipson, 2009; Sternberg, 2001). A wide interest and curiosity are essential attributes for gifted individuals (Clark, 2000; Goertzel et al., 2004; Tuttle & Becker, 1983), and are considered to be independent predictors of giftedness (e.g. Gottfried et al., 2005).

From several studies on eminence, we can learn that following the desire to learn, motivational factors such as drive and persistence are predictors of eminence (Csikszentmihalyi, 1992; Friedman-Nimz & Skyba, 2009; Howe, 1999; Renninger, 2009; Simonton, 1994; Treffinger et al., 2002; Trost, 2000; Walberg et al., 2003). Empirical studies have shown that talented professionals must practice within their domain for several hours per day for a full decade before their latent capacity becomes actualized (Roe, 1953; Ericsson, 2006; Howe, 1999; Walberg et al., 2003), showing that persistence is another predictor of excellent achievement. Roe (1953) examined the antecedents of 64 living eminent scientists in various fields using interviews and tests. Although there were marked differences between the groups from different disciplines, all of the scientists were driven by their absorption in their work. A landmark retrospective study on this topic was conducted by Cox (1926). She investigated the lives and achievements of 301 eminent people. Using primary sources such as publications, medals, citations, scores on intelligence tests as well as secondary materials such as biographies, she estimated their intelligence and personality traits, taking into account environmental factors. She concluded that high intelligence alone is insufficient for high achievement in adult life; those who became eminent all had certain personality traits in common, the most important of which were persistence, drive, and passion. The psychologist Howe (1999) traced the lives of eminent men, including for example Charles Darwin.

Referring to childhood traits that contribute to creative productiveness, Howe concluded that motivational aspects like persistence, the capacity to concentrate and to resist distractions and intense curiosity are better predictors of eminence than measures of early intelligence. Walberg conducted several retrospective studies on eminent men and women (Walberg & Stariha, 1992; Walberg et al., 1981), and found that the top ranking traits which eminent men and women showed during childhood, apart from high intelligence, were perseverance and diligence (Walberg et al., 2003).

The best-known longitudinal study on the school and work careers of gifted people was conducted by Louis Terman (1954). Terman kept track of the careers of over 1,500 highly intelligent children over the course of the lives, starting in 1921. Most of these promising children did not fulfill Terman's expectations. In a follow up study, Terman and his colleagues tried to identify the non-intellectual factors that had influenced their levels of achievement. They concluded that the superiority of the successful people was especially marked in volitional traits, such as perseverance and the drive to excel (Terman, 1954). In acknowledgment of these various findings, motivation has therefore been incorporated as a vital factor into many theories of giftedness (for instance, Dai, Moon, & Feldhusen, 1998; Gottfried et al., 2005; Lens & Rand, 2000; Ziegler & Heller, 2000; Renzulli, 1986).

Summarizing, within Renzulli's three rings, six characteristics are present, predicting excellence in professional life: intelligence, creative thinking, openness to experience, desire to learn, drive to excel, and persistence. In the next section, we review studies at university level regarding differences on these characteristics between honors and non-honors students.

Differences between honors and non-honors students

Considering the literature regarding excellence, we found that there was a reasonable amount of literature regarding the characteristics of gifted and talented students in primary and secondary school settings. Research on the characteristics of talented students at university level, however, is limited. Rinn and Plucker (2004) considered the literature on academically talented college students to be outdated, as little work has been published in the last two decades. Achterberg (2005), in her review of the literature about differences between honors and non-honors students in higher education, detected a 'severe lack of descriptive evidence, comparisons, or empirical data based on respectable sample sizes' (p. 5). A study that withstands this critique is that of Long and Lange (2002), who explored personality differences between honors and non-honors bachelor students in a large regional university in the United States. The student group was largely Caucasian, and covered a variety of disciplines, as well as liberal arts. In this study, honors students scored higher on the personality scales of conscientiousness and openness to experience, and were more likely to prepare for class, focused more on grades, and participated more in extracurricular activities than their non-honors counterparts. However, according to the researchers, the magnitude of the differences was small.

Since 2005, a few studies have compared honors students and non-honors students in terms of specific traits. Comparing honors and non-honors students of the entire freshman population at Louisiana Tech University, a comprehensive public university, Kaczvinsky (2007) found that honors students were more academically confident, more intellectually interested, and more open to new ideas than their non-honors peers were. Rinn (2007) focused on the motivational differences between honors and gifted non-honors students, in a sample of 294 bachelor students in various disciplines in a large university in the Midwest of the United States. She found that honors students appeared to be more confident in their abilities. The scarcity of information on the differences between honors and non-honors students that has been gathered in the last two decades suggests that there is a need for additional studies. Moreover, the few existing studies on the differences between honors and non-honors students were conducted in the United States, and therefore could be regionally biased. Furthermore, none of these studies compared honors and non-honors students in terms of the combination of traits which have been found to be predictors of excellent achievement in professional life, which after all is a major objective of honors programs.

In this study, we compare honors and non-honors students on six characteristics that have been found to be central for creative productive professionals (intelligence, creative thinking, openness to experience, desire to learn, drive to excel, and persistence). Additionally, the development of a reliable questionnaire that measures these six characteristics was included as the first objective of this study.

Research questions are:

1. Can a questionnaire be developed for measuring the six characteristics which are found to be central for creative productive professionals that (a) is psychometrically sound, and (2) is brief and readily usable?
2. Do honors students differ from non-honors students with respect to the combination of intelligence, creative thinking, openness to experience, desire to learn, drive to excel, and persistence?
3. Which of these characteristics primarily differentiate between honors and non-honors students?

2.2 Methods

Participants

The participants in this study comprised 1,122 students, including 467 honors students and 655 non-honors students of Utrecht University, a large research university (30.344 students). In this university, about 21 honors programs and four honors colleges have been developed in the last 15 years, disciplinary as well as interdisciplinary. The participants were undergraduate students (41% 1st year students, 37% 2nd year students, and 22% 3rd year students) of seven different bachelor programs: Law, Humanities, Mathematics, Physics, Earth Sciences, Innovation & Environmental Sciences, and Liberal Arts and Sciences (LA&S). These programs had well-established honors programs with substantial student enrollment,

except for the Humanities honors program, which was newly developed and started with 21 students. Despite of this, we included this program in order to keep the sample broad. Student samples of the programs included in the study were chosen by the program coordinators by selecting one or more classes in all three bachelor years to hand out the questionnaires. Student attendance in the selected classes and thus the response in the data gathering was nearly complete, as these classes were either very important or compulsory, both for honors and non-honors students.

Honors programs, in the Netherlands as well as elsewhere in the world, are diverse in the ways in which they are designed. In Utrecht University, there are several models for honors programs at undergraduate level, which are comparable to those in the United States. The 3-year Liberal Arts and Sciences Honors College is a full program, designed especially for honors students. Another common model is an enrichment program for a selected group of students, including specially designed honors courses, which are additions to the regular study program. The Humanities program belongs in this category. Most programs at Utrecht University combine the two models, meaning that honors students follow part of the regular program, engage in substitute components such as research projects, and follow an additional enrichment program. The Law and Earth Sciences and Innovation and Environmental Sciences courses in Utrecht University are examples of this model. Honors students are selected for these programs based on their prior achievements (school grades) and on their motivation, which is reflected implicitly in their self-selection or explicitly checked by their application letters and/or intake interviews.

Procedures

The students completed a paper and pencil questionnaire during class or in the break between classes. The questionnaires took approximately 20 min to complete, and were semi-anonymous: the students were asked to fill in their student number so that we were able to retrieve their gender, age and year of study from the administrative system. Their names remained unknown to the researchers. Two subjects were removed from the analysis as they were missing more than 5% of their data. Eleven students were of non-traditional age (older than 27) and were also removed from the study, leaving 1,109 cases from the original 1,122.

Materials

In the absence of an adequate existing questionnaire for university level that measures all of the components of Renzulli's (1986) model, we developed a self-report questionnaire using parts of existing validated instruments. These instruments included: Goldberg's (1999) 'intellect' scale from the International Personality Item Pool (IPIP-NEO) for 'intelligence' and the Scale of Creative Attributes and Behavior (SCAB) instrument (Kelly, 2004) for 'creative thinking'. The scale measuring openness to experience was founded on a Dutch version of the Big Five Openness to Experience scale (Gerris et al., 1998), and for the 'desire to learn' scale, the International Personality Item Pool Values In Action (Peterson and Seligman, 2004) was

used as a basis. The items of the 'persistence' scale were derived from the Perseverance IPIP-VIA scale (Peterson & Seligman 2004), and the 'drive to excel' was measured using three items adapted from the work of Elliot & McGregor (2001). Based on an exploratory study, a questionnaire was developed, which consisted of 63 items. In order to ensure that the format of the items was similar across the whole questionnaire, the scales of the items were adjusted to fit a seven-point Likert type scale, ranging from 'not at all true of me' (1) to 'very true of me' (7).

Both English and Dutch versions of the questionnaire were administered. A back translation procedure (Brislin, 1986) involving four English-Dutch bilingual university instructors was applied to ensure cross-cultural conceptual equivalence. The back translations were compared until consistent meanings were obtained.

Analysis

Multivariate analysis of variance was conducted, using General Linear Models (GLM) with SPSS 16, to find out whether the (combination of) the six talent factors differ between honors and non-honors students, which was the central question of this study. The independent variables were honors (yes or no), and study program. Gender and year of study were entered as covariates. The dependent variables were the six scales; the 'talent factors' described above. The order of entry of the independent variables was study program, then honors, as our main interest was in the differences between honors and non-honors students. Any missing values (2%) were replaced using the two-way imputation method described in van Ginkel and van der Ark (2008). The results of the evaluation of normality, linearity and multicollinearity were satisfactory. As a preliminary test for robustness, Tabachnick and Fidell (2007) suggest comparing sample variances for each dependent variable across the groups. The ratio of the largest to the smallest group variance did not approach 10:1 for any dependent variable. As a matter of fact, the largest to the smallest group ratio was about 1.9:1. The sample sizes were widely discrepant, with a ratio of 4.3:1. However, with very small differences in variance, this discrepancy in sample sizes does not invalidate the use of multivariate analysis (Tabachnick & Fidell 2007).

In order to answer the second research question, concerning the importance of each of the six talent factors' unique contribution to the differences between the groups (honors and non-honors), discriminant analysis was conducted. The six variables were not independent and show significant correlations that vary between 0.17 and 0.53. Intelligence, creativity and the desire to learn have the highest correlations (between 0.41 and 0.53). As these variables are correlated, the univariate tests do not properly show the relative importance of the variables in distinguishing honors students from non-honors students. In order to be able to interpret the contribution of each of the variables to the differences between the two groups and to determine which of the variables discriminate most, direct discriminant analysis was used with the six talent scales.

In this study we used a significance level of 5%.

2.3 Results

The first section reports of the scale descriptives. In the second section, we report whether the two groups (honors and non-honors) differed from one another in terms of the six talent factors. In the third section, we examine which of the six talent factors accounted for the largest differences between the honors and non-honors groups.

Scale descriptives

The original questionnaire consisted of 63 items, divided over six scales: intelligence, creative thinking, openness to experience, desire to learn, drive to excel, and persistence. Based on reliability and factor analyses, items were deleted, reformulated, or replaced. The resulting questionnaire consisted of 31 items (see Appendix 1'). The corrected scales were separately factor-analyzed, and in all six scales, the items loaded on a single factor, showing unidimensionality of the scales (Field, 2009). Using the often used rule of thumb of 0.7 as an acceptable alpha value, and given the unidimensionality of the scales (Cortina, 1993; George & Mallery, 2003; Nunnaly, 1978), all scales are acceptable, despite the small number of items in some of the scales. Table 2.1 presents a sample item and the reliability (Cronbach's alpha) for each of the six scales.

Table 2.1
The six scales of the questionnaire with sample items and reliability scores

Scales	*Sample item*	*α*
Intelligence (six items)	I am quick to understand things.	0.71
Creative thinking (six items)	I am often able to make connections between seemingly unrelated things or situations.	0.76
Openness to experience (four items)	I am imaginative.	0.70
Desire to learn (six items)	I want to learn as much as possible.	0.78
Drive to excel (three items)	It is important for me to do better than other students.	0.83
Persistence (six items)	I give up easily.	0.73

In order to examine whether or not the items were distributed correctly over the six scales, a principal component analysis was conducted on the 31 items with oblique rotation (Oblimin). The Kaiser–Meyer–Olkin (KMO) measure verified the adequacy of the sample for analysis (KMO = 0.87). Seven components had eigenvalues over Kaiser's criterion of one, and collectively explained 54% of the variance. The pattern matrix showed seven factors, which were almost identical to the six original scales, except for the persistence scale. The persistence items loaded into two subscales: one focused on hard work and determination, and the other focused on perseverance regardless of setbacks. For conceptual reasons, these two factors of persistence were kept in one scale. With a cutoff of 0.40 for the inclusion of a

variable in the interpretation of a factor, 24 of 31 items loaded only on their original factor; one item did not reach the 0.40 cutoff, another item, 'preference for complexity', loaded higher on the creativity scale than on its original scale of intelligence. Theoretically, however, this is not surprising, as preference for complexity, when understood as tolerance of ambiguity, is related to creativity (Zenasni et al., 2008). Five items loaded into two scales, which was to be expected considering the overlapping of the three rings in Renzulli's model. The overlap was found mainly between the Creativity and Desire to learn scales. In order to keep the reliability sufficiently high for all of the scales, all items were kept in their original scales. Intelligence and Drive to excel correlated 0.30, which was modest and limited to one pair of factors; remaining correlations were low.

Differences between honors and non-honors students

The second research question focused on determining the differences between honors and non-honors students with respect to the six talent factors (intelligence, creative thinking, openness to experience, the drive to excel, the desire to learn, and persistence). Table 2.2 shows the means and standard deviations.

Table 2.2

Descriptive statistics: means and standard deviations

Descriptives		All Honors (N= 373) Non-honors (N= 493)		Law Honors (N= 71) Non-honors (N= 187)		Humanities Honors (N= 19) Non-honors (N= 91)		Physics Honors (N= 32) Non-honors (N= 48)		LA&S Honors (N= 242) Non-honors (N= 493)	
		Mean	SD	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Intelligence	Non-honors	4.8	.8	4.8	.7	4.8	.8	4.9	.9	4.8	.8
	Honors	5.1	.7	5.3	.6	4.9	.9	5.2	.8	5.1	.8
Creative thinking	Non-honors	4.9	.7	4.9	.7	4.8	.7	5.0	.7	4.9	.7
	Honors	5.4	.7	5.3	.7	5.1	.7	4.8	.6	5.5	.7
Openness to experience	Non-honors	4.9	.9	5.0	.9	4.7	1.0	4.6	.9	4.9	.9
	Honors	5.3	.8	5.3	.8	5.5	.7	4.7	.8	5.4	.7
Drive to excel	Non-honors	4.1	1.4	4.3	1.2	3.8	1.5	4.3	1.5	4.1	1.4
	Honors	4.8	1.3	5.4	.9	5.1	1.1	4.1	1.2	4.7	1.3
Desire to learn	Non-honors	5.2	.8	5.1	.8	5.5	.8	5.4	1.0	5.2	.8
	Honors	5.8	.7	5.6	.8	5.5	.8	5.8	.6	5.8	.7
Persistence	Non-honors	4.9	.9	5.1	.8	4.7	1.0	4.9	1.0	4.9	.9
	Honors	5.2	.8	5.5	.7	5.1	.9	4.6	.8	5.3	.8

Mean scores for all students show highest scores for their Desire to learn (5.2 for non-honors and 5.8 for honors groups), while the scores for Drive to excel are lowest (4.1 for non-honors and 4.8 for honors groups). These two characteristics also show the largest difference between honors and non-honors groups. The scores on Drive to excel show the largest standard deviations, especially for the non-honors group. Further, it is noticeable that honors students in the physics department assess themselves lower on Creative thinking, Drive to excel, and Persistence than their non-honors peers do.

GLM analysis using Wilks' criterion, showed that the combined dependent variables were significantly affected by the distinction in honors and non-honors groups (p= .000; partial g2= .04), the study program (p= .000; partial g2= .04), and gender (p= .000; partial g2= .05), but not by year of study. Univariate between-subjects tests showed significant differences between honors and non-honors students when adjusted for study program, gender and year of study, for five of the six variables; 'persistence' being the variable that did not show a significant difference.

As our findings showed that the study program is a significant variable, we examined how the differences between honors and non-honors students varied across study programs. We conducted separate GLM analyses for four study programs: Law (91 honors and 261 non-honors students), Humanities (21 honors and 121 non-honors students), Physics (40 honors and 59 non-honors students), and LA&S (301 students). The non-honors students of the whole data set (N = 469) were used as a control group for the LA&S honors group. The other three study programs were not included in the analysis, due to a lack of sufficient data for the honors groups. In these additional analyses, only gender was included as a covariate.

The results of the separate multivariate and univariate analyses for the four study programs are presented in Table 2.3. For all of the study programs, the multivariate analysis showed significant differences between honors and non-honors students, with effect sizes (partial g2) between 0.15 and 0.25. The results of the univariate analysis between subjects showed different results for the four study programs. For the Law students, univariate between-subjects tests revealed statistically significant differences between honors and non-honors students for all six variables, with the largest effect sizes for the desire to learn (partial g2 = 0.15) and intelligence (partial g2 = 0.10). For Humanities students, univariate between-subjects tests demonstrated that the differences between honors and non-honors students were statistically significant for openness to experience (partial g2 = 0.08) and the desire to learn (partial g2 = 0.09).

Table 2.3

Multivariate and univariate analyses of variance for the six talent factors in four study programs

	Law				Humanities				Science				Liberal Arts			
Talent factors	MS	F	p	Partial η^2	MS	F	p	Partial η^2	MS	F	p	Partial η^2	MS	F	p	Partial η^2
Multivariate	-	14.07	0.00	0.25	-	2.89	0.00	0.15	-	3.49	0.00	0.23	-	29.40	0.00	0.20
Intelligence	12.19	28.61	0.00	0.10	0.44	0.70	0.41	0.01	0.79	1.03	0.31	0.01	15.08	25.64	0.00	0.03
Creative thinking	7.09	14.23	0.00	0.05	1.07	2.33	0.13	0.02	0.44	1.02	0.32	0.01	58.41	127.31	0.00	0.15
Openness to experience	3.16	4.50	0.04	0.02	8.32	9.38	0.00	0.08	0.01	0.01	0.93	0.00	40.67	60.06	0.00	0.08
Desire to learn	56.13	46.03	0.00	0.15	22.17	11.13	0.00	0.09	0.28	0.14	0.71	0.00	51.61	27.12	0.00	0.04
Drive to excel	14.05	23.36	0.00	0.08	0.03	0.05	0.83	0.00	3.17	3.91	0.05	0.05	53.82	87.45	0.00	0.11
Persistence	9.37	8.26	0.00	0.07	1.37	1.59	0.21	0.02	2.53	2.96	0.09	0.04	21.67	30.14	0.00	0.04

Univariate between-subjects tests for the six variables did not show significant differences in the Science groups, and finally, univariate tests for the LA&S students showed a significant effect for all six variables, with the largest effect sizes for creative thinking (partial g2 = 0.15) and the drive to excel (partial g2 = 0.11).

Main contributors to the differences between honors and non-honors students

In order to be able to interpret the contribution of each of the variables to the differences between the two groups, direct discriminant analysis was used with the six talent scales as predictors of membership of the honors or non-honors group, corrected for gender. This was executed for the whole data set, as well as separately for the four study programs. The discriminant function that was calculated revealed a significant overall difference between the honors and non-honors students (p= 0.00, Wilk's= 0.83, v2(7)= 159, canonical R2= 0.17). This means that 17% of the variance can be accounted for by the combined predictors. The standardized discriminant function coefficient predictors and the discriminant function, as shown in Table 2.5, suggest that the best predictors for distinguishing honors students from non-honors students were the desire to learn, the drive to excel and creativity.

As shown in Table 2.4, separate discriminant function analyses for the four study programs revealed significant differences between the honors and non-honors groups.

Table 2.4

Discriminant function analysis for all groups and for separate study programs (Law, Humanities, Science and Liberal Arts)

	Wilk's λ	$\chi^2(7)$	p	Canonical R^2
All	0.83	159.00	0.00	0.17
Law	0.73	80.69	0.00	0.27
Humanities	0.79	25.03	0.00	0.21
Science	0.78	19.01	0.01	0.23
Liberal Arts	0.83	158.87	0.00	0,20

The standardized discriminant function coefficients of predictors, as shown in Table 2.5, show that the best predictors for distinguishing honors students from non-honors students differ for the various study programs. The main predictor for Law honors students is their drive to excel (0.60). For Humanities students, their drive to excel (0.56) and their openness to experience (0.55) contribute most to the differences between honors and non-honors students. For Physics students, their desire to learn (1.00) and creative thinking (-1.01) show the greatest difference between honors and non-honors students. The main predictor for the LA&S honors students was creative thinking (0.62).

Table 2.5

Standardized discriminant function coefficients for all groups and for separate study programs (Law, Humanities, Science and Liberal Arts) corrected for gender

	All groups Function	Law Function	Humanities Function	Science Function	Liberal Arts Function
Intelligence	0.05	0.29	-0.06	0.53	-0.18
Creative thinking	0.37	0.15	0.05	-1.01	0.62
Openness to experience	0.27	0.12	0.55	-0.02	0.33
Drive to excel	0.36	0.60	0.56	-0.01	0.21
Desire to learn	0.41	0.18	-0.34	1.00	0.35
Persistence	0.12	0.33	-0.10	-0.71	0.07

2.4 Conclusions and discussion

The central question of this study was whether honors students differ from non-honors students with respect to the talent factors (intelligence, creative thinking, openness to experience, persistence, the desire to learn and the drive to excel) that have been found to be essential for exceptional accomplishment in professional life. The results showed that honors students are significantly different from non-honors students in terms of the combined as well as the separate variables, with the exception of persistence. The effect sizes were medium when measured within the separate disciplines, but small for the group as a whole. Two issues need to be taken into account here. First, honors and regular programs do not unambiguously distinguish talented students from regular students in their selection processes. As Rinn and Plucker (2004) state, groups of non-honors students will also include a proportion of potential honors students. Such potential honors students could have various reasons for not applying to an honors program. They might not want to belong to an 'elite' group or might not want to spend extra time on a study program. In addition, these students may simply not be aware of the existence of honors programs, which is plausible given that some of these programs are relatively new. A second issue to take into account is that effect size in the combined analysis has been affected by the different trends we found across study programs.

The second research question concerned which of the talent characteristics contribute most powerfully to the differentiation between honors and non-honors groups. The strongest distinguishing factors for honors and non-honors students appeared to be the desire to learn, the drive to excel and creative thinking, while intelligence and persistence did not differentiate groups very much. The strong distinguishing value of creative thinking was unexpected, as creativity is not an explicit selection criterion for most honors programs. Intelligence was the weakest factor. This was surprising, as honors students are selected on their average school grades, and grades are affected by intelligence (Lubinski, 2004). An explanation for the negligible distinguishing value of intelligence could be found in the fact that average students are more likely to overestimate their ability than gifted students,

whereas gifted students tend to base their judgment of their ability to succeed more accurately on the actual difficulty level (Dai et al., 1998; Pajares, 1996). Persistence was assumed to be necessary for high achievement in secondary education, which is in turn a selection criterion for many programs. Chamorro-Premuzic and Arteche (2008) suggested that persistence and intelligence compensate for each other in terms of achievement, which could explain the low distinguishing value we found for persistence. This would imply that these honors students were intelligent enough to achieve well in school without having to put much energy into their work. The other two motivational factors however, did differentiate between honors and non-honors students. In comparison with the non-honors group, these honors students of Utrecht University had a greater desire to learn, which concurs with Kaczvinsky's (2007) findings in a similar study in the US. The drive to excel is not very dominant in the Dutch educational culture, where students are not as much grade-oriented as students in for example the United States are. It is therefore not surprising that the mean scores for 'drive to excel' are lower than the scores on any of the other factors. The drive to excel however was one of the important traits differentiating honors from non-honors students, indicating that honors students aim to get the most out of their education: in learning profits as well as the recognition of it.

As we found substantial differences between honors and non-honors students across disciplines, we further explored these differences within four of the study programs. In all programs, honors groups differed significantly on the combined talent factors from their non-honors peers. The profiles of these differences, however, varied according to the study program. While Law and Humanities honors students differed from their non-honors peers in terms of their drive to excel, Physics honors students were primarily more eager to learn than their non-honors peers were, while the LA&S honors group distinguished themselves predominantly in terms of creative thinking. The variation in results between the study programs could be explained firstly by the differences in selection procedures used in the various programs. In the Law and LA&S programs, a combination of grades, application letters and interviews are used, while in the Physics and Humanities departments, selection procedures predominantly rely on self–selection. The differences between honors and non-honors students in both the Law and LA&S groups were relatively strong, with significant differences in each of the six talent factors. In the two programs with less comprehensive selection methods, the differences between honors and non-honors students were less convincing, with Humanities students showing significant differences for only two of the talent factors, and Physics honors students not scoring significantly higher for any of the factors. It may well be that comprehensive selection procedures account for more distinctive differences between honors and non-honors students. A second explanation for the differences we found between study programs could be found in the diversity of the academic cultures of disciplines (Becher & Trowler, 2001). Differences in academic cultures imply different habits, values, and rules, serving as a background reference point for students' assessment of their characteristics. In a competitive culture for example, students might assess their own drive to excel lower than in a culture with an egalitarian tradition.

The unusual scores of the Physics honors group necessitate further reflection. Physics honors students did not assess themselves to be significantly more intelligent than their non-honors peers, and their score for 'persistence' was noticeably lower. These low scores for persistence are surprising, as in their honors program, Physics students combine two bachelor programs (physics and mathematics) which can be assumed to be quite demanding. It is possible that these students are not accustomed to working hard, because there was no need for them to do so at school as they were intelligent enough to succeed with little effort (Chamorro-Premuzic & Arteche, 2008). The insignificant differences in intelligence scores between Physics honors and non-honors students were also unexpected, while honors students had significantly higher average school grades (8.1 for honors and 7.6 for non-honors). A clarification might be found using the 'big fish little pond' effect, which will be explained below.

Limitations

Although this study provides important insights in the differences between honors and non-honors students across their study programs, there are several limitations of this study that need to be mentioned. First, this study used a sample of students at one university, and so the results cannot be generalized unproblematically. Second, the data presented in this study are all self-reported. Although this is a logical and defensible methodology in its own right, self-report questionnaires could threaten the validity of this study. One threat is the general tendency for individuals to respond in socially desirable ways. The anonymity of the survey, which allowed respondents to respond truthfully, was intended to mitigate this threat. Further, the influence of the reference group could influence students' self-assessment. Although the questionnaire focused on measuring personal dispositions, their peer group may well influence students' assessments. This effect, referred to by Marsh (1987) as the 'big fish little pond effect', concerns the selected group of highly capable peers that students are part of once they enroll in an honors program. According to Marsh, students tend to assess their own competence in comparison with their peer group, which, for honors students, is a selective group of highly able students. We tried to reduce the influence of this effect by avoiding items that directed students to compare themselves to their peers and by focusing our attention on relatively stable dispositions. Nonetheless, there are indications that the students used their peer group as a point of reference. The Physics honors group, for example, assessed their intelligence lower than the Law honors group did (5.2 and 5.4), whereas the Physics students' average school grades were a great deal higher (8.1 and 7.5). Assuming that the honors groups assessed their qualities as such because of their highly able reference groups, the differences between honors and non-honors students per program may have been underestimated in this study.

Implications

As the honors students in this study differed significantly from non-honors students in terms of the characteristics which predict excellent professional achievement, it seems justifiable to provide special attention for the education of these students. After all, the effectiveness of education depends to a large extent on the fit between the learning environment and the abilities, interests and motivation of the students (McKeachie, 1986; Snow, 1986; Pascarella & Terenzini, 1991). These students need to be challenged appropriately and they require a learning environment that matches their abilities, interests and level of motivation. Honors students, however, are not a homogeneous group; the differences we found between these programs emphasize the value of insights in the characteristics of honors students. These differences across disciplines provide grounds for program leaders to reflect on their selection methods. The programs with the most extensive selection procedures showed the strongest differences between honors and non-honors groups. Self-selection seems to attract honors students with a strong drive to excel, while the desire to learn and persistence are presumably more important motivational factors for excellent achievement in professional life (Ericsson 2006; Howe 1999). Furthermore, the differences between the student groups across disciplines indicated that the appropriate learning environments and teaching methods could differ between the various programs, providing reasons for the importance of matching the programs to the characteristics of the students in the honors groups.

The findings of this study provide three directions for further research. One direction would be the selection of honors students. The results of this study indicate that a more comprehensive method for the selection of honors students could provide a sharper distinction between honors and non-honors groups. In future research, it is worthwhile to focus on the selection processes and the effectiveness of the various elements of the methods of selection. A second area of interest is the relationship between the (interacting) talent characteristics and success in professional domains. The results of the present study suggest that honors groups do have the characteristics needed for excellent performance in professional life, but we do not know whether they have these traits in sufficient measure and in successful combinations. Furthermore, as we found different talent profiles across disciplines, it could be valuable to relate the relative importance of these characteristics to academic domains. A third field of interest is the interaction between the talented students and their learning environment. This study focused on the characteristics students require for excellence in their professional lives, leaving out the crucial role of their environment in developing students' talents. A developmental perspective is needed to provide insights into how to teach honors students effectively. A better understanding of the interaction between students' motivational, creative and intellectual aptitudes on the one hand, and the learning environment and teaching methods on the other hand, could lead to better learning in honors groups.

3. Challenging high-ability students[3]

Abstract

The existing literature on indicators of an optimal learning environment for high-ability students frequently discusses the concept of challenge. It is however not clear what, precisely, constitutes appropriate challenge for these students. In this study, we examined an undergraduate honors course, Advanced Cell Biology, which has succeeded extremely well in challenging students. Methods included interviews with teachers and students, analysis of course materials, and observation of class meetings. As part of their course, the students developed a research program according to national scientific standards, which they did successfully, according to an external jury of experts in the field. The challenge faced by the students comprised the complexity of the task, the high expectations placed upon them, and the lack of teacher direction. Our results indicate that students' perceived learning peaked in a period of over-challenge and, although students felt worried and frustrated in this period, their efforts increased.

[3] This chapter was published as Scager, K. , Akkerman, S. F., Pilot, A., & Wubbels, T. (2012). Challenging high-ability students, *Studies in Higher Education*, DOI:10.1080/03075079.2012.743117

3.1 Introduction

The existing literature on indicators of an optimal learning environment for high-ability students frequently discusses the concept of challenge (Clinkenbeard, 1994; Freeman, 1990; Heller, 1999; Lens & Rand, 2000; Marra & Palmer, 2004; Reis & Renzulli, 2010; Sayler, 2009). In comparison to average ability students, high-ability students are quicker thinkers, more flexible in their use of strategies, have better memories, know more, and prefer complexity (Freeman, 1998; Shore & Kanevski, 1993; Wallace, 2000). Given their ability, it was found that these students in courses aimed at average students often experience boredom (Gallagher, Harradine, & Coleman, 1997), leading to a loss of motivation (Hoekman, McCormick, & Gross, 1999; Lens & Rand, 2000), which in turn can lead to underachievement (e.g., Gentry, Gable & Springer 2000).

Although the need for challenge is often mentioned in studies on learning needs of gifted students, we have not been able to find a definition of the concept in educational literature. Therefore, we used Webster's Dictionary, in which challenge is defined as a 'test of one's abilities or resources in a demanding but stimulating undertaking'. This definition corresponds with the well-known 'flow-model' of Csikszentmihalyi (Csikszentmihalyi, 1975; Csikszentmihalyi, Abuhamdeh, & Nakamura, 2005). Flow is a state of intrinsic motivation in which people are fully engaged in a task for the sake of the activity itself. A crucial condition for intrinsic motivation in flow theory is the compatibility of the perceived levels of skills and challenge. The importance of balancing challenge and ability is also emphasized in several contemporary theories of intrinsic motivation (Deci and Ryan, 1985; Pekrun, 2006). Pekrun's control-value theory of motivation posits that boredom can be experienced if demands are either too high or too low. Deci and Ryan (1985) in their theory of intrinsic motivation, discuss the importance of 'optimal challenge' in sustaining intrinsic motivation. More recently, Keller and Bless (2008) tested the balance hypothesis, and concluded that 'the perceived fit of skills and task demands is important with respect to the emergence of intrinsic motivation' (p. 207).

The level of challenge thus needs to match the level of skills. A person can also experience 'over-challenge', meaning that the level of challenge exceeds the level of ability. Over-challenge can lead to negative emotions like worry and anxiety (Csikszentmihalyi, 1975, p. 49), which, in turn, may hamper motivation (Pekrun, Goetz , & Titz, 2002; Pekrun, 2006; Sansone & Thoman, 2005). Creating challenge for students is, therefore, not a straightforward matter. In this study, we wanted to determine how challenge could be created for high-ability undergraduate students in such a way that their (perceived) learning, as well as their motivation, is stimulated.

Challenging high-ability students

Many universities within and outside Europe have separate honors programs for their top students in order to maximize development of their talents. Although these programs testify to the value universities place on their high-ability students, much is still to be understood

about how education should be specifically designed for these students: appropriate challenge seems to be a crucial element in understanding effective education for high-ability students.

Research on the experiences of gifted students at secondary school level indicates that they do not feel challenged by 'jumping through the hoops' in the pre-structured courses that dominate most of education (Reis & Renzulli, 2010). How, then, should learning environments in university honors courses be designed to better challenge high-ability students? Empirical studies on challenging learning environments for high-ability students at university level are scarce; the vast majority of research has been conducted in primary and secondary education. Moreover, research on the effectiveness of learning environments for gifted students in primary and secondary education has mainly focused on the curriculum level, specifically on the effect of grouping, acceleration, and enrichment, and on combinations of the three. In this study, we focus on the learning environment at the course level in classes where gifted students are already grouped, and we are interested in what aspects of the learning environment constitute challenge. This focus reduces the number of prior research studies available to us, although some findings in studies of gifted students at school have been helpful.

Based on a synthesis of research on gifted students in primary and secondary education, Rogers (2007) points to five lessons, which education practitioners should take into account, three of which seem to be related to challenge: offering students consistent challenge, providing opportunities to work independently, and focusing on depth and complexity. 'Consistent challenge' refers to a progressive increase in the level of difficulty. Based on analysis of empirical studies on the efficacy of curriculum models for gifted students in primary and secondary education, Van Tassel-Baska and Brown (2007) suggest emphasizing higher-level skills and inquiry-based learning. From a theoretical point of view, Shore and Kanevsky (1993) arrive at similar conclusions. Based on their review of differences in thinking processes between gifted and non-gifted learners, they state that gifted students benefit from learning from the standpoint of an inquirer. In a similar vein, Snow and Swanson (1992) propose discovery learning, less scaffolding and less structure. Gruber and Mandl (2000) suggest that situated learning - learning that takes place in the same context in which it is applied - is especially suitable for gifted students. In their study, Kanevsky and Keighly (2003) interviewed gifted high school students focusing on the reasons why they were bored. They found that what these students needed to overcome boredom was more choice and control over their learning, a higher level of challenge and complexity, and caring teachers.

Only a few studies focus on how to teach high-ability students at university level. Marra and Palmer (2004), in their study on the learning preferences of university students with high and low developmental levels, conclude that high-ability students prefer more independence, less structure, as well as more challenge. Their findings focus on both learning tasks, which should be challenging, and the role of the teacher, which should allow student autonomy. Conversely, in a similar study of students aged 14 to 24, Freeman (1990) found that those with a high IQ prefer more communication with the teacher and more feedback, findings that seem

inconsistent with the conclusions reached by researchers like Marra and Palmer, who suggest that less 'scaffolding' and more student autonomy are essential to optimum learning. 'Challenge', in these studies, is not clearly defined, but in all cases seems to refer to difficulty, including difficulty of the task (e.g., complex and unstructured) and difficulty of the process (e.g., an independent way of working).

Although the above studies point in same direction when it comes to organizing challenge in education for gifted students, we see two limitations. First, the factors identified as constituting challenge are of a rather general nature. Second, the studies have not considered the level of challenge nor how this level of challenge relates to learning and motivation. As we will describe in the next section, it is necessary to consider the level of challenge in relation to the level of ability, since it is this relation that has been theorized as affecting motivation as well as learning.

Flow-model related to learning and motivation

Csikszentmihalyi's (1975) flow-model is founded on the assumption that, in a state of flow, learning in itself can be intrinsically motivating. In order to maintain the flow state, tasks should increase in their level of difficulty as the students' skills, and thus their level of ability, improve. As long as the level of challenge is at, or just above, the ability level, learning and intrinsic motivation go together.

Empirical studies have shown that flow experience is, indeed, associated with positive outcomes such as improved performance (Nakamura and Csikszentmihalyi, 2005; Engeser & Rheinberg, 2008; Klein, Rossin, Guo, & Ro, 2010; Volmeyer & Rheinberg, 2006). Volmeyer and Rheinberg (2006), for example, carried out two studies examining the effect of university students' flow experiences on their performance in final examinations and they conclude that the degree of flow experience positively affects achievements in this context. Klein, Rossin, Guo and Ro (2010) also investigated the effects of flow on learning outcomes, in this case on a university management course. Their findings show that a state of flow positively affected students' perceived learning of the subject matter.

A state of flow thus relates to enhanced learning and motivation. The equilibrium between perceived challenge and abilities, however, is fragile. 'Over-challenge' as well as 'under-challenge' can cause negative emotions (Csikszentmihalyi, 1975) and, in turn, negative emotions interfere with motivation (Pekrun, Goetz, & Titz, 2002; Phillips & Lindsay, 2006; Shernov, Csikszentmihalyi, Schneider, & Shernoff, 2003). Under-challenge (tasks that are too easy) has been found to lead to boredom (Csikszentmihalyi, Rathunde, & Whalen, 1993). In their study of gifted students at high school level, Hoekman, McCormick, and Gross (1999) found that students' levels of stress were substantially higher when they were placed in unchallenging classroom settings. Conversely, stress was considerably reduced for these students when they were subjected to higher levels of challenge and they were subsequently successful in meeting the challenge.

According to flow theory, challenge levels that exceed the perceived ability level too much lead to feelings of worry and anxiety (Csikszentmihalyi, 1975). Research on the influence of negative emotions on learning has primarily focused on test anxiety. The findings of these studies show that worry is associated not only with loss of motivation, but also with low performance (Hembree, 1988, 1990; Pajares & Urdan 1996). In the last decade, research on the influence of emotions on motivation and learning has expanded from outcome emotions, such as test anxiety, to activity emotions pertaining to ongoing achievement related activities, including frustration when dealing with difficult tasks (Pekrun, 2006).Findings suggest that emotions affect interest, metacognition, and effort, which in turn mediate students' performance (Pekrun, 2005).

For the purposes of our research, Csikszentmihalyi's flow theory (1975) was used as a conceptual framework (see Figure 3.1). The theory is relevant to our study as it allows us to conceptualize the relationship between the challenge-ability level of a learning environment and the cognitive and emotional state of motivation and learning. We slightly adjusted the original model. Whereas Csikszentmihalyi's model focuses on the experience and feelings of intrinsic motivation (flow), we are specifically interested in intrinsic motivation and learning as outcomes of the challenge-ability balance.

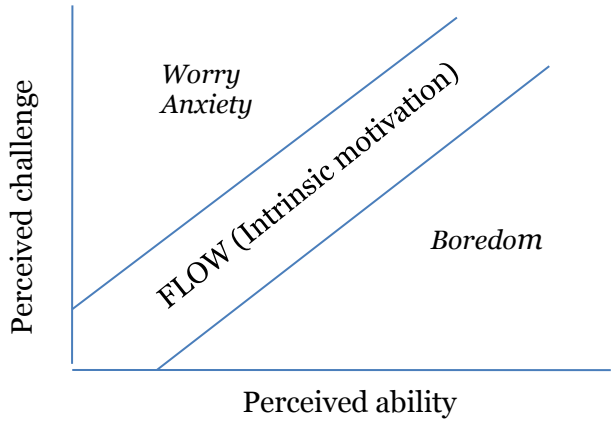

Figure 3.1 *Flow model*

Understanding the challenge-ability balance seems particularly relevant for honors programs, which often aim for high standards. For this study, we selected a course that, according to previous student evaluations, seemed to succeed extremely well in challenging honors students to work very hard and achieve beyond their expectations (Wiegant, Scager, & Boonstra, 2011). It was apparent, therefore, that studying this learning environment would help to both identify and characterize elements that challenge students to produce their best work.

Three research questions are central in this study:

1. How do students experience the balance between levels of challenge and ability, throughout the course?

2. How do students' perceived learning and motivation relate to the perceived challenge and ability levels?
3. What factors constitute challenge in this course?

3.2 Methods

Context and course

Advanced Cell Biology is a third year course of the University College Utrecht (UCU), an international Liberal Arts and Sciences honors college. UCU students are carefully selected on their academic excellence, their curiosity, and motivation. Students complete four courses per 15-week semester, each course having a workload of about 200 hours (7.5 European Credit points) consisting of 60 contact hours (i.e., 4 contact hours per week in two sessions) and approximately 140 hours of self-study. The central task for students in the Advanced Cell Biology course was, to formulate, in groups, a research program consisting of three separated doctoral research proposals. In addition, the research program needed to meet the criteria used for professional research proposals. At the end of the course, students presented and defended their proposal before an expert jury. The course was strongly student-driven. The teachers devised the broad parameters but, within these general guidelines, the students were free to make their own decisions about the content of their research project, the organization of processes, as well as the nature of class meetings. The course comprised four phases: (1) orientation; (2) formulating research questions; (3) finding and choosing methods and techniques; and (4) writing, presenting and defending the research program. Table 3.1 provides a brief description of the four phases.

Assessment by the jury of the novelty, originality, feasibility, and readability of the final product, the quality of the presentation, and the way the students defended their work before the jury all formed important parts of students' final grade. It should be stressed here that each of the students was responsible for the whole program and that the jury could question them on all aspects of the program. A more detailed description of this course can be found in Wiegant, Scager and Boonstra (2011).

Participants

Participants of the study are the two teachers who delivered, and the 13 students who took the course. All 13 students who entered the Advanced Cell Biology course in 2010 are life science majors who had completed the prerequisites of an introductory biology course (100 level) and a textbook-based cell biology course (200 level). However, the students had virtually no laboratory experience, with the exception of a 200-hours science laboratory module.

Table 3.1
Description of the four phases of the course

Phase	Aims and Activities
Orientation Week 1 to 4	The aim of this phase was for students to master the current state of knowledge about the theme, learn the methods of researchers in the field, and to develop a critical approach. The teachers defined the theme; this year's theme was 'epigenetics'. In the first week, the course procedures (session 1) and the content (session 2) were introduced by the teachers. After these introductions, students received a series of primary articles and review articles on the theme, and were expected to present the findings in pairs during the first 3 weeks. This phase ended with the students' decision on the topic of research.
Research questions Week 5 to 8	In this phase, students focused on the research field covering their interest by searching primary literature and review papers. The findings were presented and discussed in class. From this phase on, the students were responsible for the organization of class meetings. At the end of the second phase, the students were supposed to have formulated their research topic (title of the projects) and their research questions, and to have produced a background paper.
Research methods Week 9 to 11	In the third phase, students needed to identify research methods and techniques, and plan and budget their projects. This required more reading, focusing on the methods sections, and finding expert help by contacting scientists in the field through visits to laboratories. Findings, ideas and proposals were discussed in class and in group meetings outside class hours.
Presenting and defending Week 12 to 14	In the final phase, students needed to relate the group projects to a whole, put it all on paper, and prepare for defending the program for the expert jury. The jury defense was formal; the students first presented their proposals in about 2 hours, and defended their proposals during a 1.5-hour questioning session by the jury members.

Data collection

A case study approach was used, with four methods of data collection, including the course design (provided to the students in written form), an interview with the teachers, course observations and student interviews, see table 3.2.

Table 3.2
Overview of research questions and methods

	Storylines	Student interview	Observa-tions	Teacher interview	Course materials
RQ1 How do students experience the balance between levels of challenge and ability, throughout the course?	x	x			
RQ2 How do students' perceived learning and motivation relate to the perceived challenge and ability levels?	x	x			
RQ3 What factors constitute challenge in this course?		x	x	x	x

*Course design***.** The course design as described in a paper (Wiegant, Scager and Boonstra, 2011) was used to provide contextual data and to inform the researchers. The documentation includes descriptions of the aims, learning activities, and results of the course, as well as evaluations by students, alumni, and jury members who had participated in the past.

Interview with the teachers. The teachers were interviewed by the first author before the course started. The interview was semi-structured, and questions concerned the aims of the course, the course design, the role of the teachers, and the student-led approach of the course.

Course observations. The course was observed by the first author, who attended 19 of 29 class meetings (about 40 hours in total), with field notes and audio-recordings taken. Field notes (reported in 58 pages) contained information on all interactions, reporting specifically who said what at what time. Contributions observed included the following: (a) a student presentation; (b) a teacher's instruction; (c) a question and answer session between students and teachers; (d) a discussion between students, and between students and teachers; (e) a suggestion, by either students or teachers. All contributions were noted in an abbreviated way, as were notable signals of non-verbal communication.

Interviews with students. All 13 students were interviewed individually in the week after the course ended, in a 45-minute semi-structured interview that was audio-taped. Students were informed that their individual statements would not be discussed with their teachers, and that their experiences and direct quotes would be used anonymously in subsequent reports. All students gave their consent.

Four themes were central to the interviews: students' learning, their motivation, their perceived ability, and perceived challenge. The perceived challenge was conceptualized as the level of difficulty experienced by students. The stimulating aspect of challenge, referred to in Webster's definition, was considered an effect of the difficulty-ability-balance. The perceived level of ability was conceptualized as the confidence students had in successfully completing the course. The passive role of the teachers in this course was an important feature, and it was

assumed to be a key factor that affected the students' experiences. Therefore, when students did not mention the role of the teachers spontaneously, they were asked, specifically, how they experienced the role of the teachers. At the end of the interview, students were asked about their future plans.

At the time they were interviewed, students had just finished the course having been praised highly for their work and presentations by the jury. The 'storyline method', which has been used successfully to retrospectively improve teacher's learning experiences (Henze, 2010; Meijer, 2011; Beijaard, 1999), was used to allow students to re-conceptualize their experiences over the period of the course. In the storyline method, students draw a line, visualizing their experiences along that timeline. In our interviews, students were asked to draw four lines, capturing the development, throughout the course, of their perceived learning, motivation, confidence (ability), and course difficulty (challenge). They were provided with pre-structured graphs showing diagrams, with the 14 weeks and 4 phases of the course set out on the horizontal axis, and a scale from -5 to +5 on the vertical axis. An example of such a graph is shown in Figure 3.2.

The students were asked to think aloud while drawing the lines. The thinking-aloud procedure gave an insight into how the various processes (learning, motivation, confidence, and course difficulty) developed, and allowed the students to question what affected these processes. In order to help students articulate potentially unnoticed factors, such as those that may be deemed too obvious to mention, they were also asked what could have caused their learning to decrease to levels below the lowest points on their graph.

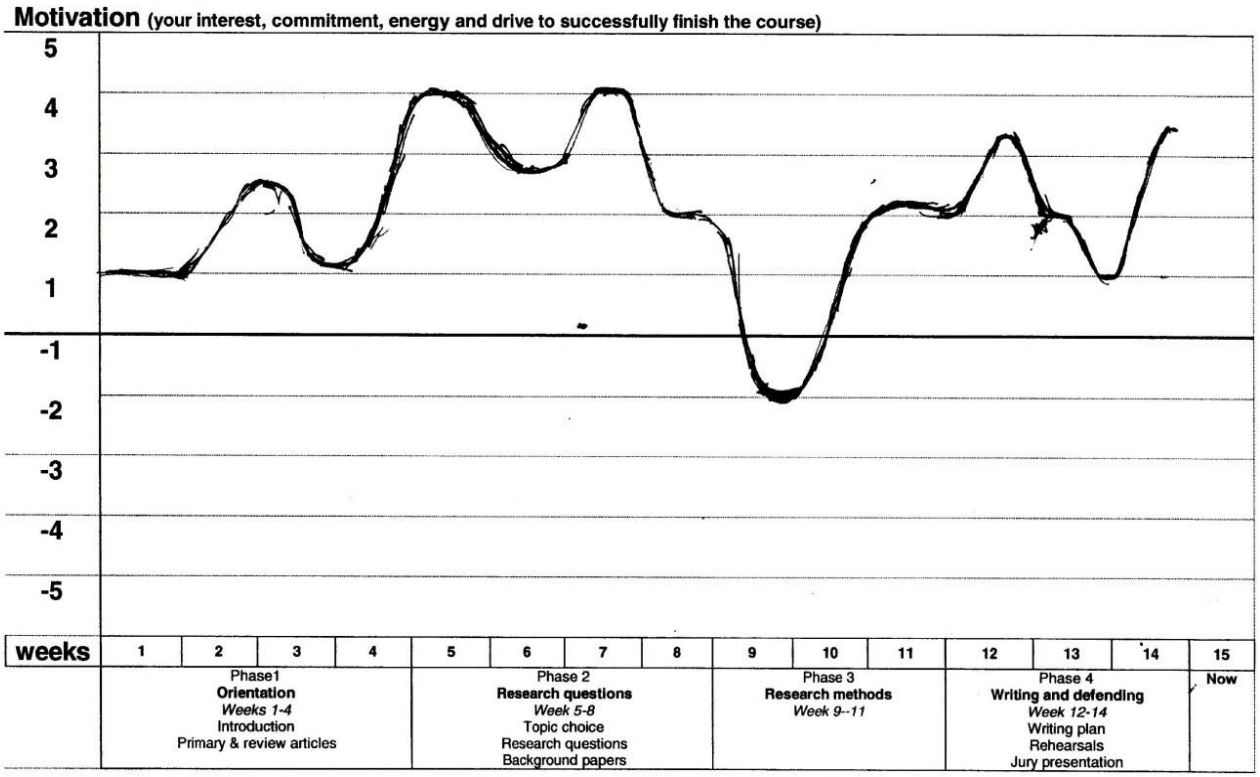

Figure 3.2. *Example of the 'storyline' graph of one of the students*

Data analysis

To answer the first two questions, about students' experience of the balance between challenge and ability levels, and the relation with learning and motivation, we analyzed the four storylines that were drawn and explained by all the students in the interviews. The storylines were not only used as means to activate students' memories, but also to generate quantitative results by averaging the weekly scores in the graphs of the 13 students on the scale from -5 to +5. The scale was regarded as an interval scale, with the 14 weeks as points of measure. The weekly scores from the 13 students were taken from their drawings in order to discover the mean lines, standard deviations, and correlations. The visualization of these mean lines provides the trends in students' experiences of learning, motivation, ability, and challenge in the four phases of the course. To complement the quantitative results, the transcribed interviews were read with care, and students' comments on the four pre-structured interview themes (learning, motivation, ability and challenge) were selected and coded, using NVivo (a qualitative data analysis computer software package).

To answer the third question, regarding the challenge-factors, we conducted an analysis in three steps. First, we examined the full transcriptions of the interviews with students, searching for elements in the learning environment that students referred to as difficult or challenging. Although their perception of course difficulty was the most explicit topic in the fourth interview theme, students also mentioned course difficulties they had experienced in earlier parts of the interview (while talking about learning, motivation, or confidence during the course). All the students' comments about the level of difficulty and challenge they experienced were selected and coded, using a sequence of open, axial and selective coding (Boeije, 2010). The final code list was a result of the iterative process of testing, discussing and adapting codes until the first and second author reached consensus on the codes to be used (Boeije, 2010). This process lead to three codes representing challenge factors: task complexity, lack of guidance, and high expectations. In a second step, student comments were categorized and subcoded, during which the first and second author compared and discussed their categorization until consensus was reached.

In step three these codes were used to explore the three other data sources to ground the realization of these three challenge factors in the learning environment: the field notes of the class observations, the transcribed interviews with the teachers, and the course description. These additional data sources were used to check whether the experiences of the students corresponded with the teachers' aims as described in the course description and/or reported in the interview, as well as with the observer's notes. Further, the field notes and recordings of the observations were used to determine the contributions of the teachers during class time.

3.3 Results

The results section opens with a description of the students' experiences of the processes of learning, motivation, ability and challenge throughout the four phases of the course (research question 1), and describes how the processes relate to each other (research question 2). Next,

we consider the third research question focusing on what constitutes challenge for students on this course. Typical students' quotes are included throughout to support and illustrate the findings.

Reported levels of challenge and ability, and their balance throughout the course

The students' challenge scores show, on average, an increase through phases 1 and 2 of the course, climbing to its peak in the third phase of the course, where students needed to find ways to answer their research questions (see Figure 3.3). The level of challenge decreased in the final week, when students had to present and defend their proposals before the jury. The standard deviation of the challenge scores was about 1.0 in all phases, except for this last week, where it rose to 1.8. Some students (N=3) experienced the presentation and defense for the jury to be very difficult, but most (N=10) reported a decrease in challenge in week 14 because they had practiced their presentations several times in the previous weeks. When looking at the individual graphs, all but one student located the highest levels of challenge somewhere during phase 2 (research questions) and 3 (research methods) of the course.

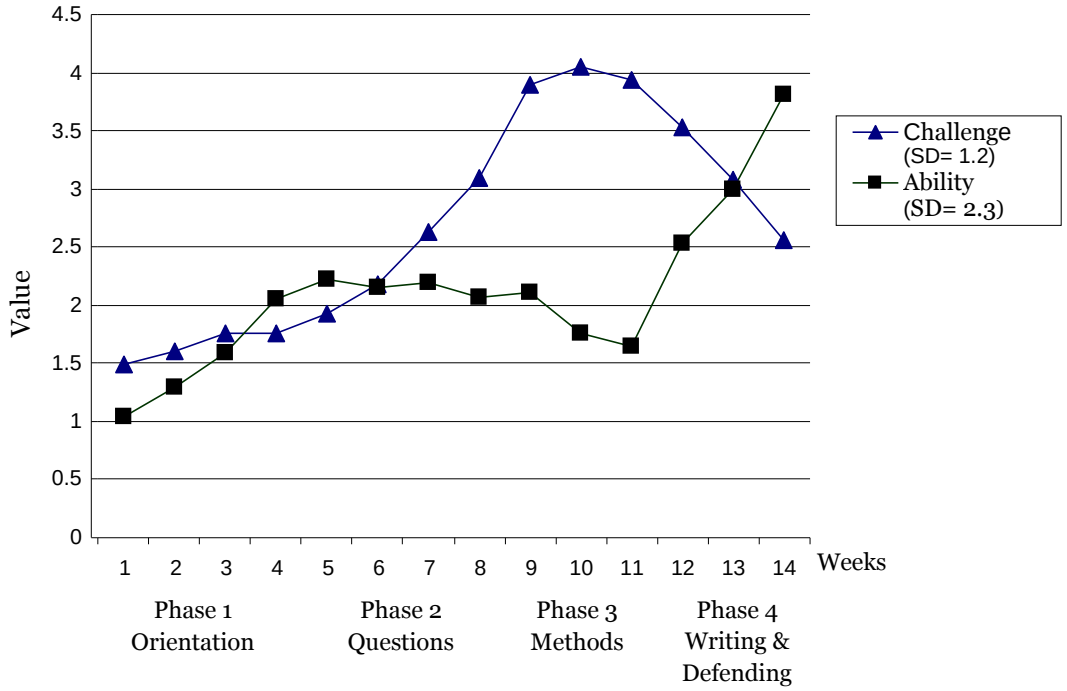

Figure 3.3. *Mean graphs for 'challenge' and 'ability' throughout the 14 weeks*

The mean scores of 'ability' showed large standard deviation (SD=2.3). Only two students reported stable confidence in their ability to successfully complete the course from the beginning to the end. All the other students drew erratic graphs for their perceived ability. The ability lines varied not only between students, but also for individual students between the four phases. Perceptions of ability vary for different reasons, depending both on personal and group differences. Students' perceived ability at the beginning of the course could vary

because of their perceived level of basic knowledge of cell biology, in comparison with their peers.

High challenge and low perceived ability: worry and frustration[4]

When we look again at Figure 3.3, the largest distance between students' challenge and ability lines are between week 8 and week 12. Although the students' ability graphs show significant variations, 11 of the 13 students' drawings show this gap between ability and challenge in this period. Emma's comment on this period illustrates students' experiences:

> I think that was the up and down part. At some points, for some questions, it was really easy to solve them, with other things it was way too hard, and we had no idea how to do it. We were emailing with people, and reading maybe 50 articles, and still we weren't sure how to do it. [...] At some point, we thought it was impossible.

Here, Emma describes a period of worry and frustration experienced by her group. The other two groups also reported that they had one or more frustrating periods during phase 2 or 3. Students' confidence dropped as a result of failing to formulate appropriate research questions, or ways to answer their questions, and could subsequently increase steeply when a good question or a feasible method was found.

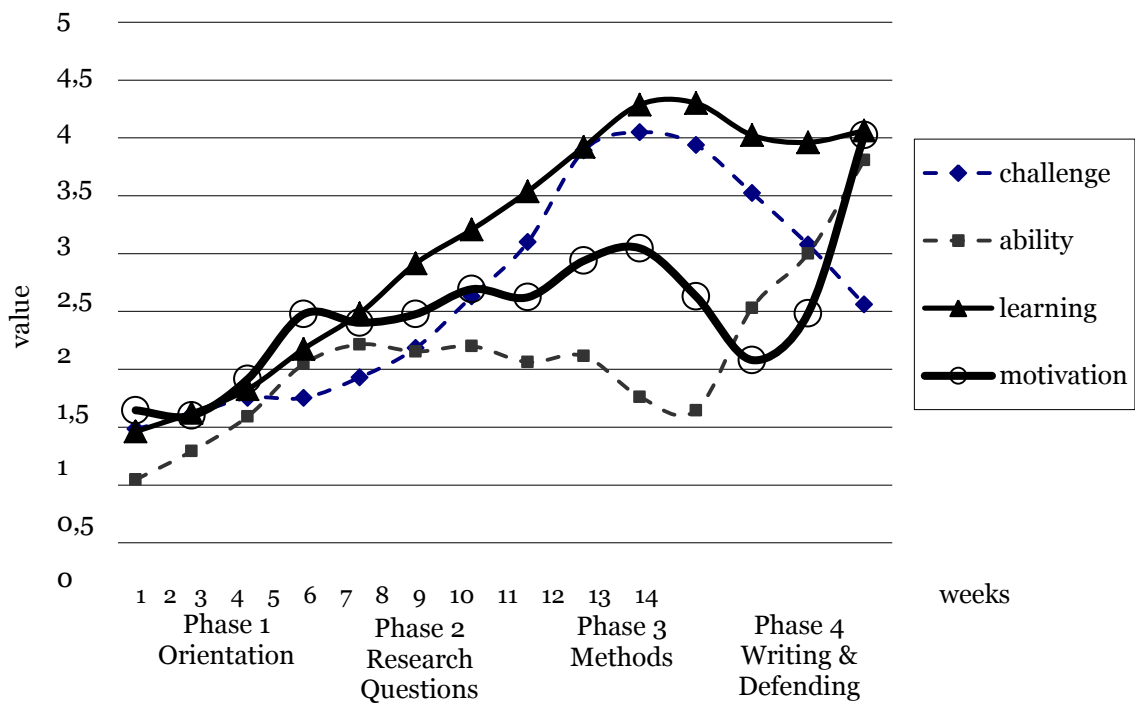

Figure 3.4. *Means of the graphs for 'learning', 'motivation', 'challenge', and 'ability' throughout the 14 weeks*

4 We use the word 'frustration' instead of Chickszentmyhalyi's 'anxiousness', as frustration better captures the experiences of the students in this study.

Students' learning increased steadily to a peak in the third phase. In this phase, where students had formulated their research questions and needed to find methods to be able to answer them, students (N=12) perceived that they learned the most. In the last phase, when students wrote down their research program, scores for learning dropped a little. Students' motivation started low as well in the orientation phase: they reported that they were not yet motivated because they did not yet have a clear picture of what was going to happen. Students were relatively passive during orientation, working much less hard than they would later on. For most students, their motivation increased in the second and third phase of the course, although progress was erratic. The standard deviation was quite high (1.6).

Figure 3.4 shows that students perceived that their learning gain, as well as their motivation highest during week 8 to 12, when the challenge was at its peak, while their confidence in their ability was at its lowest point. To check whether this pattern reflected the quantitative data retrieved from the graphs, correlations were calculated. A correlation of .51 was found between learning and challenge ($p<0.01$), but not for motivation and challenge or motivation and learning. This correlation pattern is also reflected in students' descriptions, for example Susan's:

> During research questions, we were researching, but we're not interpreting as much as we should have been. But then [during research methods] it kicked in. Every single thing that I was doing was completely involving all that I knew. That's why I was putting 5's here. I was engaging all my abilities that I already had here, and the ones that I needed to develop. [...] We were not able to grow any more when we were doing the research methods.

Susan, like many other students, assessed her learning in this period with a 5 (on a 5 point scale), and she describes an intense learning experience, resembling the flow experience Csíkszentmihalyi (1975) describes. The learning curve of students remained high during these weeks; their motivation, however, fluctuated in this period, and even drops below zero for some students, including Susan:

> Oh, and I think somewhere here during the methods, it should go below zero a little bit. This was when nothing was going right. We were looking at methods and we were not finding what we were looking for, and we were again rephrasing questions, we weren't getting any criticism from other groups, with peer review, we wouldn't get any feedback, and that was very frustrating, and when there is nothing, nothing coming your way, that is extremely frustrating. It was entirely unmotivating. Oh, but this [end of phase 3] was high again [drawing].

Obviously, during this period Susan was not experiencing flow in the sense of enjoyment. Although she was fully immersed in the task during this period, she lacks the positive feelings that accompany flow. This pattern of increased drive to work and temporary loss of enjoyment was found in most of the other students' stories and storylines as well (N=10) during this stressful period. These students described feelings in terms of "anger", "frustration", and perceived the task as "impossible" in that period. Nonetheless, students reported an increase of time investment and learning gain during this period.

According to the students, their motivation in these low periods was affected by the group interactions. In two groups, social support and the encouragement of peers helped them through the dips, whereas in one group the tensions increased because of problems between team members.

Students not only learned a lot on this course, they also performed very well. According to the jury, the students produced excellent work. All three of their research proposals were relevant, innovative, very well-motivated and presented, and grant-worthy, except for the feasibility aspect: according to the jury, the proposals were still too ambitious.

Factors constituting challenge

Three factors could be distilled from the aspects students mentioned were constituting the challenge they experienced in this course: complexity, lack of teacher direction, and high expectations of both teachers and students. In Table 3.3 below, the number of utterances and students are shown.

Table 3.3
Challenge codes, sub codes, number of utterances, and students

Codes	Sub codes	Number of utterances	number of students
Complexity	Novelty of material	14	13
	Interplay between standards	13	5
	Dynamic of the process	13	4
Lack of teacher direction	Lack of structure	13	10
	Lack of feedback	25	13
High expectations	Teacher expectations	6	6
	Student expectations	8	4

Other challenge factors that students mentioned were as follows: the group work (N=2), the professional level of the task (N=2), the assessment by external experts (N=1), and the restricted time frame (N=1).

In the following, we describe how students experienced these three factors, and compare their views with the findings of the analysis of course materials, the interview with the teachers, and the observation data.

Complexity

The task of writing a research program was complex for the students in several ways: the novelty of understanding and evaluating primary papers, the interplay between the various standards of a research proposal, and the dynamic of the process, generated by the need for groups to collaborate. The first complexity factor was the novelty of reading and judging primary articles, especially the methods sections, as described by Taniah:

> Research methods, that's a whole different story; that was hard, very very hard. Sometimes you read these techniques, and then you think, how did they [the authors] come from this step to the next, they [the authors] don't explain, and you go to the supplementary information, and then that's just not English.

Taniah refers to the 'leap' students needed to make on this course, from the relatively simple task of learning from books on previous courses to determining what information was needed, searching for it themselves, and judging its value. The ways in which the students described the novelty of understanding and critically assessing primary articles can also be recognized in the course design and the observations. Firstly, the reading of primary articles is one of the planned learning goals of this course. As explained in the course description (Wiegant, Scager, & Boonstra, 2011), students on this course need to bridge the gap between reading textbooks and research articles. From textbooks, students receive information that is already selected, organized, interpreted, and explained by the authors of the book. However, primary articles on cell biology explore the subject in much greater depth and are quite technical and jargon-laden, making them difficult to read and understand.

Unlike textbooks, the validity of knowledge presented in primary articles lacks consensus amongst scientists, and this requires the students to read critically, detecting inconsistencies in the current knowledge (which is new to them) and use this for their own research questions. The emphasis on critical reading was visible during the observations. In the first phase of the course, teachers stimulated critical reading by modeling the behavior, asking detailed questions about the methods researchers used, or the conclusions they drew, and at the end asking students how they assessed the paper.

A second aspect of the complexity mentioned by the students was the interplay of the various standards of a research proposal, as Samuel explains:

> The problem-solving aspect is something I really, really like. [...] You have to think about a lot of practical things, it has to be feasible. Think of a lot of other possibilities, like 'what if this is the case, how would that affect my methods, or questions?'

This aspect of complexity was reflected in both the course description and the observations. According to the course description, the knowledge from the literature needed to be translated into questions that have not yet been answered but for which the solution can be obtained in the laboratory, using appropriate methods and techniques. This interplay between demands demarcates the main assessment criteria for research funding agencies in general, as well as for the students' proposal. This aspect of complexity could also be recognized in the course observations. During class meetings, the teachers would frequently question the interplay of the demands of a research proposal. The novelty demand, for example, was subtly inserted into the group discussion about a possible research question, when the teacher would ask: "Did someone check how many hits you get on that mark in Pubmed?".

A third aspect of complexity for students was the need to adjust their plans to changes in the environment during the process. Groups needed to change their questions and methods on several occasions in order to keep the connection between the 3 proposals or because new

research studies were found which partly answered one of their questions. Field notes of the observations, as well as the course description, reveal that the teachers deliberately planned the interconnectedness of the 3 groups as part of the course design. Teachers told students on several occasions that researchers in the field of science work in groups on interconnected research programs. The interdependency between groups was reinforced by giving students responsibility for the whole program and not only for their own group project.

Lack of teacher direction

The course was designed to be largely student-led, which students experienced generally as a lack of direction by their teachers. This lack of direction comprised two aspects: a lack of structure and a lack of feedback.

Lack of structure. As students explained during the interviews, they are used to teachers taking the lead and structuring the course by setting the deadlines, examinations, and organizing class meetings. On this course, students were supposed to take over this role, which they found hard, specifically in the first phase of the course. Jesse explained the way in which students perceived this problem:

> I think it was the lack of idea of what we should do. It was really not directed, we weren't that guided, how to do it. And maybe because we didn't really put strict deadlines, we constantly moved them. So it was the lack of structure (...) the lack of teacher control. They weren't organizing the classes, it was all on us, we had to organize, pick the deadlines. I think the feeling of responsibility was the hardest thing.

Despite this perception, the course description and observations revealed some general guidelines. The teachers set out the assignment, the topic of epigenetics, the group work requirement, the date of the jury presentation, and the structure of the first few weeks, in which students were supposed to familiarize themselves with the field by presenting research papers in pairs. Furthermore, students received a list of tips from the previous group, and the final products of six of the previous classes. Within these parameters, students had to make their own decisions about the content of their research project, the organization of processes, and the class meetings. During the observations, the teachers' intention that the process should be student-led was clearly visible. Teachers kept silent for most of the class time, occasionally posing critical questions. After this first week, the teachers literally 'handed over the chalk' to the program leader, took their place at the side of the class, and stayed there throughout the rest of the course. Even in the observed class meeting where the teachers were most active, they did not speak for more than 5% of the meeting. Furthermore, we observed that teachers never intervened in decisions concerning the process or the content of the task. They offered suggestions now and then, such as "why don't you talk about this some more in the bar or some other non-formal setting?", but that was one of the most direct suggestions the teachers provided.

Lack of feedback. One of the most difficult aspects of the course for students was the lack of feedback (N=13). Especially in the first three phases of the course, students said that they had been ”frustrated”, ”grumpy”, and ”complaining” about the lack of feedback from their teachers. In the last phase, students did receive feedback on both their draft proposal and their rehearsal of the jury presentation. The lack of feedback was expressed in various ways.

Firstly, students mentioned their need for the teachers’ expert judgment, as Faizal explained: ”It’s just the thing that you need experience there. Our teachers would know what we were doing wrong, but sometimes we waited for their comments, and we figured out ourselves that we were wrong”. Secondly, there was the need for acknowledgement, as Jesse said: “We were still learning to be independent, and then you sometimes wanted a bit of acknowledgement”. Thirdly, students felt that they would have benefited from some confirmation, as Alex noted: “I think here, with the questions, it would have helped if they would have said, ‘’ok, go for this’”. These utterances reveal a need for teachers’ judgment on students’ performance, and some confirmation or clues about the direction of their next step. The teachers did however offer critical questions throughout the process, which was understood as feedback by some students, as Sarah explained during the interview: “In asking those questions they also let us know in a small way that it was time to start with the next step, with the methods for example”.

The analysis of the observations, as well as the interview with the teachers, concurred with the experiences of the students. In the interview, the teachers explained their approach: “if we do give feedback, students are inclined to sit back and rest, and we don’t want that. Moreover, we think that it is important that it will be their own product, something that they can be proud of”. Teachers explained they were aware of the frustration students experienced, as they explained during the interview. Their seven years of experience with this course taught them that the lack of feedback would eventually challenge students to develop the critical attitude of a researcher and that was one of the main objectives of the course. Observations of class sessions also correspond with the experiences of the students. During the class meetings, the teachers, indeed, did not provide explicit feedback in terms of giving information on the quality of the students’ performance and progression, nor did they provide confirmation or reassurance. We observed that the teachers did however facilitate students to assess their own performance, by handing students some examples of products of previous years’ students, by asking critical questions, and by initiating a peer-review system.

Although students lacked the feedback from their teachers in the first part of the course, in hindsight all of them understood the approach of the teachers, as Taniah explained:

> I think feedback would have helped. But then again, I think it was a good thing that they didn’t. We didn’t need that reassurance from people; we should be able to know if we were doing good or not.

Students conveyed ambiguous feelings about the lack of teacher guidance. At the time, the quietness of the teachers was difficult for the students, but after the success experienced

during the positive assessment by the jury, feelings of pride and satisfaction replaced earlier frustrations.

High expectations

A third factor constituting the challenge, according to the students, was the high expectation of their teachers as well as of the students themselves. The teachers conveyed their expectations by mentioning the achievements of previous groups, as Sarah explains: "The fact that the previous years did good and the teachers told us that they all did good and that none of them did bad during the final presentation, that was putting the standards high. You need to keep up with that…".

The course description and the interview with the teachers were not entirely in line with the students' experiences at this point. Teachers assumed that the students' performance goals would be raised primarily by the expert jury assessment, and not so much by their own expectations. For Alex, however, the jury assessment primarily added to the feeling that they had to live up to the expectations of the teachers: "[…] You know that the level needs to be high. He [the teacher] emphasized that we needed to deliver high quality, otherwise they would look bad as well". Students did not only attribute the high expectations to their teachers: the majority also wanted to achieve for themselves. Firstly, students, for example Marieke, felt that having high achievement goals is a part of the UCU culture: "the fact that we really wanted to do well [Why?]. I think it's just a trait a lot of people at UC have". Other students, like Gizelle, related her aim to do well to the student-driven nature of the University: "I believe that it works very well that it is all up to us, our own initiative: you are not forced to work, but you do it anyway. We have put a lot of time into the project, not because it is compulsory, but because you want to". Here, Gizelle also related two of the challenging aspects of the course to each other: the autonomy (lack of teacher direction) and her high expectations. The autonomy apparently stimulated her to focus on her own goals and standards, working hard not because she was told to, but to satisfy her own need for mastery of the writing of a research proposal.

3.4 Conclusions and discussion

Three questions about challenging gifted students were central in this study. Firstly, using Csíkszentmihalyi's (1975) flow-model, we explored how students experienced the balance between perceived challenge and ability, and secondly, how this balance related to learning and motivation. Thirdly, we examined what characteristics of the learning environment challenged these high-ability students. We will now discuss the findings, starting with the first two questions as their findings are related.

Students' reports of the processes of their learning, motivation, challenge, and ability showed an interesting pattern. Students' perceived learning was at its peak in the period in which the challenge level most exceeded the ability level. According to Csikszentmihalyi (1975), if challenge exceeds perceived skills, people feel worried and anxious, which are

negative emotions that decrease feelings of flow and intrinsic motivation. Students in this case study felt that there was a significant difference between their perceived ability level and the level of challenge, especially in the third phase of the course, and indeed reported feelings of worry and frustration. Students' motivation graphs fluctuated heavily during this period. A closer inspection of students' stories about their motivation however revealed that, although their enjoyment decreased during these periods of decline in motivation (as shown in their graphs), their drive did not weaken, and their efforts even increased. Students apparently interpreted the concept of motivation as a combination of drive, which they emphasized in their stories, and enjoyment, which guided their graphs. Students' learning graphs were at their peak during this period of fluctuating motivation. This finding is unexpected, given that negative emotions have been understood to impede intrinsic motivation (Csikszentmihalyi, 1975; Larson, 1988; Meyer and Turner, 2006). An explanation for students' increased efforts, despite a lack of enjoyment, could be found in the support that was experienced from their peer group. Firstly, students in their project groups cooperated in a face-to-face mode, reading, writing, emailing and discussing together. In the busiest phase of the course, students kept working most evenings, and all through the weekend. Face-to-face settings for group work have been found to support intrinsic motivation significantly better than virtual group work (Järvela, Järvenoja, & Veermans, 2008). Secondly, the groups could have developed a sense collective efficacy. Collective efficacy -a group's shared belief in their collective power to produce desired results- affects group motivational efforts, and engenders persistence in the face of difficulties (Bandura, 2000; Goncalo, Polman, & Maslach, 2010). Thirdly, several students reported that the social support of their team members helped them during the moments of crisis they experienced during the difficult phase in the course. We accordingly assume that social support can prevent or overcome situations of over-challenge and the accompanying feelings of worry and frustration. This assumption concurs with findings of Bakker and Demerouti (2006) who studied factors that cause or prevent burnout in work settings and found social support to be one of the most important factors that balance out high job demands. Thus, although the enjoyment associated with intrinsic motivation and flow was not there in this difficult phase of the course, the drive, and persistence were maintained, generated by the social support of the group.

The third research question focused on what challenged students to work so hard and perform so well on this course. The complexity of the task, the lack of guidance, and high expectations were the three factors that constituted the challenge for the students. The complexity of the task is conceptually closest to the 'challenge' axis in Csikszentmihalyi's (1975) flow-model. On this course, students described the task complexity in terms of novelty, interplay of standards, and dynamism.

The lack of teacher direction (both structuring and feedback) was the second factor we found. The lack of structure refers to the student-led character of the course. In the first half of the course, students experienced a 'lack of guidance'. They were supposed to regulate their own learning and research process, which confused them at the beginning. The way this lack

of direction was experienced, however, evolved during the course into more positive feelings of independence, responsibility, and ownership of their project. Research on self-regulating learning shows that students' perception of internal control results in more engagement, higher achievement, and higher self-esteem, whereas external control results in lower achievement and passivity (Pintrich & Zusho, 2007), and this research is borne out by our results. Although from phase 2 onwards the course was led by the students, they did not convey a perception of control: students felt more 'lost' than 'in control' in that period. This indicates that student autonomy did not automatically or immediately lead to feelings of internal control and its positive outcomes; a period of adaptation was needed.

The way in which a lack of teacher feedback was perceived to constitute part of the challenge was an interesting finding. In educational literature, feedback is described as one of the most powerful elements of an effective learning environment (Black & William, 1998; Hattie, 2009; Sadler, 1989). In our study, students did express the need for feedback from their teachers, but at the same time, they proved to be able to perform exceptionally well without it. Apparently, students learned to judge the quality of their work themselves, which was exactly what the teachers were aiming for. Withholding feedback is advocated by several scholars (Nicol & Macfarlane-Dick, 2006; Sadler, 1989; Yorke, 2003), who argue that teacher-transmitted feedback keeps students dependent on their teachers and interferes with self-regulated learning. According to Sadler (1989, p. 121) "students have to be able to judge the quality of what they are producing and be able to regulate what they are doing during the doing of it". In this view, students need to be empowered to judge the quality of their work, which is precisely what the teachers of this course did. Students were provided with exemplars of good work and critical questions, and they were stimulated to critically assess each other's work. Teachers did not impose their views, but they were available if students needed them. The students, in hindsight, valued the restrained attitude of their teachers because it increased their sense of ownership and pride. At the time, however, the lack of teacher guidance was not beneficial for students' enjoyment. The students, especially in the first half of the course, had felt lost and were not very happy with their teachers, saying they had been feeling "frustrated", "grumpy", and were "complaining" about the lack of feedback and structure. The teachers, however, held back from intervening, despite the obvious need from students for feedback. This allowed students to adjust to their independence and take the responsibility for their project. These negative feelings might be expected to have a detrimental effect on the interpersonal relation between students and teachers, which is found to be an important element of effective teaching (Hattie, 2009; Wubbels & Brekelmans, 2005). However, our study shows that the students' appreciation of their teachers was positive, which could be explained by the fact that, in hindsight, they understood that the teachers' approach was an empowering one.

The high expectations the students experienced both from their teachers as well as from themselves was the third factor we found. High, but appropriate, teacher expectations have been found to be effective for all students (Hattie, 2009). In our study, students felt strongly

that their teachers were critical and expected high standards in their work, and they wanted to meet these expectations. According to Phillips and Lindsay (2006), high expectations have the risk of generating stress and anxiety, an emotional state interfering with motivation. On this course, the high expectations contributed to the challenge and associating feelings of worry and frustration.

In conclusion, the combination of the complex task, the lack of teacher guidance, and the high expectations increased students' worry and insecurity during the course. In hindsight however, these factors made up the challenge, and increased their efforts and perceived learning. The results of this study allow us to propose a different view of Csikszentmihalyi's (1975) flow-model, one that is specifically interesting in the context of teaching high-ability students. The processes students described led to the notion that the balance between challenge and ability was lacking during several weeks; students were over-challenged in that period, during the course, experiencing various emotions in the process. Positive emotions were experienced when students made progress, which concurs with Pekrun's (2006) control-value theory of intrinsic motivation. However, maximum effort and (perceived) learning were experienced when the perceived challenge exceeded the level of ability the most, and students felt worried and frustrated. These findings indicate that the balance between ability and challenge is vital for enjoyment, but not necessarily for effort, persistence, and learning. Specifically, the concept of challenge, a commonly used but seldom operationalized concept in the field of giftedness, has been central to our study. We recommend further research to discover whether the three-factor model of challenge (complexity, lack of guidance and high expectations) applies to challenging university courses in other disciplines.

4. How to persuade honors students to go the extra mile; creating a challenging learning environment[5]

Abstract

Although universities aim to challenge honors students to develop their talents to the full, the question of how to create such challenges remains unanswered. In this study we examined six different honors courses to discover which factors in the learning environment specifically designed for high-ability students challenged these students and determined how such challenges were established. Perceived challenges (factors and intensity) in the different courses were identified from focus group interviews with students. The interviews were guided by a storyline method which retrospectively discussed students' experiences of course activities. In addition, course materials and audio-recorded classes were analyzed to describe how challenge was established. Results revealed three challenge factors: autonomy, complexity, and teacher expectations, similar to the ones found in a previous, single-case study. Results showed that the greatest challenge was experienced when the three factors were simultaneously experienced in the learning environment. These three challenge factors were manifested in various and sometimes very subtle ways through course design (e.g. open structure of assignments) and class interactions (e.g. particular ways of questioning).

[5] This chapter is based on: Scager, K., Akkerman, S. F., Pilot, A., & Wubbels, T. (2013). How to persuade honours students to go the extra mile; Creating a challenging learning environment. *High Ability Studies,* DOI: 10.1080/13598139.2013.841092.

4.1 Introduction

"Most of his students are willing to go an extra mile for him. There are only a few teachers who can get that out of students" (an honors student).

Many universities within and outside Europe have recently developed honors programs for their high-ability students (Long & Mullins, 2012). High-ability students are considered to be students who are talented, driven and persistent (Friedman-Nimz & Skyba, 2009; Renninger, 2009; Simonton, 1994; Trost, 2000), as well as committed to work hard (Ericsson, Roring, & Nandagopal, 2007; Howe, 1999; Walberg, Williams, & Zeiser, 2003). In focusing on these students, honors programs are explicitly intended to provide opportunities for them to develop their talents to the full, enabling them to make significant contributions to science and society.

Despite the increase in availability of honors programs across European universities, empirical studies on effective learning environments for high-ability students at university level are rare. Most research on high-ability students has been conducted in primary and secondary education and has focused on the curriculum level, specifically the effects of grouping, acceleration, and enrichment (Ziegler & Phillipson, 2012), when high-ability students are in a mixed-ability classroom. Therefore, these results are of limited value for honors education in universities where students are grouped more homogeneously.

Recognizing that practice in honors programs is ahead of research on the topic of creating effective learning environments for high-ability students, Scager, Akkerman, Pilot, & Wubbels (2012) analyzed the characteristics of a motivating honors course with high learning results. They found the concept challenge to be crucial in explaining the positive results of that course. The present study aimed to examine what constitutes challenge or the lack of challenge in a variety of honors courses.

Challenge in gifted education

Research on gifted education has generated suggestions for educational practice that point consistently to the importance of challenge (Eddles-Hirsch, Vialle, Rogers, & McCormick, 2010; Gallagher, Harradine, & Coleman, 1997; Heilbronner, Connel, & Reis, 2010; Marra & Palmer, 2004; Noble & Childers, 2008; Rogers, 2007). What challenges high-ability students is often not explicitly mentioned but implicitly suggestions for creating challenge indicate two directions. One is the curricular objectives, for example offering complex problems, promoting higher-level thinking skills rather than the mastery of facts, and the development of creative problem-solving (Gallagher et al., 1997; Kanevsky & Keighly, 2003; Rogers, 2007; Shavinina, 2009; Sternberg & Lubart, 1993; Tassel-Baska & MacFarlane, 2009). Whether providing high-ability students with opportunities for self-regulation and independence is useful is debated (Freeman, 1990; Marra & Palmer, 2004; Sontag, Stoeger, & Harder, 2012). The second direction points to various educational models appropriate for high-ability students, such as discovery or inquiry learning, problem-based learning, or learning in authentic or contextualized settings (Gruber & Mandl,

2000; Hertberg-Davis & Callahan, 2008; Heller, 1999; Mönks & Katzko, 2005; Shore & Kanevsky, 1993; Snow & Swanson, 1992).

The concept of challenge

Though challenge seems to underpin many suggestions for gifted education, only few of the aforementioned studies rely on empirical research to support their claims regarding challenging learning environments. Hence, it is relevant to conceptualize and operationalize challenge in order to strengthen practice and theory more systematically in university gifted education. For our conceptualization we draw from psychological theories that include the notion of challenge, such as the control-value theory of motivation (Pekrun, 2006), the self-determination theory of motivation (Deci & Ryan, 1985), and the flow theory (Csikszentmyhalyi, 1975). These theories express the inherent relationship between the levels of challenge and ability, indicating that studies of challenge also have to include ability. In addition, these theories stress the need for balance, with a challenge level slightly above the ability level for positive effects on motivation and learning. Csikszentmyhalyi (1975) assumes that an appropriate balance between challenge and ability creates a state of flow; accordingly, tasks should increase in their level of difficulty as the student's level of ability improves. This postulate of optimal challenge in flow theory is consistent with the need for competence in the self-determination theory (Deci & Ryan, 2000): To create a feeling of competence, tasks need to be optimally challenging, implying a balance between challenge and abilities. In the control-value theory of motivation (Pekrun, 2006) it is posited that a lack of balance, caused by demands that are too high or too low, causes negative emotions that can impair intrinsic motivation.

The hypothesis in literature on gifted education seems to be that the level of challenge needs to be increased for these students to match their higher ability (Heller, 1999; Lubinski & Benbow, 2000; Snow & Swanson, 1992). High-ability students, however, also differ qualitatively from average students in being quicker thinkers, more flexible in their use of strategies, having better memories, and preferring complexity (Freeman, 1990; Shore & Kanevsky, 1993; Wallace, 2000). Given these abilities, challenging high-ability students is not simply a matter of increasing the level, but should also take into account these qualitative differences.

In line with the flow model, our previous case study (Scager et al., 2012) showed that high-ability students reported the highest motivation when (perceived) levels of challenge and ability were in balance. Yet, in contrast with the flow model, we found that students learned the most and worked the hardest when the level of perceived challenge exceeded the level of ability beyond the comfort zone, leading to instances of worry and frustration. The (temporary) feelings of worry and frustration decreased students' enjoyment, but not their learning and persistence, suggesting that the optimal level of challenge for these students exceeds their ability. Three factors were found to induce the challenge in that course, respectively high levels of complexity, student autonomy, and teacher expectations. As the previous case study concerned a rather idiosyncratic course, the present study considers six honors courses of varying levels, in order to explore which factors

constitute challenge or a lack of it and to specify how challenge is established in the learning environments. These aims lead us to the following research questions:

1) Which factors in the learning environment constitute the (lack of) challenge in six honors courses varying in discipline, level, and teaching methods?
2) How are these factors manifested in the design and implementation of the courses?

4.2 Methods

Context and courses

The six courses were taught at the University College Utrecht (UCU), an international liberal arts and sciences honors college. UCU students are carefully selected for their academic excellence (minimum GPA of 3.0), curiosity, and motivation. Students complete four courses per semester (15 weeks), with a workload of about 200 hours (7.5 European Credits). Courses are given on three levels, the 100, 200 and 300 levels. As a rule, a 100-level course in the same disciplinary track is required for entry to a 200-level course, and a 200-level course is required for the 300-level courses.

Sample

To establish a wide range of learning environments, courses were selected from all three undergraduate years, representing the fields of science, social science, and humanities, and providing a wide range of pedagogical approaches. Within this scope, the UCU director of education was asked to select six teachers and ascertain their willingness to participate. All six teachers agreed.

Data

Data were collected in each course by class observations, gathering the course materials, and focus group interviews. Table 4.1 provides an overview of the data collection.

Observations and course materials

For each course the first author observed and audio-taped four to six classes throughout the 15 weeks across various stages of the courses. Field notes included the time (episodes of instructions and discussions in minutes, waiting time (in seconds after the teacher had asked a question), the type of interaction of the episode (lecturing, discussion, teacher questions, answers and uptakes, and student questions or comments), and observer memos. Audio-recordings were used to transcribe teacher questions and responses to student answers, and to summarize episodes of lecturing. Course materials (outlines, readings, assignments, exams) were also collected. Observations and course materials were used, first to familiarize ourselves with the courses and provide a frame of reference, and second, to provide a more rounded picture of the phenomena studied.

Table 4.1.
Course list and data

	Observations		Course materials	Focus group interview
	classes	hours		
100 level Science	5	10	Course outline Readings Assignments Exams	8 out of 26 students
100 level Humanities	4	8	Course outline Readings Assignments Exams	5 out of 24 students
100 level Social Sciences	4	8	Readings Assignments Exams	5 out of 24 students
200 level Social Science	4	8	Course outline Readings Assignments Exams	5 out of 24 students
300 level Science	5	10	Course outline Readings Assignments Exams	5 out of 12 students
300 level Social Science	6	12	Course outline Readings Assignments	4 out of 21 students

Focus groups

Focus group interviews were used to determine students' perceived challenging factors in the learning environment. The presence of these factors could increase and absence could decrease the challenge. For each course focus group interviews of about 60 minutes were held, with four to eight students, ensuring that the sample size was large enough to capture a range of voices and small enough for everyone to have their voices represented (Langford, Schoenfeld, & Izzo, 2002). Students were selected by the moderator after consulting the teacher on the basis of differentiation in gender, age, and activity in class discussions, allowing variation and representativeness. All students gave their consent to audio-taping of the focus group interview and the use of their quotes, albeit with altered names. Students already knew the moderator (first author) as an observer in class sessions. The role of the moderator was to stimulate and moderate discussion, ensuring depth as well as diversity in perspectives on the course. An assistant took notes to identify the participants and summarize the content of the discussions.

The interviews were semi-structured by using as a start the “storyline method” that has been used successfully to discuss teachers' learning experiences retrospectively (Beijaard, van Driel, & Verloop, 1999; Henze, van Driel, & Verloop, 2009). In this method respondents drew lines representing the level of challenge across time throughout the 15 weeks on a scale of -5 to +5 (y-axis) on transparencies on which weeks and activities of the course were printed on the x-axis (See Figure 1). Students were asked to draw their challenge storylines, and given time to think back without being influenced by their peers. Then the interview started with elaboration on the variety of experiences. As a trigger for elaboration, the storylines were displayed simultaneously using the overhead projector. Students were asked to explain the changes in their lines (i.e. “In week four

your lines go up… what happened?”). When the graphs showed conflicting patterns, students were invited to share their different experiences, and when their lines corresponded, attention was paid to stimulate broad participation, to prevent “group think” in an early stage of the discussion. In the second part of the focus group interviews, students drew storylines capturing their perceived ability. These lines were collected as data, but were considered too personal to be exposed and were only discussed briefly.

Data analysis

The average difference between the perceived level of challenge and the perceived level of ability was determined by averaging the weekly individual student scores of the storylines for each course (see Figure 4.1). Perceived challenge and ability levels varied considerably within and across courses. In most courses (five out of six), students experienced one or more periods where the level of challenge was clearly above or below the level of ability, indicating respectively a situation of perceived challenge or perceived lack of it.

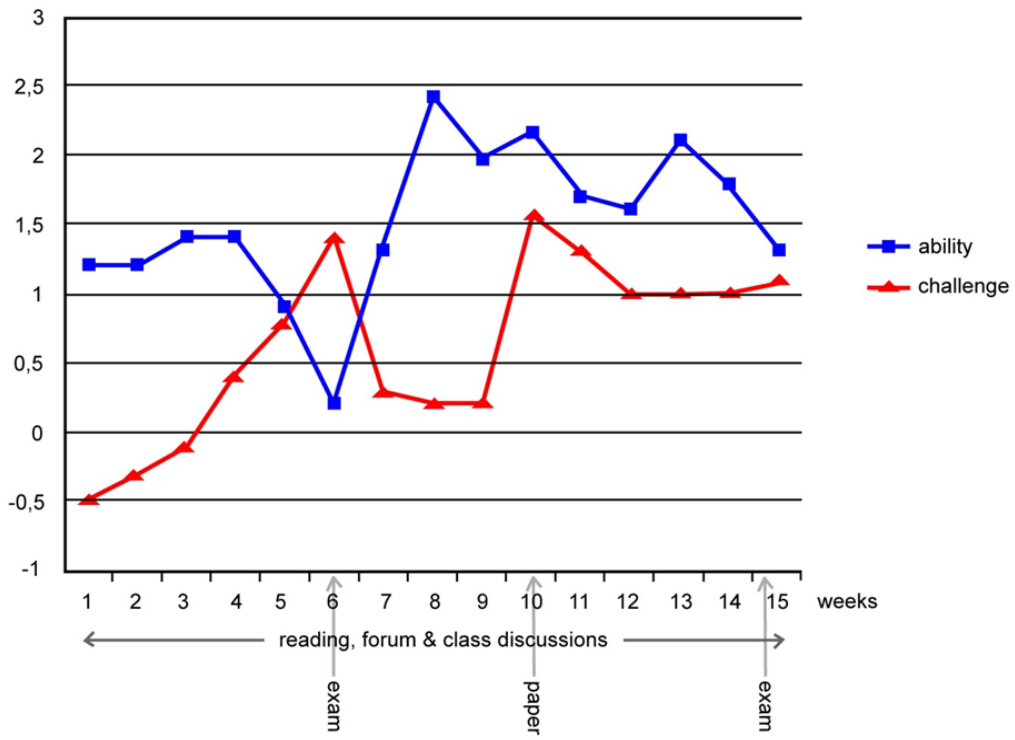

Figure 4.1. *Example of averaged storylines of students’ perceived challenge and ability levels.*

To find factors constituting the challenge in these courses (research question 1) data were analyzed by the first author in two partially overlapping stages. Stage one comprised reading and re-reading the transcripts in order to identify text units relevant to the subject of challenge. Given the aim of the focus group interviews, this meant removing small talk. The unit of analysis for the remaining data was the storyline and discussion of a specific particular course activity. Stage two comprised identifying and coding themes with the three challenge factors (autonomy, complexity, and teacher expectations) we found in a previous case study (Scager et al., 2012) as sensitising concepts

(Bowen, 2006). The three categories appeared sufficient to describe the challenge factors.[i] A second researcher coded 33% of the data which yielded a kappa coefficient of 0.80 across all categories, which was considered an adequate level of inter-rater reliability.

To examine the way in which the three remaining factors were manifested in the design and execution of the course (research question 2), the text units for each challenge factor were sub-coded into indicators of encouragement in the learning environment, respectively autonomy, complexity and teacher expectations (Table 4.2). Student references to their sense of control over their learning, as opposed to teacher control and extrinsic incentives were categorized under the concept autonomy, thus following Ryan and Deci's (2006) 'self-governance' interpretation of autonomy. For sub-coding of autonomy, cues in the focus group transcripts included 'choose', 'freedom', 'responsibility', control', 'decide', 'allowed', and 'requirements'. Cues for sub-coding of complexity included 'difficult', 'complex', 'ambiguous', 'level of thinking', 'easy', and 'straightforward'. For sub-coding teacher expectations words such as 'expectations', 'standards', and 'demanding' were used as direct cues, and phrases such as 'you want to go the extra mile for this teacher', or 'you did not need to prepare' were used as indirect cues. The data were coded by a second coder employing the indicator categories, which resulted in a kappa of 0.82.

Table 4.2
Three challenge factors and their indicators

Challenge factors	Indicators
Autonomy	1. Choice in assignments
	2. Control over achievement
	3. Student responsibility
	4. Non-directive feedback
	5. Openness to student initiative
Complexity	1. Novelty of the learning material
	2. Requirement for critical thinking
Teacher expectations	1. Teachers' reputation
	2. Announcement of challenge
	3. Teacher as co-learner
	4. The need for active participation
	5. The need to prepare for class

Guided by the indicators developed from the focus group interviews, we examined the observational data and the course materials, to specify how factors can be seen in for example teacher-student interactions, assignments, or exams. Analysis of these data involved mainly counting (e.g. numbers of questions), for example in analyzing teacher elicitations. Elicitations are forms of communication that invite students to exchange academic information about facts, interpretations, reasons, or opinions (Mehan, 1979). In the analysis we focused on teachers' elicitations in interactive lectures by studying the proportion of lecturing versus questioning, the number of (follow up) questions, and the wait time after a question. We analyzed two 50-minute lessons of each course. Determination the level of teacher questions was based on a revised version of Bloom's taxonomy of cognitive learning objectives (Anderson & Krahtwohl, 2001). Level 0 =

non-content; level 1= remembering; level 2= comprehension/application, and level 3= analysis/synthesis/evaluation. Since the analysis of levels of questions entailed interpretation, reliability was checked by a second coder, resulting in a kappa coefficient of 0.83.

4.3 Results

First we provide an overview of the (combinations of) challenge factors during the various course activities (research question 1). Then we present a description of the way in which challenge factors are visible in the learning environment (course materials and class observations; research question 2).

Challenge factors

The challenge factors we found in a previous study (Scager et al., 2012) were sufficient to describe the challenge in the six courses of this study: autonomy, complexity, and teacher expectations. Table 4.3 shows the challenge level, and the presence or absence of the challenge factors during the six courses, with positive values indicating challenge and negative a lack of challenge. The average level of challenge was calculated by subtracting average students' perceived ability from the perceived level of challenge. Table 4.3 shows that the level of challenge was always a result of the presence or absence of two or three of the challenge factors (autonomy, complexity, and teacher expectations). The highest levels of challenge (rows one to 11) were usually experienced when the three factors occurred simultaneously. The three exams in the top 11 rows were an exception to this rule. The level of challenge students experienced varied both across and within courses. In the Sci300 course, for example, students experienced both high and low levels of challenge, whereas in the SSc300 course the relative challenge was consistently positive.

Table 4.3
Analysis of challenge factors per course activity and the level of challenge between perceived challenge and ability in sequence of level of challenge

	Course	Week	Level of challenge	Course activity	Challenge factor			
					Autonomy	Complexity	Teacher Expectations	Combination
1	SSc 300	4	2,7	Paper set-up	+	+	+	C+A+E+
2	SSc 300	15	2,3	Final paper	+	+	+	C+A+E+
3	Sci 300	14	2,3	Exam	-	+		C+A+/-
4	Sci 300	13-14	2	Research design	+	+	+	C+A+E+
5	SSc 200	7	2	Exam	-	+/-		C+/-A-
6	SSc 300	1-3;6-8;10-12	1,4	Readings & class discussions	+	+	+	C+A+E+
7	SSc 200	12-13	1,4	Research design	+	+	+	C+A+E+
8	SSc 300	7	1,2	Case study	+	+	+	C+A+E+
9	SSc 200	3	1,2	Article analysis	-	+	+	C+A- E+
10	Hum 100	6	1,2	Exam1	-	+		C+A-
11	SSc 300	11-12	1,1	Essay	+	+	+	C+A+E+
12	Sci 300	7	0,3	Group presentation	+/-	+		C+A+/-
13	Sci300	10-11	0,3	Individual orals	+/-	+		C+A+/-
14	Hum 100	14-15	-0,5	Exam2	+/-	+		C+A+/-
15	SSc 100	8	-0,2	Exam	+	+		C+A+
16	Sci 100	9	-0,4	Debate	-	+		C+A-
17	Hum 100	10-12	-0,5	Paper	+	+/-		C+/-A+
18	Hum 100	1-5;7-10; 12-14	-0,5	Reading and class & forum discussions	-	+	+/-	C+A- E+/-
19	SSc 100	4-14	-1,3	MC tests		-	-	C- E-
20	SSc 100	1-3;5-7;9-15	-1,5	Reading & class discussions	+	-	-	C-A+ E-
21	Sci 100	4+8+15	-1,8	Exams 1&2&3		-	-	C-A-
22	SSc 200	1-3,4-7,10-11	-2,5	Reading & class discussions	+/-	-	-	C-A+/- E-
23	Sci 300	3-6;8-9;11-2	-2,7	Reading & class discussions	+/-	+/-	-	C+/-A+/- E-
24	Sci 100	1-4;11-14	-2,9	Reading & class discussions	+/-	-	-	C-A+/-E-
25	Sci 100	13	-3,2	Bryson Essay	+	+/-	-	C+/-A+ E-

Note. The units of analysis are the course activities (column 5). Courses and course periods in which these activities took place are indicated in columns 2 and 3 (SSc = Social Science, Sci = Science, and Hum = Humanities). Columns 6 to 8 show whether students considered the presence of the factor to contribute to experiencing challenge (+), or considered the absence of the factor to contribute to experiencing a lack of challenge (-). Empty boxes indicate that the factor did not occur in the discussion of the course activity, and +/- that some students mentioned this factor in relation to experiencing challenge and others in relation to experiencing a lack of challenge. The last column shows the combination factors (C=Complexity, A= Autonomy, and E= Teacher Expectations). The rows are ordered by the challenge level, which is a comparative score representing the difference between perceived challenge and ability (column 4).

Considering the presence or absence of challenge factors, we clearly see different patterns (Table 4.3). Whereas autonomy occurred in student activities with both positive and negative challenge levels, complexity appeared more often in challenging periods and was more often absent during a lack of challenge. The factor 'teacher expectations' was only found in course activities with a challenge level above one. When we compared course activities with the highest challenge level (top 11 rows) with activities of low challenge, the main difference was the presence of high teacher expectations. In the bottom seven rows, complexity or autonomy was in some cases high, but teacher expectations were invariably low or absent.

Manifestations of the challenge factors

To answer the second research question on the manifestation of challenge factors in the learning environment, we describe the manifestations of the factors mentioned by students, followed, when relevant, by an examination of the course materials and/or the observations.

Autonomy

Five indicators for establishing autonomy, or the lack of it, were found: choice in assignments, control over grades, student responsibility, non-directive feedback, and openness to student initiative. Table 4.3 shows that all or some of the students experienced autonomy in most (17 of 24) of the course activities.

a) Choice in assignments.

Most of the assignments mentioned in Table 4.3, such as writing or presenting, included according to students some level of choice of topics. In two of these assignments, students felt that there was no choice. In the Hum100 course, for example, students talked about the lack of challenge of participating in a digital forum discussion: "I felt that the forum should have been more student-driven, there was not enough debate but rather answering the questions posted by the teacher." (Susan)

To analyze manifestation of choice in the course materials we looked at the degree to which topics and sources were predefined by the teachers. In contrast to students' experience in the forum discussion assignment (Hum100), the assignment as described by the teacher did allow students to choose between "responding to the study questions, discussing each other's response, or posting a question." Although the assignment did offer some choice, students apparently did not perceive this.

b) Control over achievement.

A second indicator of autonomy is the command students had over their achievement. At UCU, grades were typically assembled from three or more graded assignments and/or exams. These assignments and exams differed largely in the degree of control students experienced. Assignments such as writing essays or preparing class presentations generally provided students with perceived control over their achievement, as Frank (SSc200) explained: "The presentations were way easier [than the exam], because you had spent much more time on the actual analysis. So, you know what you are doing, I could have talked twice as long on it." Frank obviously felt

good and in control during the presentation of his article analysis. When preparing for exams, on the other hand, students generally experienced a lack of control over their achievement: "I don't know what to anticipate [...] it makes me nervous." (Sophie SSc200) Exams clearly caused stress, which decreased students' enjoyment, whereas graded assignments such as presentations generated effort as well as enjoyment. The rudimentary delineation between students' perception of having or lacking control was evident between exams and other graded assignments. The degree of autonomy students experienced during exam preparations varied across exams, however.

The varying degrees of students' perceived control over their achievement in the exams were displayed in the course materials, which disclosed that differences in the form of the exams explained why some exams listed in Table 4.3 created more feelings of autonomy than others. Two exams (Table 4.3, rows 3 and 14) were open-book, which apparently provided some students with feelings of control, and one exam (row 15) was a take home exam, which could be compared to writing an essay, and also provided students with control over their achievement.

c) Student responsibility.

On the interpersonal level, autonomy was mentioned with regard to the responsibility teachers allowed students regarding class attendance. As Tobias (SSc100) explained: "He [the teacher] says it's up to you basically [...] to choose to be there", which for Tanya added to the challenge: "The freedom you get in my case stimulated the learning process." Manifestations of the responsibility students were allowed were found in the other data. First, some course manuals stated rules and expectations very clearly. Some were strict, using controlling language: "You are required to attend all lectures. Attendance registers will be taken" (SSc200), whereas others were more autonomy-supportive: "It is my goal to take you seriously as an honors student. Because of this, I will not bother you with insipid tests in which you reproduce knowledge that I handed to you beforehand. I will also not police you for classroom presence or having read the materials." (SSc300)

Another indicator of the responsibility students were given was found in student attendance rates and numbers of latecomers. Attendance rates and number of latecomers differed across courses, which aligned with the responses of teachers on student behavior. The two teachers who were lenient, according to students, had lower attendance rates than the other four, and waited a little longer (± 5 minutes) than the other teachers for students to come in, filling the time by getting some coffee and chatting with the students who were already there[6]. Second, they used to react in a relaxed way to students dropping in after the teacher had started the lesson ("Good morning, Simon", in a friendly tone of voice). Students in these two courses perceived autonomy during 'reading and class discussions' (Table 4.3, rows 6 and 20). In the other four courses, student attendance was noted more strictly, and latecomers were scarce. Perceived autonomy was +/- in these courses (Table 4.3, rows 18, 22, and 23).

d) Non-directive feedback. Another indicator of facilitating autonomy mentioned by students was the non-directive, informative way in which teachers provided their feedback on draft versions of

6 Attendance rates were checked by counting the turnout, and the waiting time by counting the minutes the teacher started later than the official time.

their assignments. In the Sci300 course, for example, students had to design their own research proposal. Intermittent feedback of the teacher was clearly focused on maintaining students' autonomy, as Peter explained: "The feedback sessions consisted of giving us more things to think about rather than narrowing down stuff." As the feedback on individual and group assignments was planned outside class meetings and therefore not observed, this aspect cannot be concretized by means of the other data.

e) Openness to student initiative. Receptiveness to student input is another sign of autonomy-facilitating teacher behavior (Reeve, Bolt, & Cai, 1999) mentioned by students: "He really thinks about what you say, he is interested in our questions and opinions." (Alena Sci100)

To find manifestations of openness to student input, class observations were studied. First, we examined how teachers responded to student questions and comments. Class observations revealed no differences between teachers, however; all six teachers welcomed and seriously responded to student questions and comments. Courses, however, did differ in the number of times students actually took the initiative by asking (uninvited) questions or offering opinions. In Table 4.4, the number of student initiatives in two instruction-based lessons (50 minutes each) of the six courses is shown.

Table 4.4
Number of student initiatives in two instruction-based lessons[7]

Course	Student input		
	Questions	Opinions	Total
Sci100	4	5	9
SSc100	1	9	10
Hum100	15	4	19
SSc200	6	6	12
Sci300	4	3	7
SSc300	0	31	31

Table 4.4 shows considerable variations in student input during instruction-based lessons which do not concur with the students' perceived autonomy during class discussions as shown in Table 4.3 (rows 6, 18, 20, 22, 23, and 24). Overall, autonomy was clearest in the SSc100 and SSc300 courses, whereas student input was highest in the Hum100 and SSc300 courses. The fact that student input was higher in these two courses specifically could be related to the fact that students were better prepared for these classes, as we will show in the next section.

In sum, student autonomy was experienced in many ways. Students were given some degree of autonomy, in assignments or interpersonally, during lessons in all courses. Lack of autonomy

7 Student input was operationalized as the number of non-teacher initiated student contributions (questions, comments and opinions).

was felt most profoundly during exam preparation, which undermined the perceived control over their achievement.

Complexity

Complexity was, as shown in Table 4.3, mentioned most frequently as a challenge factor. According to the students, complexity was found in the novelty of learning material or the level of thinking that was required, or both.

a) The novelty of the learning material. Students mentioned the novelty of the material as a challenge factor, particularly in the Hum100, Sci300, and SSc300 courses. The Hum100 and SSc300 courses had a philosophical perspective, which was new to students. In the Sci300 course students needed to read research articles which were quite technical, as Peter (Sci300) said: "At our level of understanding the field, reading goes at a pace of one hour per page, so it takes a lot of time. Once you get better and understand the techniques, it gets faster." The novelty made it difficult mainly at the beginning of courses. Discrepancies between challenge and ability levels, as shown in the student lines, were indeed higher in the beginning of these three courses.

b) The requirement for critical thinking.

The call for critical thinking emerged as a theme in all focus group interviews. In three courses, students needed to critically review research articles, which they found challenging. In the more philosophy-oriented courses students were stimulated to argue or apply concepts, as Alek (Hum100) explains: "I felt challenged the most if you were forced to think for yourself, about something that arises from the discussion, when you have to find arguments and relate the concepts to things outside it." The discussion about the level of thinking also evolved when students discussed course activities where they did not experience the need to think critically, as Nicholas (SSc200), explained: "To me it [his storyline] shows I had to do something, but there was not a lot of thinking involved."

Manifestations of the level of thinking in the other data were found in class observations. As students referred to the questions teachers asked during lessons, we analyzed the level of the questions from two instruction-based meetings (50 minutes each) of each of the six courses (see Table 4.5). Questions in the three 100-level courses were mainly focused on knowledge retrieval and comprehension/application, which aligned with their course objectives. In the higher-level courses however, the levels of questions varied more between courses. Questions in the Sci300 course mainly concerned retrieval of knowledge, and in the SSc200 course the teacher asked questions predominantly focused on comprehension and application, whereas the questions of the SSc300 teacher were largely at level 3: analysis, synthesis, or evaluation.

Table 4.5
Level of teacher questions

Question	level 0	level 1	level 2	level 3
Sci100	0%	50%	34%	16%
Hum100	14%	18%	43%	25%
SSc100	0%	0%	41%	59%
SSc200	7%	16%	67%	12%
SSc300	2%	6%	19%	70%
Sci300	5%	47%	32%	16%

Note: Levels are based on a revised version of Bloom's taxonomy of cognitive learning objectives (Anderson & Krahtwohl, 2001). Level 0 = non-content; level 1= remembering; level 2= comprehension/application, and level 3= analysis/synthesis/evaluation.

The novelty of the course material combined with the high level of questions of the SSc300 teacher might explain the differences in the level of complexity between the three 300 level courses during "class discussions" (see Table 4.3). The complexity caused by the novelty of the philosophy-oriented content of the SSc300 course was reinforced by the level of the questions. In the Sci300 course, the reading of research papers was novel, but the level of questions asked in lessons focused on the retrieval of knowledge, which could explain the difference in opinion on the perceived complexity (+/-).

Teacher expectations

Course activities with a challenge level of 1.0 or higher always included high teacher expectations (see Table 4.3), suggesting that teacher expectations were a decisive challenge factor. Course activities with perceived high teacher expectations (the top eleven rows of Table 4.3) included several research assignments, such as designing a research plan. The challenge these assignments generated was not surprising; as was stated in the introduction, inquiry-based and authentic tasks are specifically recommended to challenge high-ability students (Hertberg-Davis & Callahan, 2008; Mönks & Katzko, 2005; Snow & Swanson, 1992). During the activities "reading and class discussions", challenge levels varied widely across courses. In one course (SSc300), students experienced readings and class discussions as quite challenging. The focus group interview revealed that the high challenge level in this course was mostly induced by the expectations of the teacher.

In the focus group interviews students mentioned five aspects of teacher behavior most visible in class discussions that contributed to their perception of his or her expectations: the teacher's reputation; the announcement of challenge; the teacher as co-learner; the need for active participation, and the need to prepare for class. The majority of references are drawn from the focus group of the SSc300 course because in this course students experienced teacher expectations most explicitly.

a) The teacher's reputation

In the advanced courses, students generally know or know of their teacher's reputation. Jamal (SSc300) observed: "I also took the 200-level course with this teacher, and I knew he would be theoretical." In 100-level courses, students generally do not know their teachers in advance, but some actively seek information, as Jan (Sci100) revealed: "The teacher has a really high ranking as a scientist. I looked it up on SCOPUS, which is a website where you can find the number of citings and so on, and he was ranked 21, which is really high; he had had many articles published." Apparently, students actively sought information about the reputation of the teachers, and started the courses with their own anticipation of their teacher's expectations. Students had therefore already created a representation of their teacher's behavior which guided their interpretation of that behavior. Teacher reputation as an indicator of high expectations was not visible in the course materials or the observations.

b) Announcement of challenge

According to students in the SSc300 course, their perception of teacher expectations was already evident in the first meeting. Kim said: "In the first class we were told by the teacher: 'You guys are going to be completely intimidated and at times feel completely lost, but by the end you'll see how worthwhile it was'". The fact that Kim literally quoted what the teacher had said 15 weeks ago indicates that students strongly sensed their teacher's ambition with respect to their learning processes.

Display of the teacher's announcement of challenge could also be found in the course observations and materials. During class sessions, the teacher repeatedly mentioned the difficulty level of the book, and pointed to additional lunch meetings for those who had problems understanding the material, and his availability in case students should feel "panicked". Additionally, the course manual comprised the following announcement of challenge:

> This is not an easy course, we will be exploring many areas of knowledge that are new, and which may involve some philosophy of science and social science methodology. Answers may seem unclear at times- or even most of the time. As a result, this course may produce side effects such as academic dizziness and a sense of being lost. However, after some perseverance, student's minds tend to clear up significantly towards the end of the course.

Course manuals of the other courses did not contain announcements of challenge.

c) Teacher as co- learner.

In the SSc300 course, the feeling of being treated as (near) experts by their teacher made students experience teacher expectations and create an image they wanted to live up to, as shown in the next excerpt.

> Kim: And also the way he introduces the discussion topics, it starts already in a way that… Claire: Yes… the stories he tells us sitting on his bike, about these conferences. He's asking us what we think, as equal people, expecting us to have equal knowledge and skills. Or asking us about a couple of books he's reading.

> Kim: He puts himself on your level, but it's a level that's not really yours; he asks students what they think on various topics, and you want to get to that level. It's in between somewhere.

Class observations corroborated with the teacher positioning himself and the students as peers. In the first part of the lessons invariably the teacher started to share his thoughts and ideas with students, for example about his research, and asked them for suggestions and opinions. The focus group interview continues:

> Eric: He [the teacher] asked a lot and taught little. He presents himself just as lost as we are, puts himself on an equal footing with us, and from that you start a journey together.
> Jamal: And that's intimidating for many students, because you are used to someone who gives you confirmation, reassuring. [The teacher] doesn't do that, he is searching, saying he's lost too at one point, wants you to go on a journey with him.

In the observations we noticed that the teacher invoked discussions frequently, offering his opinions and responding respectfully to students' opinions. Students for example would occasionally disagree with him, which he accepted gracefully with "point taken".

d) The need for active participation

Teacher expectations also arose from the interactional approach of the teachers during classes. One indicator students mentioned was the proportion of instructing and questioning. Additionally, the teacher's uptake on student's answers affected the perception of teacher expectations. Claire (SSc300) said: "He would ask questions, and he kept on asking questions, deeper every time, until I had no idea what to answer to these questions". The time the teacher waited to allow students to think of an answer also contributed, or diminished, students' perceptions of teacher expectations. Anna (Sci300) explained that she did not feel that the teacher expected an answer: "What happened now was like 'does anybody have any questions?', and we were sitting there waiting until [the teacher] would say something, like 'well, see you next week'." Teacher expectations apparently were perceived through the elicitations during instruction and discussions, which align with research findings on the manifestations of high teacher expectations (Brophy, 1983; Rubie-Davies, Hattie, Townsend, & Hamilton, 2007). In Table 4.6, we confront student perceptions with an analysis of the observations focusing on teacher elicitation. Elicitations are forms of communication that invite students to exchange academic information about facts, interpretations, reasons, or opinions (Mehan, 1979).

Teachers of three courses held the floor by devoting 78 to 93% of the time to lecturing, which is not surprising, since we selected instruction-based lessons. Three teachers, however, did not dominate classroom talk, lecturing 31% to 48% of the available classroom time and reserving more than half of the lesson for discussions.

Table 4.6
Comparison of teacher elicitations in the six courses

Course	Lecturing time	Number of questions			Wait time in seconds	
		Total	Initial	Follow up	M	SD
Sci100	31%	57	41%	60%	1.2	0.57
Hum100	78%	25	52%	48%	1.3	0.85
SSc100	90%	23	61%	39%	1.2	0.59
SSc200	48%	57	39%	61%	1	0.19
Sci300	93%	27	85%	15%	1	0.57
SSc300	42%	47	47%	55%	1.6	1.46

Student talk was managed by these teachers by way of asking questions (57, 57, and 47 respectively). Only one teacher (Sci300) did not frequently ask follow-up questions, which could also be a consequence of the factual character of the questions (see Table 4.5). Questions of teachers of the Sci100, SSc200 and SSc300 courses were largely follow-up questions. The character of the uptake, however, differed; the SSc300 teacher generally asked for further explanation or more depth ("Why?"), whereas the Sci100 and SSc200 teachers more often asked for other, or better, answers ("What else?"). The time teachers allowed students to think about an answer was mostly fewer than two seconds. The teacher of SSc300 waited relatively long (up to nine seconds) when students did not immediately respond. These results show that none of the individual three factors of teacher elicitations convincingly explained the differences in student-perceived high teacher expectations of these courses. When, however, we add the level of the questions (Table 4.4) to the patterns, differences become clearer. The high-expecting SSc300 teacher appeared to instruct little and ask a lot of questions, although not the most, on a higher level. Furthermore, he waited longer for answers, and took up on student answers with follow-up questions which required more thinking.

e) The need to prepare for class

Students' active participation in class is graded at UCU, and students also assessed their teacher's expectations by the necessity to prepare for class. Kim (SSc300) said: "It is almost impossible to contribute [to class discussions] if you haven't read the material." In the Hum100 course, students were expected either to contribute to the internet forum (prior to class) or in class. All of them prepared for class by participating in the forum discussion: "For me, it forced me to read; I had to at least do enough to answer the questions." (Alek) In the other four courses, class discussions were not very challenging. Students in these courses experienced a lack of teacher expectations, as Mark (SSc100) explained: "Well, I did do the reading at first, but not that much afterwards, I just stopped reading. It wasn't difficult to participate in the class discussions; you didn't need the reading for that [...]. It was interesting but not too demanding."

In conclusion, high teacher expectations were displayed in several ways, and it seemed that the perception of high teacher expectations in the SSc300 course was built on a combination of all

indicators. For these students, the teacher expectations interacted with the other two challenge factors, autonomy and complexity, as is shown in the next excerpt:

> Claire: I chose the difficult article because I knew the teacher would like it. (*So you chose this article for the teacher as well?*) Yeah, and also for myself.
> Jamal: Definitely, it's not that he would say that your proposal wasn't good, if you do a simple regression analysis…But then I was thinking: that would be quite straightforward. He sets high expectations and you want to live up to that, you want to reach a level that is really deep. The teacher has very high expectations and this makes it challenging. […]
> Eric**:** Yeah, [*the teacher*] is really very theoretical, and most of his students are willing to go an extra mile for him. There are only a few teachers who can get that out of students.
> Jamal**:** Yeah, you definitely want to go an extra mile.

This fragment demonstrates that the high expectations of the teacher affected the choices students made; they used their autonomy to increase the complexity level of their course assignments to accommodate the assumed expectations of their teacher.

4.4 Conclusions and Discussion

In the introduction we argued that challenge has been an underlying characteristic in effective gifted education but has not been researched explicitly. Building on the results of a previous single case study of an extremely challenging honors course (Scager et al., 2012), the present study examined six honors courses, questioning (1) which factors in the learning environment constituted (lack of) challenge for high-ability students and (2) how these factors were manifested in the design and implementation of the courses. To answer the first research question, we determined students' perceived levels of challenge and ability using a storyline approach, and studied the challenge factors that explained experiencing (a lack of) challenge using focus group interviews. Students experienced the highest levels of challenge when three factors (autonomy, complexity, and teacher expectations) occurred simultaneously. This result corroborates previous results of a case study (Scager et al., 2012). We found that the challenge level varied both between courses and within courses (between different course activities). Autonomy was experienced in both high and low challenging course activities, indicating that autonomy cannot directly explain the level of challenge. Findings of our previous study (Scager et al., 2012) showed autonomy to be an essential component for explaining the challenge. The autonomy students experienced in that course, however, was quite extreme; students described it initially as "lack of guidance", whereas in retrospect they valued the sense of ownership the autonomy allowed them. Possibly, the autonomy students felt in the current six courses was required for their motivation, but was not sufficient to challenge them to "go the extra mile". The third factor, teacher expectations, appeared to be a crucial addition to the other two factors. Course activities with the highest levels of challenge were all characterized by high teacher expectations. In the cases where teacher expectations were lacking, students could be motivated by the complexity and autonomy, but additional reinforcement was needed to challenge them to the utmost. The strong impact teacher expectations have on students is surely not a new finding, but research on the influence of teacher

expectations has been conducted mostly on primary school level (e.g. Babad, 1993; Brophy, 1983; Weinstein, 2002). In research on older students (Rubie-Davies, Peterson, Irving, Widdowson, & Dixon, 2010), it is assumed that the influence of teacher expectations decreases as students become older and rely more on self-assessment. The results of the current study contradict this assumption, indicating that even students on the tertiary level are susceptible to the expectations of their teachers.

Concerning the second research question, a variety of indicators was found, making clear how the design and implementation of the course foster autonomy, complexity and teacher expectations. Novel and abstract content and a required high level of thinking in particular make courses complex. Mere understanding of difficult readings did not challenge students; neither did critical analysis of simple material. Autonomy was manifested in a variety of aspects of the learning environments. An obvious manifestation of autonomy was the level of choice students were allowed in assignments such as being able to choose a topic of their own interest for a paper. Just as important was the extent to which they felt they had control over their achievement/grades (e.g. writing a paper) as opposed to teacher control (exams). Further, allowance of student responsibility, non-directive feedback, and the teacher's openness to student input indicated students' autonomy. All of these manifestations match the descriptions of autonomy-supportive teaching behavior of Reeve, Bolt, and Cai (1999). High teacher expectations were conveyed through various teacher behaviors. First, it is noteworthy how the teacher's reputation shaped students' ideas about the expectations the teacher might have, which is beyond course design and realization. Second, teacher expectations were formed by the teacher's announcement of the difficulties students were expected to face. Third, the teacher's position as a co-learner, manifested for example when one teacher asked for students' ideas about his own research, reinforced students to perceive high teacher expectations. According to Heller (1999), such role-taking is encouraging for high-ability students. Fourth, teacher expectations manifested in the need to participate actively, conveyed by actions such as asking a lot of questions, waiting for answers, and up taking on students' answers, confirm findings of other research on manifestations of high and low teacher expectations (Babad, 1993; Brophy, 1983; Rubie-Davies et al., 2010).

Challenge and the self-determination theory

We now discuss part of the results from the perspective of Deci and Ryan's (1985) self-determination theory (SDT) of intrinsic motivation. This theory posits that intrinsic motivation stems from the need to feel competent, autonomous, and related. The motivating effect of feelings of autonomy found in our study aligns with findings of researchers of the self-determination theory who found that controlling social contexts (e.g. exams, deadlines) may undermine intrinsic motivation, whereas autonomy-supportive social contexts (e.g. offering choice) facilitate intrinsic motivation and self-determination (see for reviews Deci, Koestner, & Ryan, 1999; Deci & Ryan, 2000, 1991; Vallerand, 1997). Intrinsic forms of motivation have been found to lead to greater interest, and greater effort, whereas less self-determined types of motivation are negatively related to these outcomes (Benware & Deci, 1984; Ryan, Koestner, & Deci, 1991; Ryan & Stiller, 1991). Our

results only partly agree with these findings. The less self-determined activities, such as exams, encouraged students to work, but the accompanying stress seemed to hinder their enjoyment. Also, one of the assignments, the forum discussion in the Hum100 course, did offer students the choice between answering the pre-posted questions or initiating a debate. As participation was graded, however, students merely answered the pre-posted questions and did not perceive it as a choice. On the other hand, autonomy in the sense of allowing students responsibility to prepare, combined with complexity, enhanced students' enjoyment, but did not raise a level of challenge that inspired students to go the extra mile. High teacher expectations were needed to engage these high-ability students to the full. The (high-expecting) teacher of the SSc300 course, for example, managed to motivate students in reading the material without having planned exams. Our results for teacher expectations might be associated with the need for relatedness, which Ryan and Deci refer to as the desire to feel connected to others (Deci & Ryan, 2000). The motivation of students to meet their teacher's expectations therefore might originate from their need to be recognized or appreciated by their teacher. Research by Assor, Roth, and Deci (2004) on the relationship between children and parents captured the parental use of conditional love as a form of external control, undermining autonomous motivation. Although we did not find that students' autonomy was undermined by the high expectations of the teacher, it is important to realize that students might tend to perceive their teacher's regard as dependent on their ability to meet their expectations. Further, Deci and Ryan (2000) identify the need for competence as a basic condition for intrinsic motivation, which is represented as optimal challenge where levels of difficulty and ability are in balance. This balance might be necessary for maintaining intrinsic motivation.

Implications

A first implication of our results for teaching high-ability students is the need to aim for a high level of complexity, specifically the level of thinking. Second, students can be challenged if teachers let go of controlling behavior and stimulate students to take responsibility for their learning. Third, students seem to need some reinforcement from their teachers, which can be found in the form of exams, though high teacher expectations and open assignments are preferred, because these are less detrimental to intrinsic motivation (Deci & Ryan, 1985).

Although the same three challenge factors were found in our previous case study, the pedagogical models of the courses in these case studies were completely different. The course in the previous case study was strongly project-based, whereas the most challenging course in this multiple case study (SSc300) was more traditional, being instruction-based. Apparently, for creating a challenging learning environment for high-ability students the pedagogical model is less important; the pattern of challenge factors (autonomy, complexity, and teacher expectations) can be orchestrated in more than one way. This finding has potential implications for the way in which learning environments are designed. For example, whereas research-based designs could offer challenges, the challenging potential of the complexity of such assignments might be increased or decreased by the level of autonomy and expectations teachers provide. More attention is therefore needed to the teacher role in the learning process.

Limitations and suggestions for further research

Although this study provided insights for creating a challenging learning environment for high-ability students in higher education, conclusions must be regarded as tentative. The findings were related to courses of a single college. One direction for future research, therefore, is to examine the challenge factors found in this study across a wider range of honors programs.

In the current study, we averaged students´ storylines of perceived challenge and ability. Individual discrepancies between challenge and ability could have been more extreme, causing a level of challenge that would pull students too far out of their comfort zones, which could in turn lead to too much stress and anxiety. The support teachers offered their students prevented that happening in the courses we studied. Letting go of controlling behavior (e.g. exams), increasing expectations, and relying on students´ motivation and will to learn call for an established support system.

Given that students are responsive to challenging conditions in the learning environment, it might be oversimplifying matters to interpret our results as a one-directional effect of instructional behavior on the students. First, students actively construct challenge by interpreting their teachers' behavior, and, second, the engagement of the students also affects their teacher's motivation and instructional behavior (Pelletier, Seguin-Levesque, & Legault, 2002; Reeve & Tseng, 2011). Specifically, the intelligent, creative, and motivated population of honors students has potential to enhance the learning environment when teachers allow them to (Heller, 1999). Heller's proposed role transition of teachers of high-ability students to that of co-learners could stimulate students to contribute actively to the instructional flow (Heller, 1999), and subsequently influence the teachers' engagement (Pelletier et al., 2002). Combining these two ideas we recommend further research to address challenge as an aspect of gifted education as it comes about during the process of continuous, responsive and anticipatory teacher-student interactions.

5. Teacher dilemmas in challenging high-ability students in higher education[8]

Abstract

High-ability students need to be challenged to prevent boredom and stimulate their learning. Applying challenge in practice is however not straightforward, because challenging students could conflict with other teacher responsibilities, causing dilemmas. This study aimed at disclosing dilemmas teachers encounter when challenging high-ability students as well as the considerations accompanying their choices. Considering the evidence in the 12 interviews with teachers, seven main categories of dilemmas were found. The diversity of choices and considerations brought forward by the teachers, indicate that it matters who the teacher is, what he or she stands for, and is able to do.

[8] This chapter has been submitted as Scager, K., Akkerman, S. F., Pilot, A., & Wubbels, T. (submitted). Teacher dilemmas in challenging high ability students in higher education.

5.1 Introduction

Research on education for high-ability students has consistently shown the need to challenge high-ability students to prevent boredom and stimulate their learning (Eddles-Hirsch, Vialle, Rogers, & McCormick, 2010; Gallagher, Harradine, & Coleman, 1997; Heilbronner, Connel, & Reis, 2010; Marra & Palmer, 2004; Noble & Childers, 2008; Rogers, 2007). Previous studies have revealed three factors that in higher education are conducive to challenging high-ability students to give their best work: high levels of complexity, student autonomy, and teacher expectations (Scager, Akkerman, Pilot, & Wubbels, 2012, 2013). One might argue that accomplishing teaching practices with these features is relatively straightforward. However, challenge is likely to be but one of many objectives in teaching that follow from teachers' conceptions about teaching and learning or the college environment. Research has shown teaching to reflect a complex mix of responsibilities, therefore it inherently involves dilemmas (Billig et al., 1988; Enyedy, Goldberg, & Welsh, 2006; Kelchtermans, 2009; Lampert, 1985). As we will explain, realizing challenge in educational practice also easily evokes dilemmas for teachers. The current paper investigates implicit and explicit dilemmas underlying challenging high-ability students in higher education in an honors college environment. This allows us to specify the particular complexity of choices and considerations that are part of teaching in high-ability education. Considerations entail the weighing of the perceived benefits and losses of the options to choose from.

In the following, we first attend to the theory of dilemmas in the field of teaching before focusing on the concept of challenge in higher education and its dilemmatic nature.

Teacher dilemmas

Conceptually, dilemmas pose a choice between alternatives with equally perceived profits and losses (Billig et al., 1988). Other than a problem that can be solved, a dilemma cannot be fully resolved without leaving some residue (Cushman & Young, 2009; Denicolo, 1996). Accordingly, a teaching dilemma can be conceptualized as a situation in which teachers must choose between two or more relevant courses of action, when residue on each side hinders the decision (Berlak & Berlak, 1981). Dilemmas are subjective, as one person's dilemma might not be another's, because the profits and losses will be valued differently. Hence, dilemmas require a choice between courses of action, based on considerations of the perceived values of the profits and losses of both choices.

A body of research recognizes and describes the dilemmatic nature of teaching, relating to the contradictions individual teachers experience in their daily practice as a consequence of the complex and ambiguous nature of teaching (cf. Cabaroglu & Tillema, 2011; Enyedy et al., 2006; Helsing, 2007; Kelchtermans, 2009; Lampert, 1985; Lyons, 1990). In line with this, Lampert (1985) positioned dilemmas at the heart of teaching, regarding the teacher as a dilemma manager, "balancing a variety of interests that need to be satisfied in classrooms" (p.190). According to this view, teachers cannot evade dilemmas in practice, but they can find

various strategies to handle them, depending on teachers' values, priorities or knowledge, and their awareness and ability to reflect on alternatives (Kelchtermans, 2009).

Research on teacher dilemmas covers a wide range of topics, most of which can be considered ideological or ethical in nature (Ben-Peretz & Kremer-Hayon, 1990), although there is no clear line between the two. Ideological dilemmas in teaching include both educational content and concepts about the teaching and learning process (Billig et al., 1988). Ideological dilemmas concerning the content are mostly related to issues of autonomy versus conformity to external curricula (Ben-Peretz & Kremer-Hayon, 1990; Colnerud, 2006; Lampert, 1985). Dilemmas concerning concepts about the teaching and learning process represent the disconnections between teachers' conceptions of learning and their actual teaching practice (Enyedy et al., 2006; Lyons, Freitag, & Hewson, 1997; Tillema, 2004), and are often evoked by the implementation of new learning concepts, such as constructivism, or student-centered learning (Billig et al., 1988; Wang, 2011; Windschitl, 2002). A second type of dilemma focuses on ethical dilemmas of the profession of teaching (Colnerud, 1997, 2006; Shapira-Lishchinsky, 2011), often revolving around issues of care, justice, and power (Buzzelli & Johnston, 2001; Colnerud, 2006; Noddings, 1992; Oser, 1994). Dilemmas between caring and justice, for example, can occur in maintaining (organizational) rules and standards, where the teachers must choose between individual needs and applying equal and uniform rules (Noddings, 1992; Pope, Green, Johnson, & Mitchelle, 2009; Shapira-Lishchinsky, 2011). Care for students can also lead to student dependency, accordingly creating dilemmas between the wish to help students and to empower students to become self-regulated (Sumsion, 2000).

Although teacher dilemmas are widely acknowledged as being central to teaching, to our best knowledge there is no research considering the dilemmas teachers experience when challenging students. Offering challenge is an important factor for the learning of all students, and in particular high-ability students. An increased understanding about the considerations of teachers when choosing for or against maximizing challenge may enable teachers to better manage and change their work, and shed light on the competing values, beliefs, demands, and practical constraints that may complicate teaching high-ability students.

Examining teachers' considerations is not straightforward. Many studies on teacher dilemmas have used direct interview methods, asking teachers what dilemmas they experience. This can be problematic as teachers are not equally sensitive to dilemmatic issues and able to reflect on them (Billig et al., 1988; Oser, 1994). According to Billig et al. (1988), dilemmas can be expressed explicitly or implicitly, and in the latter case, discourse should be analyzed and interpreted. In the present study, teachers were asked to reflect on their practice, which brought to light explicit as well as more implicit dilemmas. Further, the status of a dilemma could vary with respect to the urgency the dilemma holds for teachers. An analytic framework was developed in this study for identifying dilemmas in teachers' discourse, which is also sensitive for the urgency status of the dilemmas.

The dilemmatic nature of challenge

Findings in literature on gifted education point to the importance of increasing the level of challenge for high-ability students compared to mainstream students (Heller, 1999; Lubinski & Benbow, 2000; Snow & Swanson, 1992). Providing high levels of challenge therefore is assumed to be an objective for teachers of high-ability students. Increasing the challenge, however, may contradict other objectives, such as keeping the study time doable, and subsequently cause dilemmas for teachers. For our conceptualization of challenge we draw from psychological theories that include the notion of challenge, such as the control-value theory of motivation (Pekrun, 2006), the self-determination theory of motivation (Deci & Ryan, 1985), and the flow theory (Csikszentmyhalyi, 1975). These theories express the inherent relationship between the levels of challenge and ability, indicating a need for balance, with a challenge level slightly above the ability level for positive effects on motivation and learning. Csikszentmyhalyi (1975) assumes that an appropriate balance between challenge and ability creates a state of flow; accordingly, tasks should increase in their level of difficulty as the student's level of ability increases. This postulate of optimal challenge in the flow theory is consistent with the need for competence in the self-determination theory (Deci & Ryan, 2000): to create a feeling of competence, tasks need to be optimally challenging, implying a balance between challenge and abilities. In the control-value theory of motivation (Pekrun, 2006) it is posited that a lack of balance, caused by demands that are too high or too low, causes negative emotions that can subsequently impair intrinsic motivation. Hence, finding a balanced level of challenge is a delicate matter.

In line with the flow theory, a study by Scager et al. (2012) showed that high-ability students reported the highest motivation and learning effects when (perceived) challenge exceeded the level of (perceived) ability. We found three factors evoking the challenge, respectively high levels of complexity, student autonomy, and teacher expectations. While high-ability students reported maximum learning in periods of high challenge, a disbalance of perceived challenge and ability induced feelings of worry and frustration (Scager et al., 2012). Negative emotions have been found to damage student motivation (Linnenbrink, 2006; Meyer & Turner, 2002; Pekrun, 2006). When students within one group differ in their perceived ability, a particular challenge level can increase motivation for some students, while for others the challenge might be too high and subsequently might induce negative emotions and impair student motivation. This suggests that teachers need to counterbalance the challenge for some students by creating support and a safe classroom environment (Clough, Berg, & Olson, 2009).

Heterogeneity of groups of high-ability students in an honors college environment can thus be expected to evoke dilemmas when creating a challenging learning environment for all. Heterogeneity in student groups is omnipresent, but in mainstream college courses such student differences do not necessarily lead to dilemmas for teachers, because it might be expected that students feel comfortable in courses that allow all of them to perform well effortlessly. In challenging courses, however, students are more tested, and some students

may feel that they cannot cope with the demands, and subsequently lose their motivation. Increasing challenge thus brings differences between students to light and accordingly generates dilemmas for teachers in creating a learning environment which is challenging for all students. Maximizing challenge, after all, is not the only responsibility teachers have; they also pursue non-academic objectives, such as caring for the well-being of individual students and developing a good learning climate (Butler, 2012; Oser, 1994). We imagined that specifically high teacher expectations would be the challenge factor causing dilemmas for teachers, because applying high expectations has a more (inter)personal and less technical character than increasing autonomy or complexity. Teacher expectations are communicated in direct interaction with students, for example by way of asking many questions and waiting for answers, which could cause discomfort in class and as a result evoke dilemmas.

With respect to the practical relevance of the study, we believe that a study of dilemmas could be fruitful for understanding the important aspects of the nature and complexity of teachers' work in general and specifically when handling phenomena such as challenge. Insight into the way experienced teachers reflect on these dilemmas can potentially provide important information for beginning teachers of high-ability students in higher education. Our aim is to disclose dilemmas relating to challenge so that they become visible in the complexity of practice, with respect to both the content of the dilemmas and their considerations. Although this study is situated in honors education, we believe that the findings have relevance for teaching higher education in other contexts as well. Given the dilemmatic nature of teaching and the assumption that creating challenge might evoke teacher dilemmas, this study aims to answer the following research questions: What dilemmas do experienced teachers encounter when challenging high-ability students, and what considerations do they have in making choices related to these dilemmas?

5.2 Methods

We initially interviewed teachers to reflect on whether and how they realized challenge. These reflections appeared to lead teachers to talk about the dilemmas they experienced in relation to challenging their students. This caught our interest, as it made it evident that increasing student challenge is complicated to put into practice. Therefore, we decided to focus on the parts in the interviews that revealed dilemmas. We then checked the transcripts to see whether we could find a wide variety of dilemmas within and across the interviews.

Context

The teachers we selected all taught at an international liberal arts and sciences honors college of a research university in the Netherlands. Students at this college are selected for their academic excellence (minimum GPA of 3.0), curiosity, and motivation. Courses are given on three levels, the 100, 200, and 300 levels. Courses at the 100 level are accessible for students of all majors and years, entry to a 200-level course requires a 100-level course in the same disciplinary track, and entry to a 300-level course requires a 200-level track course.

Participants

Twelve teachers were chosen to participate in this study. The first author also observed the courses of six of these teachers as part of a previous study (Scager et al., 2013). The second group of six teachers had not been part of the previous study, and were chosen because student evaluations showed that their courses were experienced as challenging. The indicator of the level of challenge was the evaluation item: 'How difficult was this course?' All 12 teachers were experienced and were evaluated very positively in student evaluations (>4.2 on a 5-point scale).

Courses

Although all teachers were responsible for more than one course at the college, they were asked to talk specifically with one of their courses in mind. Courses were well spread over the departments and comprised all three undergraduate years:

- At the 100 level: one science, one humanities, and two social science courses;
- At the 200 level: one humanities and three social science courses;
- At the 300 level: two science, one humanities, and one social science course.

Interviews

The first part of the interview was open, the second part of the interview was structured around the three challenge factors, complexity, high teacher expectations, and autonomy, and factors that counterbalanced high challenge (support and safety). Teachers were asked to assess their levels on these five factors in the particular course on a scale from one to five. The first six teachers were interviewed a second time, several months later. In this interview teachers reflected on feedback from the interviewer on their courses. This feedback was based on analyzing student focus group interviews with their students (see Scager et al., 2013) and looked at students' comments about their perceived level of challenge. The interviews took place between January and May 2012 and took at least 1.5 hours.

Analyses

The interviews were transcribed verbatim. The first author read the interviews to find, refine, and elaborate the dilemmas. Steps in the dilemma analysis then comprised 1) identification and selection of dilemmas, 2) detecting themes and status of dilemmas, and 3) uncovering the direction of and the considerations for the choice between the options comprising the dilemma.

Step 1 Identification and selection of dilemmas

As a first step, all fragments comprising dilemmas were selected from the interviews. We used a reasoning chain as the defining cut for the fragments (Chi, 1997), sometimes combining separate parts within one interview when they concerned the same dilemma. Dilemmas not relating to challenge were removed. To be able to identify dilemmas, we developed a list of linguistic cues that reveal dilemmas. As Engeström and Sannino (2011: 369) state: "A crucial

point is that contradictions cannot be observed directly; they can only be identified through their manifestations." According to Engeström and Sannino, contradictions, or dilemmas, are expressed in the form of hedges and hesitations, recognizable in linguistic cues such as: "On the one hand" or "Yes, but…" Based on our data, we expanded their list of linguistic cues, as shown in Table 5.1.

Table 5.1
Linguistic cues and frequency of observation in the interviews

Linguistic cues (speech acts) indicating a dilemma	Frequency	Teachers
1. Hesitations: (maybe, I am not sure, I think)	9	5
2. Deliberations (contrasting, 'if I do this, it endangers that'; 'on the one hand.. on the other hand', Talk of discussions with colleagues, 'but')	45	12
3. Problematizations (difficult, challenging, labeling his/her own behavior negatively)	12	5
4. Questions or Solutions: (rhetorical questions, questions to the interviewer; 'maybe I could, 'I'm thinking of')	12	8
5. Eliminations ('I do not want to…')	12	5

Note: Fragments could comprise more than one linguistic cue.

The second linguistic manifestation – deliberations – was observed most frequently and in fragments of all 12 participants, although participants all had their own manner of speaking; some showed more hesitation, or tended to talk in terms of questions and solutions, while others seemed to prefer deliberations.

Step 2 Detecting themes of the dilemmas

In a second step, all fragments comprising a dilemma were listed and coded on two aspects: the theme and the urgency status of the dilemma. Fragments often comprised multiple themes, such as A versus B and C. In these cases, we distinguished between the primary dilemma, which was the dilemma most prominently referred to as opposing challenge (A versus B), and the secondary dilemma (A versus C) within a fragment, resulting in 28 secondary dilemmas. For example, a teacher could mention hesitating over increasing the difficulty level of his questions in class because students might feel embarrassed when failing to answer (challenge versus safety dilemma), but also because this teacher behavior could affect the relation between teacher and students negatively (challenge versus a positive relation dilemma). We identified seven recurrent dilemmas from the transcripts: maximizing challenge versus (1) caring for students' general concerns; (2) maintaining students' safety; (3) maintaining a positive relation with students; (4) keeping all students aboard; (5)

maintaining students' enthusiasm; (6) matching teachers' self-understanding, and (7) complying with (perceived) expectations of others.

Step 3 Determining the urgency status of the dilemma

The urgency status of the dilemmas varied. In some cases, teachers were experiencing the dilemma here and now, while in other cases, the teachers' position seemed to be set. We differentiated three distinct statuses dilemmas could have for teachers, summarized in Table 5.2: urgent, in deliberation, and decided.

Table 5.2

Urgency status of the dilemma

Status of the dilemma	Position	Cues
1= Urgent	Position uncertain Teacher chooses A, but is not happy with A, or with the alternative B.	Speaking in the present Providing recent examples Speaking of emotions Repetitions and/or length of discussion
2= In deliberation	Position hesitant Teacher chooses A, but in the future considers B.	Speaking in the future Speaking of ´could´ rather than ´should´
3= Decided	Position certain Teacher chooses A, is satisfied with choice, and does not consider B (anymore).	Speaking in terms of they, he, she, instead of I Speaking in the past Theorizing, generalizing

Step 4 Uncovering the direction of and the considerations for the choice between A or B

In a fourth step, the content of the dilemmas was further analyzed, in search for positions teachers took, and the considerations teachers expressed when elaborating their choices. Considerations were understood as the weighing of the gains and losses of the choice to be made. The grain size of the analysis in this step was cut to the level of individual sentences. The direction of their choice was coded as in favor of challenge, or against when teachers favored other responsibilities over challenge.

To ensure the reliability of the analysis in the four steps mentioned above, we first practiced the analysis by coding and discussing five fragments with all four authors. In a next stage, 13 ambiguous cases (30% of the fragments) were checked by the second author, and discussed until agreement was reached. As a third and final check, an inter-rater reliability analysis was conducted to assess the degree to which a colleague consistently coded the linguistic cues, the theme of the dilemma, the direction, and status to 25% of the fragments that were not yet checked. The resulting Cohen's (1960) kappa coefficients indicated substantial agreement (Landis & Koch, 1977): κ=0.7 for linguistic cues; κ=*0.8* for dilemma themes*;* κ=0.7 for direction of choice*;* κ=*0.8* for challenge factor, and κ=*0.8* for status of the dilemma.

5.3 Results

Overview of the dilemmas

Seven dilemmas, relating to maximizing the challenge, were found, as shown in Figure 5.1.

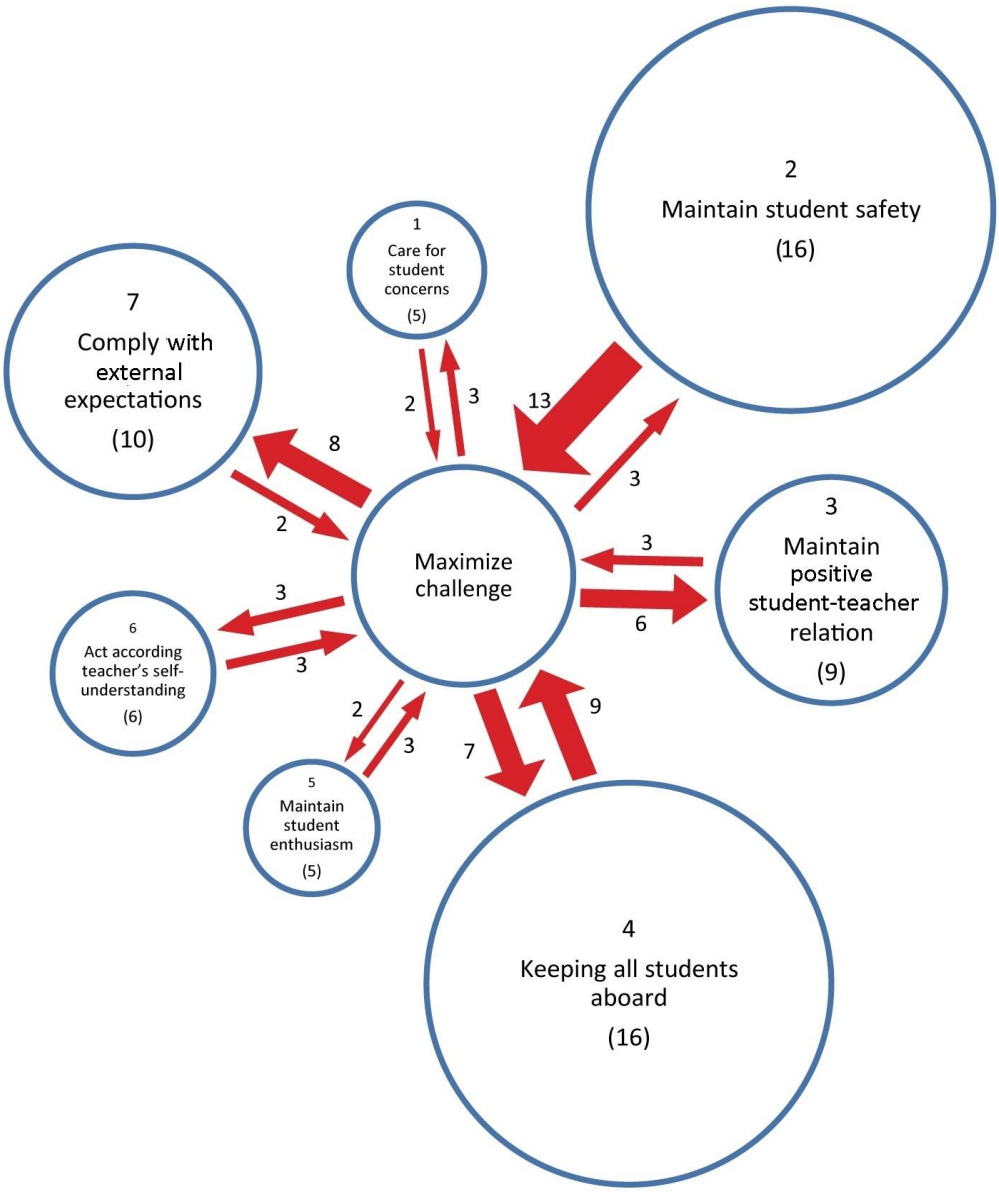

Figure 5.1. Diagram representing the seven dilemmas related to maximizing challenge.

Note: The surface area of the circles represents the relative amount of times the dilemma was mentioned, and between brackets the frequencies of the dilemma, mentioned as a primary dilemma, are given. The width of the arrows symbolizes the frequency of the choices pro and contra challenge.

In Table 5.3, the frequency of these seven dilemmas, the number of teachers who experienced each dilemma, the direction of the position teachers chose between the two sides, and the urgency status of each dilemma are shown.

'Maintain students' safety' (dilemma 2) and 'keep all students aboard' (dilemma 4) were the dilemmas teachers experienced most. Least experienced were the dilemmas 'care for students'

concerns' (dilemma 1) and 'maintain students' enthusiasm' (dilemma 5). The number of teachers mentioning the dilemmas was lower than its total frequency, indicating that teachers sometimes mentioned a dilemma more than once, albeit in a different situation and with different considerations. For example, student safety could be threatened by asking students difficult questions, or by openly providing critical feedback on their work. Teachers tended to choose to maximize challenge slightly more often than not (35 and 32). In the dilemma 'maintain students' safety' teachers mostly chose to maximize the challenge, while in choices between challenge and 'comply with external expectations' (dilemma 7) teachers opted primarily for the latter.

Table 5.3
Overview of the dilemmas: frequencies, number of teachers mentioning the dilemma, direction of choice, and urgency status

	Dilemma	Frequency[1]	Challenge Factor[2]				Teachers	Direction of choice[3]		Status[4]		
			A	C	HE	Combi-nation		↑	↓	1	2	3
1	Care for student concerns	5 (1)		1	4		4	2	3	5		
2	Maintain students' safety	16 (2)	2	1	11	2	9	13	3	5	6	5
3	Maintain a positive relation with students	9 (5)			3	6	6	3	6	5	1	3
4	Keeping all students aboard	16 (5)	1	11	1	3	8	9	7	4	9	3
5	Maintain students' enthusiasm	5 (4)		1	3	1	4	3	2	1	1	3
6	Act according teacher's self-understanding	6 (8)		1	5		5	3	3	3	3	
7	Comply with external expectations	10 (3)	5	3	2		6	2	8	2	4	4
	TOTALS:	*67(28)*	*8*	*18*	*29*	*12*		*35*	*32*	*25*	*24*	*18*

Note. 1 'Frequencies' shows the number of times the dilemma concerned a primary dilemma, and between brackets the number of times the dilemma were indicated as secondary dilemmas.
2 'Challenge factors' are A: Autonomy, C: Complexity, HE: High Expectations, or a combination of these factors.
3 An upward arrow represents a choice for challenge, while a downward arrow means that the teacher chose for the alternative instead of challenge.
4 Three urgency statuses are indicated: 1) urgent, 2) in deliberation, and 3) decided.

The direction of the choices also differed between course levels: in 100-level courses teachers more often chose *against* challenge (10:21), while teachers of 300 courses more often chose *for* challenge (16:5). In 200-level courses, differences were smaller (9:6). Most of the dilemmas were urgent (status 1), meaning that the teachers were currently uncertain about their chosen course of action, although the frequency differences between the three statuses were small. Further, the different dilemmas were well spread over the teachers, and therefore

also over the levels of the courses. The urgency status of dilemmas was also well spread over the teachers (and course levels).

With respect to the challenge factors, we found that 'high expectations' was, as expected, the challenge factor which was most frequently (29) involved in the dilemmas teachers described, followed by 'complexity' (18) and 'autonomy' (8). Dilemmas involving high teacher expectations were also more often urgent than when challenge concerned autonomy or complexity (17 urgent, 5 urgent, and 3 urgent, respectively).

Description of the dilemmas and considerations

1) Maximizing challenge versus caring for students' general concerns

This dilemma entails a choice between maximizing challenge and caring for general concerns of students, such as their anxiety level, or their freedom to prioritize courses or course elements. Challenge in this dilemma represents the high expectations of the teachers. The status of the dilemma was 'urgent' in all five cases.

Two teachers reported choosing to maintain their high expectations, although they were concerned about the negative impact the challenge could have on students' stress levels. Not willing to reduce their learning objectives was one of their considerations. Second, some stress was assumed to be inevitable for students to meet the learning objectives. A third consideration referred to the motivation of students, as shown in the following example.

> T: You need some stress for people to do the work, but actually the motivation is already high, the adrenaline is very high, the desire to produce something that is of very high quality, and they want to write those papers and get them right, they want the teacher feedback.
> I: So they don't need the additional stress?
> T: No, but I do occasionally. I got questions like "what do you have to do to get an A?", this kind of stuff, and I often go like, "O, no idea, I don't care, I'm not interested in an A, I'm interested in you being able to do this stuff," and that's not a constructive answer either. So you know I, I need to think about this course to make it so that there aren't people who only go for the grades.

In this excerpt, the nature of student motivation was at stake. This teacher valued learning over performance, while her class seemed to consist of both students focused on learning and students focused on grades. She hesitated about her challenge level when she described students who were already motivated to produce high-quality work and did not need additional stress. Then again, she did not want to lower her expectations to avoid students focusing on the grade.

Two teachers reported choosing to keep the level of challenge for their course manageable, accepting that students might prioritize other obligations. This concerned two teachers of 100-level courses, and as one of them explained:

> I don't know how important this course is for their future. But I don't want to give them the impression that they have to focus on this course and not on other courses. I don't think I should be offended that some students rather devote their time on another course. I think that that is their responsibility. [...] The best thing I get from students is that they say: "I did not feel pressure, but I liked it a lot."

Care for students' well-being, particularly their pressure level, was the first consideration of this teacher. A second consideration was the value these teachers awarded students' freedom to take the responsibility to choose their own priorities, instead of forcing them to dedicate their full effort to their course. The word 'offended' in the excerpt indicates that the teacher's own feelings were involved, considering a negative effect of his choice. Some satisfaction was apparently wished for by this teacher, and if not found in the energy students put into the course, satisfaction should be derived from students' positive evaluations of the course. The fear that students could be bored by the lack of challenge was mentioned as a second negative effect, creating doubt over the choice these teachers had made.

2) Maximizing challenge versus maintaining students' safety

A second dilemma encompasses the choice between maximizing challenge and maintaining students' feelings of safety, for example to contribute freely in class discussions, and avoiding students 'feeling stupid.' This dilemma was one of the most frequently mentioned ones (16 cases by 10 teachers). The challenge in this dilemma referred to complexity, high teacher expectations, and student autonomy, as well as to a combination of two or three of these factors.

In most cases (13 of 16), maximizing challenge was chosen over maintaining students' safety. These teachers considered it important to pull students out of their comfort zone to stimulate thinking, by asking many and difficult questions and waiting for answers and/or providing difficult assignments. There were several considerations for choosing challenge over safety. First, asking difficult questions and expecting answers was assumed to stimulate learning and deepen the level of class discussions. Second, questioning communicates the need to prepare for class, which was also meant to help students keep up with the material and prevent procrastination, thereby supporting students' achievement in the exams later on in the course. Third, discussions (in class and through the Internet) were used to stimulate students to share their insights, and particularly share their problems and faults, capturing an orientation to learning as a social process. A byproduct of the intensive interaction in class, according to one of these teachers, was that students learned to know each other, which might contribute to the group cohesion. The loss of the choice to maximize challenge was considered to be the risk that students would feel embarrassed when failing to answer, and as a consequence shy away from sharing their thoughts and questions in class. Further, two teachers considered questioning students directly by name too 'schoolish' for this level of education, negatively affecting students' own responsibility. Challenge thus was considered to stimulate shared learning but at the same time was expected to cause danger for sharing and disrespect for students' responsibility.

In three cases, teachers reported choosing safety over maximizing challenge. Two of these cases concerned courses at the 100 level. A consideration of these teachers was that the subject was new to these students and the aim of the courses at this level was to motivate students for the subject. For these teachers this included diminishing initial insecurity students might feel, which, as one of these teachers argued, was not achieved by scaring them

by asking questions that were too difficult. Other considerations of these teachers mirrored the losses mentioned above: not wanting to affect the open class atmosphere and students' own responsibility to prepare for class. The loss of this choice for all three teachers was the acceptance that students did not always prepare for class.

3) Maximizing challenge versus maintaining a positive relation with students

A third dilemma entails the choice between maximizing challenge and maintaining a positive relation with students. In this dilemma, challenge was mostly described as having high expectations. In three cases, teachers reported choosing to maximize the challenge, with the same considerations mentioned as in the previous dilemmas: attaining the high level of class discussions and individual student work they aim for, helping students to avoid procrastination, and encouraging shared learning. In addition, these teachers believed that interesting class discussions also contributed to their own enjoyment which had subsequent positive effects on students' enjoyment. The perceived loss of their choice was the way students might perceive them as authority, as one of them said: "I am a very demanding, insisting teacher; if I had had myself as a teacher, I would not have been able to get on with him." One of these teachers also expected that the possibly disturbed relation would show up in the students' evaluations of the course.

Five teachers described prioritizing a positive relation with students over enforcing their expectations. One consideration of these teachers was that these are university students, who should be expected to take responsibility for their own learning. Also, these teachers considered it 'not their job' or 'not their style' to control students' preparation. Further, teachers spoke about preferring a 'non-authoritarian' or 'relaxed class atmosphere,' encapsulating an unwillingness to use their power. Nonetheless, all five teachers showed uncertainty about their choice, wondering whether students actually took the responsibility to do their readings.

4) Maximizing challenge versus keeping all students aboard

The fourth dilemma concerns the differentiation of students' (perceived) ability within groups, referring to the choice teachers had to make between maximizing challenge or tuning in on less able students in order to keep all students aboard. This is one of the most frequently mentioned dilemmas (16). The challenge in this dilemma concerned increasing the complexity. The direction of the teachers' choices was nearly evenly spread (9:7), and the considerations of losses and benefits accompanying their choices were diverse, and not solely dependent on the direction of their choice.

Teachers choosing to maximize challenge considered the risk that students for whom the challenge level was perceived too high might lose their self-confidence as the most important loss. Students subsequently would lose their motivation for the course and the desire to participate in class. One of the benefits of setting the bar high according to these teachers was that this enhances the depth of discussions in class and the quality of students' papers, which

made teaching more interesting and rewarding. Another benefit, mentioned by two teachers, was that a temporary loss of confidence caused by high challenge was a learning experience: students sometimes need to realize that they have to raise their efforts. Lastly, for one of the teachers, a consideration to maximize the level of challenge was to provide students with the opportunity to gain access to a top-level master's course.

Teachers reporting choosing to tune in on the students with lower (perceived) ability considered the risk of under-challenging the group of students with the highest (perceived) ability and accordingly losing their motivation. Some of them were even afraid that they might underestimate the whole group of students. Two teachers emphasized the more ideological side of the choice to focus on either the top or the bottom half of the class, as one of them said:

> But there is also a risk of creating a sort of sect: I, the teacher, you, the good ones, and together we form a sort of little sect. For instance, some of us continue in a social network. [..] I don't want to give them the impression of forming a legion of favorite students here.

5) Maximizing challenge versus maintaining student enthusiasm

The fifth dilemma was mentioned five times and referred to the possibility that students lose their enthusiasm when teachers and/or texts are too demanding. Challenge in these cases concerned having high expectations (3), or keeping up high levels of complexity in the course (2).

In two cases, teachers chose not to maximize the challenge, which in both cases encompassed the complexity, considering that students would lose their enthusiasm when it got too difficult. These were both cases in 100-level courses. Stimulating students' motivation for the field of study, which was new to the students, was considered an important objective for these teachers. Simplifying the content, according to these teachers, was intended to provide students with access to the material. The loss of their choice was considered under-challenging students, and not getting the most out of them.

In the other three cases, teachers reported choosing to maximize the challenge, acknowledging the risk of a (temporary) loss of students' enthusiasm. Their considerations depended on the situation teachers described. Students could lose their enthusiasm or their confidence because their research plan was criticized, or because of getting low grades. High expectations were expected to stimulate students' achievement, but at the same time, these teachers anticipated that students could also lose their enthusiasm and decrease their efforts as a result.

6) Maximizing challenge versus acting in line with the teacher's self-understanding

A sixth dilemma entailed the choice between maximizing challenge or acting in line with one's professional self-understanding. 'Self-understanding' refers to the teachers' conception of themselves as teachers, including their self-esteem, job motivation, and task perception (Kelchtermans, 2009). This dilemma was more often enunciated as a secondary than as a primary dilemma (8:6), indicating that other objectives mostly prevailed over the teachers' own interests. Challenge in these cases was high teacher expectations.

In three cases teachers reported maximizing the challenge, even though that affected their workload considerably. In these cases, teachers thought that having high expectations of students also required maximum effort from themselves. Expecting students to achieve highly on assignments, for example, required teachers to support students by supplying frequent and elaborate feedback, to be very accessible for student questions, and to provide and discuss additional materials with individual students who wanted to learn beyond the defined curriculum. These teachers said they invested hours far beyond what was expected from them formally.

In three cases, teachers described choosing not to maximize the challenge. Their main considerations were related to their own dispositions and nature, as one of them explained:

> A bit more expectation, about their behavior in class, I suppose would be good. [Gives an example of not intervening when a student was facebooking in class.] I left it, but well, it was not so much a matter of respect, to be honest, but it was more my own difficulty. Perhaps I don't dare to do it.

For this teacher, expectations would entail intervening when students do not live up to them, which was considered preferable, but not in line with his own aptitudes.

7) Maximizing challenge versus complying to (perceived) external expectations

This dilemma encompasses the choice between maximizing the challenge and complying with the expectations of others (students, colleagues), or with college regulations. Challenge in this dilemma referred to autonomy (5), complexity (3), and high expectations (2).

In eight out of ten cases, teachers described choosing to comply. Seven of these cases related to courses at the 100 level, which are supposed to cover basic concepts and theories. These teachers believed that students, as well as the teachers of subsequent 200-level courses, expected them to cover the basics of the discipline, in the notion that students' interest might decrease as a result of the lack of challenge.

In two cases, teachers described choosing to maximize the challenge despite the (perceived) expectations of others. In one of these cases, the teacher allowed students a high level of autonomy in choosing topics for their research assignments, withholding directive support. Some students, however, expected the teacher to help them choose their topic. The dissatisfaction of students in view of that, and the possible showing of their dissatisfaction in the course evaluations, were considered acceptable losses of the choice to maximize students' autonomy.

5.4 Conclusions and discussion

This study aimed to disclose dilemmas teachers in higher education encounter when challenging high-ability students and the considerations they have when making their choices, weighing perceived benefits and losses. Because analyzing the dilemmas that were more implicitly communicated required some interpretation, for which no standard instruments were available, we developed a procedure to do this. This four-step procedure, the 'Dilemma

Analysis Instrument', might be of help for other researchers in the classification of dilemmas in future studies.

We detected seven main categories of dilemmas that can be ordered into five groups: (a) choices that could benefit and harm a student at the same time; (b) choices that could advance one pedagogical objective but impair another; (c) choices that could be good for one student but harmful for another; (d) choices which are beneficial for students but undesirable for teachers, and (e) choices that benefit students but oppose colleagues or college policy.

(a) The first dilemma encompasses the choice between maximizing challenge and care for students' general concerns, such as their stress level and their freedom to choose their own priorities. In these cases, teachers reasoned that their responsibility for students moves beyond optimal intellectual development and encompasses nonacademic goals such as care for students' general well-being. In other words, what is good for one student's learning could harm the same student's well-being.

(b) The next three dilemmas (2 Maintaining students' safety, 3 Maintaining a positive relation with students, and 4 Keeping all students aboard) encompassed instructional choices. Maximizing challenge is supposed to stimulate students to bring their talents to full bloom. But at the same time, challenge could negatively affect students' feelings of safety (dilemma 2), and subsequently scare them away from contributing to class discussions. Psychological safety is the feeling of being able to show and employ oneself without fear of negative consequences for self-image or status (Kahn, 1990). Research on the influence of psychological safety has mainly focused on safety among peers. These studies indicate that psychological safety has a significant impact on students' social and academic motivation (Wentzel, Battle, Russel, & Looney, 2010), and is associated with better team learning and achievement (Edmondson, 1999). This supports the hesitation of teachers to affect students' safety. In dilemma 3, challenge was opposed to maintaining a positive relationship between teacher and students. This is in line with the findings of Williams and Wilson (2012), who encountered a similar tendency for teachers to focus on establishing positive relationships with students to the detriment of the intellectual quality of the learning experience, rooted in the belief of teachers that positive teacher-student relations are the key to managing students' behavior. Recently Winheller, Hattie, and Brown (2013) argued that this is a misunderstanding. In dilemma 4, teachers encountered challenge as a threat to students' enthusiasm, judging that enjoyment decreases when students lose their confidence when challenge exceeds students' ability. Here, the concept of self-efficacy is at stake, which has indeed been found to affect student motivation (e.g. Pintrich & Zusho, 2007). These instructional dilemmas show how one pedagogical objective might harm other pedagogical objectives.

(c) Dilemma 5 encompasses maximizing challenge versus keeping all students aboard. Students differ in their (perceived) ability, and creating an optimal challenge level for all students in class is nearly impossible. Teachers thus feel that they need to choose to serve one group to the detriment of another. Although this dilemma is probably experienced in all

education, we expected it to be more prominently encountered by teachers of high-ability classes, because high challenge is a key concept in honors education. In fact, this dilemma was one of the main dilemmas, while we hardly found it in the dilemma literature. The main theme in this dilemma is that what is good for one student could be harmful for another.

(d) All of the first five dilemmas involved frictions between various courses of action in the best interest of students. This list, however, did not fully address why certain values would be prioritized above others in the teacher's decision-making. Teachers also deal with their own needs, limitations, task perception, and personality (dilemma 6), such as maintaining their own enthusiasm. The importance of the self-understanding of the teacher is increasingly recognized, and has been found to be an additional source for dilemmas to occur (Enyedy et al., 2006; Kelchtermans, 2005). This does not necessarily mean that teachers' deliberations are selfish. For example, they argued the importance of their own enthusiasm for students' motivation and performance, an argument that can be supported by research findings (Brophy & Good, 1986; Long & Hoy, 2006; Witcher, Onwuegbuzie, & Minor, 2001). In turn, teaching highly motivated and highly achieving students may positively influence teacher enthusiasm (Frenzel, Goetz, Lüdtke, Pekrun, & Sutton, 2009), signifying the reciprocal nature of teacher-student interaction. The focal point of this dilemma is that what could be good for students might be unfavorable for the teacher.

(e) The seventh dilemma encompassed maximizing challenge versus complying with (perceived) expectations of others (students/colleagues), or with institutional regulations. According to Troman, Jeffrey, and Raggl (2007), institutional policy creates the majority of teacher dilemmas, which does not concur with our findings. One explanation could be found in the fact that the college under study allows teachers a considerable amount of freedom, which is something that, according to Colnerud (2006), initiates teachers to take their own responsibility, resulting in the internalization of dilemmas that were initially created externally. Summarizing, the central theme in this dilemma is: what benefits students or teachers could harm colleagues or institutional policy.

Overall, this study shows that applying a significant educational objective such as challenge is not straightforward. Challenging high-ability students in order to bring their talents to full bloom is an educational objective that can easily interfere with other values. Teaching implies not only a technical agenda of applying concepts and achieving the curriculum goals, but foremost a complex relationship with students, characterized by moral responsibilities and emotional experiences (Kelchtermans, 2009). The diversity of choices and considerations brought forward by the teachers in this study, point out that it matters who the teacher is, and what he or she stands for, and is able to do. Fundamentally, the moral dimension of teaching is to do what is in the best interest of students (Kelchtermans, 2009). Acting in the best interest of students is, however, not straightforward, because there can be conflicting matters which are both in the interest of a student, and moreover, what is good for some students might be detrimental for others in the same class. Accordingly, there seems no single right or wrong way to handle dilemmas such as the ones related to challenge and

described here. Our results show that explicating and reflecting on these dilemmas advances our understanding of teaching practice.

Relating challenge factors to dilemmas

As expected, high expectations created more dilemmas than the other two challenge factors (autonomy and complexity). Dilemmas concerning teacher expectations also had the status 'urgent' more often than the other two. An explanation could be found in the fact that high expectations are mostly communicated directly in interaction with students, for example in asking questions, or in commenting on students' work. It requires courage from the teacher to do that, because communicating high expectations could engender immediate uncomfortable silences, embarrassment, or feelings of anxiety.

The interactional nature of communicating expectations, therefore, is also expected to affect the teacher-student interpersonal relationship, which can be confirmed by the fact that high expectations as a challenge factor were mostly found in the dilemmas concerning 'students' safety' and the 'relation with students.' In the model of interpersonal teacher behavior (Wubbels, Brekelmans, den Brok, & van Tartwijk, 2006), two dimensions describe the student-teacher relationship: influence and proximity. Communicating high expectations implies intervening when students do not meet expectations, thus taking more control. Doing this for many teachers goes along with reducing at the same time the proximity between students and teacher in the perception of students. Lower proximity has indeed been shown to imply a lower quality of relationship (Wubbels et al., 2006). This is not, however, a necessary result of taking control. It is possible to take higher control without reducing proximity, and when teachers are able to do this, it might help them solve this dilemma.

Urgency status and direction of the dilemmas

The urgency status of the dilemmas included all possible statuses for all the teachers except for one, who experienced only urgent dilemmas. Urgent dilemmas were more discomforting than dilemmas with the status 'in deliberation' or 'decided.' Helsing (2007) detected fundamental differences in the ways teachers respond to and interpret their dilemmas. Some describe dilemmas and uncertainty as a prime cause of anxiety, frustration, burnout, and poor teaching, while others claim that the recognition of uncertainty is an important ingredient of improved practice and that it protects teachers from pessimism, guilt, and frustration (p.1328). The teachers we interviewed were all experienced, and therefore can be assumed to have a broad spectrum of courses of action, and knowledge of the consequences of their behavior. Their knowledge and experience are expected to allow them to improvise more fluidly than beginning teachers. This raises the question of whether the dilemmas of less experienced teachers are experienced more often as 'urgent.'

The direction of the choices teachers made was quite evenly distributed over for and against applying high challenge, indicating that the situations were indeed dilemmatic in nature. Based on similar considerations, teachers described dissimilar choices, weighing the gains and losses of their choices differently.

Limitations and further research

We recognize three limitations in this research. First, the sample was small and consisted of experienced college teachers. It would be worthwhile expanding the study to a larger group of teachers, including beginning teachers. Expanding the study could be used to validate, or extend our findings, revealing, for example, correlations between the types of dilemma and teacher experience. One can imagine that beginning teachers choose less often to maximize challenge to avoid dilemmas than their experienced colleagues do. Further, we would expect that beginning teachers experience more often 'urgent' dilemmas than experienced teachers do. Another relevant line of research considers the weighing of losses and benefits of teachers when making choices. When confronted with economic dilemmas, people tend to weigh the losses of a choice heavier than the profits (Kahneman, 2011). It would be important to know whether this finding also applies to teachers' considerations in educational dilemmas, since the tendency to weigh losses heavier than benefits could hold teachers back from taking risks and trying new instructional strategies.

Second, all the teachers in this study were from one college, so our findings must be viewed as being situated in this specific context. To gain a more comprehensive perspective of dilemmas related to teaching high-ability students, it would be relevant to expand this research to other honors programs, other countries, and other teaching contexts.

Third, it should be noted that the dilemmas were drawn from interviews that were conducted after the courses were finished. These interviews elicited reflections on events which happened during teaching. Conscious deliberations, however, are often not the primary driver of behavior; intuitions play a large role in the decisions people make (Haidt, 2001), which is why the teachers' reflections may not fully reflect their on-the-spot considerations.

Implications

Reflecting on dilemmas constitutes powerful opportunities for learning and development. Dilemmas are endemic to the teaching profession (Biesta, 2012; Kelchtermans, 2009; Lampert, 1985), suggesting that some contradictions cannot be resolved and that the best way to handle them is to find ways to keep the negative side effects of choices to a minimum. According to Biesta (2012), the centrality of judging situations in the teaching profession calls for developing "virtuosity for wise educational judgment." Since judgments in practical situations are often made intuitively (Haidt, 2001), it would be prudent to develop professional intuition. Intuitive decisions are based on knowledge and experience. Experts in a field make better decisions on the spot, while they have relevant experience and a larger and better-organized knowledge base in their domain than novices do (Salas, Rosen, & DiazGranados, 2010). To progress professional intuition, explicit learning from theory and expertise is needed (Salas et al., 2010). One way to develop expertise could be found in collaborative reflection on dilemmas. By their very nature, dilemmas allow the evoking of reflection and argumentation, encouraging teachers to talk about choices and considerations. Recent studies converge on the understanding that reflecting on dilemmas can have beneficial

effects on teacher learning and change, since it provides teachers with the opportunity to examine and challenge their beliefs and to explore alternatives for change and growth (Helsing, 2007; Pareja Roblin & Margalef, 2012; Yoon & Kim, 2010). Pareja Roblin and Margalef found that the acknowledgement of dilemmas enabled teachers to take a critical perspective on their educational beliefs and practice, thereby strengthening critical reflection. Reflecting collaboratively adds to the learning experience, since teachers differ in their experiences and beliefs, and these differences challenge each other's values, viewpoints, and actions (Yoon & Kim, 2010). Explicating and discussing dilemmas, therefore, could help beginning teachers to anticipate dilemmas and experienced teachers to advance their practice.

Teachers' considerations concern both the effectiveness of their teaching and the moral responsibilities for the process, such as care for students' well-being, and for the social climate in class. As situations in which judgments are made are unique, each judgment is new (Biesta, 2012). Therefore, not only are the decisions important, but also the considerations, the discourse about the values, objectives and practical constraints that are involved. This study revealed how teachers can arrive at different courses of action, based on the same considerations, depending on how they value the consequences. Knowledge of these considerations relating to dilemmas in practice could provide valuable insight into the complexities of teaching.

Acknowledgments
We thank the teachers for their cooperation and their open and frank reflections. Further, we would like to thank Jaap Milius for his assistance and valuable suggestions.

6. Conclusions and discussion

In this chapter the conclusions of the individual studies are summarized and integrated. In addition, both conceptual and methodological limitations of the studies are discussed, as well as suggestions for further research. Lastly, the practical implications of the findings for university professors, institutions and students are presented.

6.1 Conclusions

Summary of results

The main research aim guiding this thesis was to examine appropriate instructional strategies for high-ability students at the university level (honors students), addressing three main research questions: First, do honors students differ from non-honors students with respect to the qualities that have been found to be essential for exceptional accomplishments in professional life? (see chapter 2) Second, how can high-ability students in higher education be challenged? (see chapters 3 and 4) And third, what dilemmas do teachers encounter when challenging high-ability students? (see chapter 5)

Differences between honors and non-honors students

The aim of examining appropriate instructional strategies for honors students was based on the assumption that honors students differ from their non-honors peers in their potential for excellence, and therefore need differentiated instructional strategies. To test this assumption, we first inquired differences between honors and non-honors students with respect to the characteristics that are expected to be vital for excellence (**chapter 2**). According to Renzulli's (1986) three-ring model, components relating to intelligence, motivation and creativity are the most important predictors of excellent achievements in professional life. By means of a questionnaire, 1,122 students (467 honors students and 655 non-honors) at Utrecht University were assessed with a self-report questionnaire on six characteristics related to Renzulli's three components: intelligence (1), creativity (creative thinking (2) and openness to experience (3), and motivation (the desire to learn (4), persistence (5), and the drive to excel (6)).

The central question of this study addressed the differences between honors and non-honors students with respect to the six talent variables mentioned above. The results showed that the honors students differed significantly from the non-honors students in terms of the combined variables as well as for the separate variables, with the exception of 'persistence'. A next research question addressed which of these talent variables primarily differentiated between honors and non-honors students. Standardized discriminant function coefficients suggested that the best predictors for distinguishing honors students from non-honors students were the desire to learn, the drive to excel and creativity. The least predicting values were found for intelligence and persistence. The value of these variables as predictors however, differed for the various study programs. The strong distinguishing value of creative thinking was unexpected, as creativity is not an explicit selection criterion for most honors programs. The insignificant scores of 'persistence' could be explained by the possibility that honors students did not need to work very hard in secondary education to achieve well, and their persistence in the face of setbacks was not tested and developed yet (Chamorro-Premuzic & Arteche, 2008). Intelligence was also a weak variable. This was surprising, as honors students are selected on their average school grades, and school grades are usually correlated around r=.45 with intelligence (Chamorro-Premuzic & Furnham, 2005; Lubinski,

2004; Rindermann & Neubauer, 2004). An explanation for the small distinguishing value of (self-assessed) intelligence could be the fact that mainstream students are more likely to overestimate their ability than gifted students, whereas gifted students make more accurate judgments concerning their ability-related performance (Dai et al. 1998; Mabe & West, 1982; Pajares 1996). Further, honors students are part of a selected peer group, which may have influenced their self-assessment. According to Marsh (1987), these students tend to assess their competence in comparison with their peer group (the "big fish-little pond effect"), which may have led to underestimation of their intelligence.

As we found substantial differences in characteristics between honors and non-honors students across disciplines, we further explored these differences within four of the study programs. The variation in results between the study programs might be explained by the differences in selection procedures used in the various programs. In the Law and Liberal Arts and Sciences (LA&S) programs, a combination of grades, application letters and interviews were used, while in the Physics and Humanities departments at the time, selection procedures predominantly relied on self–selection. The differences between honors and non-honors students in both the Law and LA&S groups were relatively strong, with significant differences in each of the talent variables. In the two programs with less comprehensive selection methods, the differences between honors and non-honors students were less strong, with Humanities students showing significant differences for only two of the talent variables, (openness to experience and desire to learn) and Physics honors students not scoring significantly higher for any of the variables. The low scores of the Physics honors group call for further reflection. Physics honors students did not assess themselves to be significantly more intelligent than their non-honors peers, and their score for 'persistence' was noticeably lower. These low scores for persistence are surprising, as in their honors program, Physics students combine two bachelor programs (physics and mathematics), which can be assumed to be quite demanding. It is possible that these students are not accustomed to working hard, because there is no need for them to do so as they might be intelligent enough to succeed with little effort (Chamorro-Premuzic & Arteche, 2008). However, these Physics honors students did not assess themselves significantly more intelligent than their non-honors peers either. This is odd, since their school grades were considerably higher than those of non-honors students (8.1 for honors and 7.6 for non-honors). Possibly the 'big fish-little pond effect' (Marsh, 1987), described above, accounted for this unexpected result. For the LA&S students we found similar results; school grades did not correlate with self-assessed intelligence. In contrast, significant correlations for grades and intelligence were found for the Law and Humanities students. If a 'big fish-little pond effect' appeared, then it pertained to two groups, and not to the other two. This finding can be explained when we look closer into these programs. Law and Humanities students both follow their honors program in addition to the regular program, which they follow together with non-honors students. These students thus can compare their intelligence with that of their non-honors peers, and therefore, the 'big fish-little pond effect' might not apply. Physics students also group with honors as well as with

non-honors students, however the mean school grades of non-honors students in Physics are considerably higher than the mean grades of non-honors students in other disciplines (7,5 and 6,9 respectively). We therefore assume that physics honors students find themselves between intelligent non-honors peers, in which case the 'big-fish-little pond effect' does apply, as well as for LA&S students, who have only honors peers.

Since the results showed that honors students differed significantly with non-honors students with respect to the qualities that have been found to be essential for exceptional accomplishments in professional life (intelligence, creativity, and motivation), it seemed justifiable to assume that these students need distinct challenge to further develop these qualities. After all, the effectiveness of education depends to a large extent on the fit between the learning environment and the abilities, interests and motivation of the students (McKeachie, 1986; Snow, 1986; Pascarella & Terenzini, 1991). These students need to be challenged appropriately and they require a learning environment that matches their ability and level of motivation.

Challenge for students

The next study sought answers to the questions of how to challenge high-ability students, specifically focusing on the way high challenge affected high-ability students' learning and motivation, and the factors that constituted the challenge in the context of an honors college (**chapter 3**). We studied a course that was known for its high challenge. As part of their course, the students, in groups, developed a research program according to national scientific standards, which they did successfully, according to an external jury of experts in the field. Methods included interviews with teachers and students, analysis of course materials, and observation of class meetings. Findings showed that maximum effort and (perceived) learning were experienced when the perceived challenge exceeded the level of ability the most, and during which students felt worried and frustrated. These results indicate that a balance between ability and challenge is vital for enjoyment, but not necessarily for effort, persistence and learning. This finding was unexpected, given that negative emotions have been understood to impede intrinsic motivation (Csikszentmihalyi 1975; Larson 1988; Meyer &Turner 2006). This finding indicates the importance of studying effects of challenge using a longer time scale instead of one single measure. An explanation for students' increased efforts, despite a lack of enjoyment, could be found in the support that was experienced from their peer group. The social support from peers could have prevented or counterbalanced over-challenge and its accompanying feelings of worry and frustration. This assumption concurs with findings of Bakker and Demerouti (2006) who studied factors that cause or prevent burnout in work settings and found social support to be one of the most important factors that balance out high job demands.

As to the question of what constituted the challenge, we found a three-factor model of challenge: autonomy, complexity, and high teacher expectations. The complexity is conceptually closest to the 'challenge' axis in Csikszentmihalyi's (1975) flow-model. Student

autonomy, experienced by students as a lack of teacher guidance (both structuring and feedback) was the second factor we found. The lack of teacher feedback as challenge factor was a noteworthy finding. In educational literature, feedback is described as one of the most powerful elements of an effective learning environment (Black & William, 1998; Hattie, 2009; Sadler, 1989). In our study, students did express the need for feedback from their teachers, but at the same time, they proved to be able to perform exceptionally well without it. These results suggest that students learned to judge the quality of their work themselves, which created pride and a sense of ownership afterwards. Withholding feedback is advocated by several researchers (Nicol & Macfarlane-Dick, 2006; Sadler, 1989; Yorke, 2003), who argue that teacher-transmitted feedback keeps students dependent on their teachers and interferes with self-regulated learning.

The high expectations the students experienced both from their teachers as well as from themselves was the third factor we found. High, but appropriate teacher expectations have been found to be effective for all students (Hattie 2009). However, according to Phillips and Lindsay (2006), high expectations have the risk of generating stress and anxiety, an emotional state interfering with motivation. In this case study we found that the high expectations contributed to the challenge and also to the associating feelings of worry and frustration.

In the third study (**chapter 4**), we investigated the generalizability of this three-factor model of challenge. In this study, six honors courses, varying in level and discipline, were examined, questioning the factors in the learning environment that constituted (lack of) challenge for high-ability students and the way these factors were manifested in the design and implementation of the courses. Results showed that students experienced the highest levels of challenge when three factors (autonomy, complexity, and high teacher expectations) occurred simultaneously, confirming the results of the case study mentioned above. We found that the challenge level varied both between courses and within courses (between different course activities). Autonomy was experienced in both high and low challenging course activities, indicating that autonomy cannot directly explain the level of challenge. Course activities with the highest levels of challenge were all characterized by high teacher expectations. In the cases where high teacher expectations were lacking, students could be motivated by the complexity and autonomy, but additional reinforcement in the form of high expectations was needed to challenge them the utmost. Concerning the way these factors appeared in the design and implementation of the courses, we found a variety of manifestations of autonomy, complexity, and expectations. Although the same three challenge factors were found as in our previous case study, the pedagogical models of the courses in these case studies were very different. The course in the previous case study was strongly project-based, whereas the most challenging course in this multiple case study was more traditional, being instruction-based, indicating that challenge can be created within different pedagogical models. Apparently, to create a challenging learning environment for high-ability students the general pedagogical model is not the decisive factor; the pattern of challenge

factors (autonomy, complexity, and teacher expectations) can be orchestrated in more than one way to create maximum challenge.

Challenge for teachers

Some of the courses in the multiple case study were not very challenging, This suggested the question why experienced teachers in similar settings made different choices with respect to the level of challenge they provided students. To address this question, we interviewed the six teachers participating in the case studies and six more teachers, examining their considerations regarding the choice whether or not to challenge students (**chapter 5**). Findings of this study revealed seven dilemmas relating to challenge, showing that applying a significant educational objective such as challenge is not obvious. The first dilemma encompassed the choice between maximizing challenge and care for students' general concerns, such as their stress level and their freedom to choose their own priorities. Teachers, who faced this dilemma, reasoned that their responsibility for students moved beyond optimal intellectual development and also encompassed non-academic goals such as care for students' general wellbeing. Maximizing challenge was supposed to stimulate students to bring their talents to full bloom. But at the same time, teachers emphasized how high challenge might affect students' feelings of safety (dilemma 2), and subsequently scare students away from contributing to class discussions. Research on the influence of psychological safety amongst peers indicate that psychological safety has significant impact on students´ social and academic motivation (Wentzel, Battle, Russel, & Looney, 2010), is associated with better team learning and achievement (Edmondson, 1999), supporting the hesitation of teachers to affect students' safety. In the third dilemma, teachers argued that challenge might disagree with the interpersonal relationship between teacher and students. Williams and Wilson (2012) encountered a similar tendency for teachers to prioritize establishing positive relationships with students over the intellectual quality of the learning experience. Teachers encountering the fourth dilemma, described challenge as opposite to fostering students' enthusiasm, judging that enjoyment decreased when students lose their confidence in situations of challenge exceeding students' ability. The fifth dilemma encompassed maximizing challenge versus keeping all students aboard. Students differ in their (perceived) ability, and creating an optimal challenge level for all students in class is difficult. Teachers thus felt that they needed to choose for serving one group of students to the detriment of another group. The last two dilemmas encompassed maximizing challenge versus complying with teachers' own needs, limitations, task-perception, or personality (dilemma 6) and versus complying with the (perceived) expectations of others (students/colleagues), or with college regulations (dilemma 7).

The direction of the choices teachers made when coping with the dilemmas was rather evenly divided between 'for' and 'against' maximizing challenge, indicating that the situations were indeed dilemmatic in nature. Based on similar considerations, teachers apparently made dissimilar choices, weighing the gains and losses of their choices differently. The experience of

dilemmas indicates that challenging high-ability students in order to bring their talents to full bloom is an educational objective that can easily interfere with other values. Teaching implies not only a technical agenda of applying concepts and achieving the curriculum goals, but foremost a complex relationship with students, characterized by moral responsibilities and emotional experiences (Kelchtermans, 2009).We also looked at the urgency of the dilemmas, and found that the urgency statuses of the dilemmas were well spread over the dilemmas as well as over the teachers except for one teacher, who experienced mainly urgent dilemmas. Urgent dilemmas were more unsettling than dilemmas with the status 'in deliberation' or 'decided.' According to Helsing (2007), teachers differ in the way they respond to, and interpret dilemmas. In some of the studies she reviewed dilemmas were described as a prime cause of anxiety, frustration, and burnout, while in others the recognition of dilemmas are claimed to be an important ingredient of improved practice (p.1328). Lange and Burroughs-Lange (1994) suggest that comfort with the uncertainty dilemmas evoke come with expertise; over time, teachers learn to accept uncertainty as legitimate aspect of their role. The teachers we interviewed were all experienced, and therefore can be assumed to have a broad spectrum of courses of action which allow them to improvise more fluidly than beginning teachers, and respond to dilemmatic situations with less uncertainty. This raises the question for further research of whether the dilemmas of less experienced teachers are experienced more often as 'urgent.'

General conclusions

Overviewing the results of the four studies led to six lessons learned, which are discussed in relation to the three main research questions we stated in the introduction.

Lessons on challenge

The first three lessons relate to research question 2: How can high ability students best be challenged?

1. *High-ability students at university level are challenged to hit the highest notes when a disbalance is created between the levels of perceived challenge and ability*.

As the second study showed, honors students' perceived learning was at its peak in the period in which the challenge level most exceeded the ability level. This finding contradicts more harmonious theories, such as Pekrun's (2006) control-value theory, Deci & Ryan's (1985) self-determination theory of motivation, and Csikszentmihalyi's (1975) flow model, all of which assume a balance between challenge and ability to be beneficial for effort and enjoyments, as well as for learning. Although the honors students in our case studies were challenged by a fairly large disbalance, we assume that a slight disbalance might be beneficial for all students.

Working at the edge of ones abilities is not always comfortable. While the disbalance between challenge and ability was found to be beneficial for learning, it negatively affected students' enjoyment, yet temporarily. This finding is unexpected, given that negative emotions have been found to impede motivation and learning (e.g. Alexander, Jetton, &

Kulikowich, 1995; Lepper & Henderlong, 2000; Renninger & Wozniak, 1985). Our first case study (Chapter 3) revealed however that students' enjoyment fluctuated: increasing when progress was made, decreasing when they were stuck, and finally feeling intense pride after successfully achieving the course demands. One of the students expressed the intricate relation between enjoyment, effort and learning as follows:

> The enjoyment was gone, particularly during this second low point (referring to his motivational storyline) I was thinking 'why for heaven's sake did I choose for this course?' I could have chosen a history course you know, just sit back and let it all come to you. [...] But these were also the weeks that I worked the hardest and learned the most.

Thus, incorporating motivation over time suggested that enjoyment and effort, which both are associated with motivation, developed in dynamic patterns in time.

2. High teacher expectations are a crucial challenge factor.
As explained in chapter 3 and 4, high expectations have a strong impact on students' perceptions of challenge. Autonomy and complexity were important factors as well, but without high expectations students tended to –what Hattie (2012) refers to as- the 'minimax principle': maximum grade return for minimal extra effort. The strong impact teacher expectations have on students is surely not a new finding but research on the influence of teacher expectations has been conducted mostly on primary school level (e.g. Babad, 1998; Brophy, 1983; Weinstein, 2002). In research on older students (Hattie, 2009; Rubie-Davies, Peterson, Irving, Widdowson, & Dixon, 2010), it is found that the influence of teacher expectations decreases as students become older and rely more on self-assessment. The results of the current thesis indicate that honors students on the tertiary level are still quite susceptible to the expectations of their teachers.

3. Challenge is found in the combination of the challenge factors rather than in the task.
Challenge did not germinate directly from the pedagogical methods or the character of the tasks, but rather from the combination of the three challenge factors. Students could experience maximum challenge in either a research assignment (chapter 3) or in a traditional course format (chapter 4). According to Innoue (2007), teachers more often see challenge in the task itself, whereas students tend to assess challenge by the efforts needed for completing the task. To successfully complete a task, the level of difficulty is more likely to increase the challenge than the design of the task. Specific assignments, such as research tasks, may inherently be challenging, but unless the student needs to invests in the task, it may not be challenging for them (Hattie, 2012).

Lessons on dilemma's
The fourth and fifth lessons relate to research question 3: What dilemmas do teachers experience when challenging high-ability students?

4. Challenging students may also be discomforting for teachers.

For teachers, it may be challenging to allow students to be discomforted, because students' negative feelings could be believed to affect other values, such as care for a safe atmosphere, or care for the interpersonal relation with students. Teaching involves "human nurturance, connectedness, warmth and love" (Hargreaves, 1994, p. 175), and each teacher's individual beliefs about their role in caring for students guide their actions. Kelchtermans and Ballet (2002) add that political interests and personal values shape teachers' emotions and function as a rationale for their professional actions. This coheres with MacLure's (1993) observation that teachers frequently use their identity or political belief system to justify the way they choose to engage in their work. From the point of view of effectiveness, increasing the challenge might be desirable. However, from the point of view of responsibility for a warm and caring student-teacher relationship, challenge might produce an atmosphere that is discomforting for teachers. In such a dilemma, the teacher needs to weigh carefully the advantages of challenge against the importance he or she attaches to the student-teacher relationship. The teachers participating in this study experienced these and other dilemmas frequently, indicating the complexity of teaching practice in general and of challenge in particular. For these teachers it might be good to know that the students participating in the casestudies often valued the pride following discomforting periods retrospectively. As one of the students said, referring to the course described in chapter 3: "This was the most challenging and the most rewarding course I ever had". A relation between high challenge, anxiety, and pride was suggested by Meyer and Turner (1998), who concluded that students in high challenge classes experienced anxiety and high pride afterwards, and students in low challenge classes experienced positive emotions during the courses, but lower pride afterwards.

With respect to the three challenge factors, having high expectations was particularly causing dilemmas for teachers. An explanation could be found in the fact that high expectations are directly communicated in interaction with students. It requires courage of a teacher to do that, because communicating high expectations, for example by asking difficult questions and waiting long for student answers could engender uncomfortable silences, embarrassment, or feelings of anxiety. Hattie (2012) observed the same phenomenon when he states that many teachers hesitate to make students struggle when creating challenge for them.

5. The teachers' considerations and the situational factors which lead to choices are equally informative for our understanding of the complexity of teaching as their actual choices are.

When facing a dilemma, based on the same considerations, different teachers might choose differently, depending on the way the benefits and losses are weighed against each other. The diversity of choices and considerations brought forward by the teachers (chapter 5) pointed out that it matters who the teacher is, what he or she feels responsible for, and is able to do. The importance of the self-understanding of the teacher is increasingly recognized in the field, and has been found to be a source for dilemmas to occur (Enyedy et al., 2006; Kelchtermans,

2005). Fundamentally, the moral dimension of teaching refers to what is in the best interest of students (Kelchtermans, 2009). Acting in the best interest of students is however not clear-cut, because there is no agreement about what is in the best interest in students, and moreover, what is good for some students might be detrimental for others in the same class.

Lesson on differences between honors and non-honors students

A final and sixth lesson addresses the first research question: Do honors students differ from non-honors students with respect to the qualities that have been found to be essential for exceptional accomplishments in professional life (intelligence, creativity, and motivation)?

6. Honors groups are heterogeneous.

Honors students at Utrecht University indeed differ significantly from non-honors students with respect to the characteristics needed to excel in professional life. However, a closer look at our data revealed that the honors group was nearly as heterogeneous as the non-honors group regarding these characteristics, in view of the fact that standard deviations of the honors group were only slightly lower than that of the non-honors group. Since honors students are selected, we had not expected the heterogeneity of these groups to be this large. The heterogeneity could explain the fact that teachers experienced the dilemma of whether to challenge students to the full and loose some students or lower the challenge in order to keep all students aboard. Further, in the introduction it was stated that a significant difference between honors and non-honors groups would justify differential education. The heterogeneity of both groups however indicates that grouping is not sufficient in order to match students' needs and abilities, and that adaptivity of teachers to the way students respond to the challenges is required.

6.2 Reflections on the methodology

In this section we reflect on methodological issues concerning the validity, reliability, and generalizability of the research.

Self-reports

This study relied primarily on self-reports, of students' perceptions of characteristics they possess, and of their perceptions of challenge. The data on differences between honors and non-honors students on motivation, creativity and intelligence, presented in chapter 2 were based on self-reports. Although this is a common and defensible methodology in its own right, self-report questionnaires could threaten the validity of this study. One threat is the general tendency for individuals to respond in socially desirable ways. There are however several reasons justifying the use of self-assessment for these three characteristics. First, the practical advantages of using self-assessment justify its use. Rather than running subjects one-by-one in a tightly supervised laboratory setting, researchers can administer such scales quickly to large groups of subjects. Moreover, self-report questionnaires are less threatening than IQ tests and therefore more likely to elicit cooperation. Second, measures of motivation are

typical self-assessed. With respect to creativity, there is no single, authoritative definition of creativity, nor is there a standardized agreed upon valid measurement technique. Moreover, self-assessed creativity was demonstrated to be a valid measure of creativity; the relationship of self-ratings to established individual differences is similar to that found with other better-known measures of creativity and creative potential such as the Divergent Thinking Test (Furnham, Batey, Anand, & Manfield, 2008; Furnham & Zhang, 2006). The most problematic measure is that of self-assessed intelligence. In studies using undergraduate samples, significant correlations between self-assessed intelligence and standardized IQ tests are found, usually between .30 and .50 (Chamorro-Premuzic, Furnham, & Moutafi, 2004; Furnham, Zhang, & Chamorro-Premuzic, 2006; Paulhus, Lysy, & Yik, 1998). Skeptics however argue that such self-reports are undermined by a variety of biases including self-deception and impression management, and that correlations with IQ tests are often low (Mabe & West, 1982). The unexpected lack of differences between honors and non-honors students in two of the study programs, discussed in section 6.1, indeed suggested that self-assessed intelligence was subjected to some bias.

In the second and third study (chapters 3 and 4), self-reports were considered appropriate. Central in these studies were the way students perceived challenge in the learning environment, also in relation to their perceived ability.

Retrospective reflections and storylines

Another possible threat to validity was the fact that the interviews with students (chapter 3 and 4), and teachers (chapter 5) took place after the course finished. This might be a problem, since we do not know whether the retrospective reflections reflect the experiences of students and teachers on-the-spot.

Concerning the student interviews, we tried to bridge the time distance by using the storyline method. The use of visual (re)presentation in social science research is not new, it has been used to encourage memories and stories about experience to be extended and elaborated (Orland, 2000; Sheridan, Chamberlain, & Dupuis, 2011). The storylines acted as memory aid, which facilitated students to think back to the actual moment. The emotions students expressed during drawing and talking, and students switching to the present form now and then when explaining their storylines furnished our impression that students indeed were drawn back to the actual moments in the course, thereby decreasing the difference between actual experiences and retrospective reflection.

The dilemmas we found in our fourth study (chapter 5) were drawn from interviews, which were also conducted after the courses were finished. These interviews elicited reflections on events, which happened during teaching. Conscious deliberations, however, are often not the primary driver of behavior; intuitions play a large role in the decisions people make (Haidt, 2001), which is why the teachers' reflections may not fully reflect their on-the-spot considerations.

Subjective bias

A common critique on case study research is its tendency to confirm the researcher's preconceived notions, so that the study therefore becomes biased (Flyvbjerg, 2006; Yin, 2009). This critique is guided by a natural tendency in humans' understanding to be more receptive for information confirming their prepositions than to information rejecting them. As a first response on the critique of subjectivity, we would like to emphasize that in qualitative research, it is generally accepted that subjective perception is involved (Hamberg, Johansson, Lindgren, & Westman, 1994). As Lincoln and Guba (1985, p. 217) say: 'the final judgment ...is ...vested in the person seeking to make the transfer'. The process of analyzing data is regarded as a reciprocal process, as the researcher uses his/her individual knowledge, perspectives, and experience for enhancing 'theoretical sensitivity' (Strauss, p.21). Though we do not claim objectivity, we tried to manage subjectivity by keeping an open mind and searching for intersubjectivity. For example, our preconceived assumptions about the factors that challenged students in the Advanced Cell Biology course (chapter 3), appeared to be incorrect. Our presupposition was, that the authenticity of the students' assignment, combined with the assessment by an external jury of experts would be the challenging factors of this course, both of which hypotheses we had to reject after studying the case in depth. Further, several verification strategies have been built into each phase of the research, acting as self-correcting mechanisms, to ensure the trustworthiness of the interpretations. First, additional sources of data were used to confirm the emerging findings; the phenomena we found in the interviews with students were checked against the class observation notes and course materials. Secondly, each step in the analysis process was discussed with the daily supervisor, examining the plausibility of the interpretation of the data. Third, final interpretations were checked by external auditors, ensuring inter-rater reliability, and fourth, simple counts of phenomena have been included where appropriate, aimed at guarding against anecdotalism (Silverman, 2013).

Generalizability

A second issue in relation to case study research is, whether you can generalize findings from a single case. Generalizability can be defined as the applicability of research findings and conclusions from a study conducted on a sample population to the population at large (Crossley, 2007). Because generalizability in the traditional sense requires data on large populations, findings of our case studies are by definition not generalizable, while the small groups and the restricted context of University College Utrecht (UCU) are not necessarily representative of the larger population. We tried to enhance generalizability in two ways, first, by choosing our first case carefully. For our second study, we chose an extreme case that was known as one of the most challenging ones at UCU, and therefore was an excellent case to study challenge. Second, in the next study, we obtained generalization of the challenge factors found in the first study in the experiences of students in six more typical UCU courses, varying with respect to challenge levels. Although the findings of the first study were

confirmed in the second study, data still were drawn from a relatively small sample in one institution. Hence, we can only speculate whether these findings could be generalized to other honors programs and colleges.

We think, however, there are strong reasons to believe that our findings concerning challenge might be 'transferable' (Lincoln & Guba, 1985) to comparable situations. While generalizability is applied by researchers, transferability is a process performed by the readers of research (Metcalfe, 2005). Unlike generalizability, transferability does not involve broad claims, but invites readers of research to make connections between elements of a study and their own experience (Barnes, 1994-2002). According to Berliner (2002, p. 19), implementing scientific findings is always difficult in education, "because humans in schools are embedded in complex and changing networks of social interaction". Therefore, we do not claim to have produced broadly generalizable findings; but instead invite the reader to identify how the findings can be transferred to their situation. As Biesta (2007, p. 16) stated, educational outcomes inform us about the relations between actions and their consequences in a particular situation, showing us "what worked, but cannot tell us what works". We assume however, that the concept of creating a disbalance between challenge and ability levels, by increasing the autonomy, complexity and expectations provides relevant insights for discourse on effective teaching of honors students.

With respect to the fourth study about teacher dilemmas, we recognize two limitations concerning the generalizability of the dilemmas. First, the sample was small and consisted of experienced college teachers. Second, all the teachers in this study were from one college, so our findings must be viewed as being situated in this specific context. Some of the considerations these teachers articulated when weighing their choices, such as the fear to affect the positive relationships with students, or the responsibility to comply with institutional regulations however, may have a wider resonance, as they concur with findings in literature on teacher dilemmas.

6.3 Implications

Suggestions for future research

A first area of interest is the relationship between the (interacting) talent characteristics of honors students and their success in professional domains. The results of the first study suggest that honors groups have the characteristics needed for excellent achievement in professional life to a larger extent than non-honors students, but we do not know whether they have these traits in sufficient measure and in successful combinations to allow for excellent achievement in future professional life. Selection of honors students is currently based on school performance and motivation, and not so much on creativity, curiosity, and openness. These creativity-related qualities differentiate between what Renzulli (1986) termed 'schoolhouse giftedness' and 'creative-productive giftedness'. The first group is expected to excel at school, while the creative-productive group is not particularly performing well at school, but is expected to make innovative contributions to their professional domain.

Longitudinal study of the achievements of honors and non-honors students up to their professional life, taking into account environmental factors, would therefore be informative for selection of honors students.

A second theme addresses the fact that the case studies were located in one institute, indicating that we can only speculate whether these findings could be generalized. We therefore recommend further research to discover whether the three-factor model of challenge (complexity, autonomy and high expectations) applies to challenging university courses in other honors programs.

Third, the dynamic aspect of students' motivation, found in the first case study (chapter 3) will be informative for future research. Students reported ongoing persistence and decreased enjoyment during over-challenging periods, indicating that the implications of emotions are not always predictable. For example, progress may feel good, but do these positive feelings result in greater or lower subsequent effort? Further, additional work is needed to determine whether these emotions have distinct implications for learning and motivation. Rather than thinking of motivation as a static state, the inclusion of a temporal component may help lead to models of motivation that include dynamics of motivation retrospectively and over time.

Fourth, given that students are responsive to challenging conditions in the learning environment, it might be oversimplifying to interpret our results as a one-directional effect of instructional behavior on the students. First, students actively construct challenge by interpreting their teachers' behavior, and, second, the engagement of the students also affects their teacher's motivation and instructional behavior (Pelletier et al., 2002; Reeve & Tseng, 2011). Specifically, the intelligent, creative, and motivated population of honors students has potential to enhance the learning environment when teachers allow them to (Heller, 1999). Heller proposed a role transition of teachers of high-ability students to that of co-learners, to stimulate students to contribute actively to the instructional flow (Heller, 1999), which could subsequently increase the teachers' engagement (Pelletier et al., 2002). Combining these two ideas we recommend further research to address challenge as an aspect of gifted education as it comes about during the process of continuous, responsive and anticipatory teacher-student interaction.

Fifth, teachers of honors programs at Utrecht University have been allowed quite a lot of freedom to experiment with their course designs and teaching methods, thereby using honors programs as test plot for educational innovation. The Advanced Cell Biology course we studied is an example of such an experiment. The expected by-effect of such experiments is that non-honors programs also benefit from these innovative insights, as Renzulli (1998) stated: "A rising tide lifts all ships". Therefore, it would be worthwile to explore to what extent the challenge factors (high expectations, complexity, and autonomy) would also be advantageous for non-honors students.

Sixth, the seven dilemmas we found in our fourth study were mentioned by twelve experienced teachers. It is relevant to find out whether less experienced teachers experience

similar dilemmas, also in light of teachers' professional development. The framework we developed for identifying dilemmas in teachers' reflections and for determining the urgency status could be used in such research.

Implications for practice

Factors increasing the challenge

A first inference we draw from our results for teaching high-ability students is the need to aim for a high level of complexity, especially level of thinking. Second, students can be challenged if teachers let go of controlling behavior and stimulate students to take responsibility for their learning. Third, students seem to need some reinforcement from their teachers, which can be found in the form of exams, though high teacher expectations and open assignments are preferred, because these are less detrimental to intrinsic motivation (Ryan & Deci, 1985).

Balancing over-challenge

In the case studies of honors courses (chapters 3 and 4), we averaged students´ storylines of perceived challenge and ability. Individual discrepancies between challenge and ability could however be even more extreme, causing a level of challenge that would pull students too far out of their comfort zones, which could in turn lead to too much stress and anxiety. According to the students, the support teachers offered their students and the peer support students offered each other in the first case prevented that happening in the courses we studied. Increasing challenge thus calls for an established support system.

The significance of effort and persistence

Considering the objective of honors programs to development of students toward excellence, it is important to design a learning environment in which the qualities needed for excellence are nurtured. In our first study, we found that honors students perceived themselves as more creative and intelligent than non-honors students, however, their motivation showed a differentiated picture. Honors students distinguished themselves in their desire to learn and drive to excel, while they did not diffe from non-honors students in persistence. Empirical studies have shown that talented professionals usually practice within their domain for several hours per day for a full decade before their latent capacity becomes actualized (Roe 1953; Ericsson 2006; Howe 1999; Walberg et al. 2003), suggesting that persistence is crucial for excellent professional achievement. The low distinguishing value we found for persistence could imply that these honors students were intelligent enough to achieve well in school without having to put much energy into their work. Chamorro-Premuzic and Arteche (2008) suggested that persistence and intelligence compensate for each other in terms of achievement. In other words, these students might not have learned to apply effort in order to achieve high grades in school, because of a lack of challenge. Therefore, offering consistent challenge could help these students understand the additional value of effort and persistence in high achievement.

Implementation: informed discourse instead of following recipes

While challenge is important to stimulate honors students' learning, effort, and persistence, the challenge factors we found are not intended as a recipe. As we have seen in the study on teacher dilemmas, teachers have more responsibilities than maximizing challenge. Therefore, the findings of this research are to be regarded as a source, providing teachers information, which they can use in the judgment of actual situations as they arise (Biesta, 2007; Dewey, 1929). In this way, Dewey (1929:21) states, the teacher's ability to judge is enriched, providing a wider range of alternatives to select from in dealing with individual situations. As professional judgment is central to educational practice (Biesta, 2007), and judgments in practical situations are made intuitively (Haidt, 2001), it would be sensible to develop professional intuition. To progress professional intuition, explicit learning from theory and expertise is needed (Salas, Rosen, & DiazGranados, 2010). One way to develop expertise could be found in facilitating collaborative reflection on dilemmas. Recent studies converge on the understanding that reflecting on dilemmas can have beneficial effects on teacher learning and change, since they provide teachers with the opportunity to examine and challenge their beliefs and to explore alternatives for change and growth (Helsing, 2007; Roblin & Margalef, 2012; Sunley & Locke, 2012; Yoon & Kim, 2010). Roblin and Margalef found that the acknowledgement of dilemmas enabled teachers to take a critical perspective on their educational beliefs and practice, thereby strengthening critical reflection. Reflecting collaboratively adds to the learning experience, since teachers differ in their experiences and beliefs, and these differences challenge each other's values, viewpoints, and actions (Yoon & Kim, 2010). To conclude, collaborative reflection on ways to implement the challenge factors and to decrease their negative consequences would help building teacher expertise.

References

Achterberg, C. (2005). What is an honors student? *Journal of the National Collegiate Honors Council*, 6(1), 75-81.

Albert, A. S. (1969). Genius: Present-day status of the concept and its implications for the study of creativity and giftedness. *American Psychologist, 24*(8), 743-753.

Alexander, P. A., Jetton, T. L., & Kulikowich, J. M. (1995). Interrelationship of knowledge, interest, and recall: assessing a model of domain learning. *Journal of Educational Psychology*, *87*, 559-575.

Anderson, L., & Krahtwohl, D. A. (2001). *Taxonomy for learning, teaching and assessing: A revision of bloom's taxonomy of educational objectives*. New York, NY: Longman.

Assor, A., Roth, G., & Deci, E. L. (2004). The emotional costs of perceived parental conditional regard: A self-determination theory analysis. *Journal of Personality, 72*, 47-87.

Babad, E. (1993). Teachers' differential behavior. *Educational Psychology Review, 5*(4), 347.

Baird, L .L. (1985). Do grades and tests predict adult accomplishment? *Research in Higher Education*, *23*(1), 3-85.

Bakker, A. B., & Demerouti, E. (2006). The job demands-resources model: State of the art. *Journal of Managerial Psychology, 22*, 309-328.

Bandura, A. (2000). Exercise of human agency through collective efficacy. *Current Directions in Psychological Science*, *9*, 75-78.

Barab, S. A., &. Plucker, J. A. (2002). Smart people or smart contexts? Cognition, ability, and talent development in an age of situated approaches to knowing and learning. *Educational Psychologist, 37*(3), 165-182.

Barnes, J., Conrad, K., Demont-Heinrich, C., Graziano, M., Kowalski, D., Neufeld, J., Zamora, J., & Palmquist, M. (1994 - 2012). *Generalizability and transferability*. Writing@CSU. Colorado State University. Available at http://writing.colostate.edu/guides/guide.cfm?guideid=65.

Barron, F., & Harrington, D. M. (1981). Creativity, intelligence and personality. *Annual review of Psychology, 32*, 439-476.

Batey, M., & Furnham, A. (2006). Creativity, intelligence, and personality: A critical review of the scattered literature. *Genetic, Social & General Psychology Monographs, 132*(4), 355-429.

Baumert, J. N, Lüdtke, O.R, Trautwein, U., & Brunner, M. (2009). Large-scale student assessment studies measure the results of processes of knowledge acquisition: Evidence in support of the distinction between intelligence and student achievement. *Educational Research Review, 4*, 165–176.

Becher, T., & Trowler, P. R. (2004). *Academic Tribes and Territories; Intellectual enquiry and the culture of disciplines* (2nd edn.). Buckingham: Open University Press.

Beijaard, D., van Driel, J., & Verloop, N. (1999). Evaluation of story-line methodology in research on teachers' practical knowledge. *Studies in Educational Evaluation*, *25*(1), 47-62.

Ben-Peretz, Miriam, & Kremer-Hayon, Lya (1990). The content and context of professional dilemmas. *Educational Review*, *42*(1), 31-41.

Benware, C., & Deci, E. L. (1984). The quality of learning with an active versus passive motivational set. *American Educational Research Journal*, *21*(4), 755-766.

Berlak, A., & Berlak, H. (1981). *Dilemmas of schooling: Teaching and social change*. New York, NY: Methuen.

Biesta, G. (2007). Why “what works” won’t work: Evidence based practice and the democratic deficit in educational research. *Educational Theory*, *57*(1), 1-22.

Biesta, G. (2012). The future of teacher education: Evidence, competence or wisdom? *Research on Steiner Education*, *3*(1), 8-22.

Biggs, J. B. (1989). Approaches to the enhancement of tertiary teaching. *Higher Education Research and Development, 8*, 7-25.

Biggs, J. (1999). What the student does: Teaching for enhanced learning. *Higher Education Research & Development, 8*(1), 1, 57-75.

Billig, M., Condor, S., Edwards, D., Gane, M., Middleton, D., & Radley, A. (1988). *Ideological dilemmas: A social psychology of everyday thinking*. London: Sage.

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. *Assessment in Education*, *5*(1), 7-74.

Boeije, H. (2010). *Analysis in qualitative research*. Sage: London.

Borland, J. H. (2003). Gifted education without gifted children. In R. J. Sternberg, J. E. Davidson (Eds.), *Conceptions of Giftednes (*pp. 1-19).

Bowen, G. A. (2006). Grounded theory and sensitizing concepts. *International Journal of Qualitative Methods*, *5*(3), 12-23.

Brislin, R. W. (1986). A culture general assimilator: preparation for various types of sojourns. *International Journal of Intercultural Relations,* 10, 215-234.

Brophy, J. E. (1983). Research on self-fulfilling prophecy and teacher expectations. *Journal of Educational Psychology*, *75*, 631-661.

Brophy, J., & Good, T. L. (1986). Teacher behavior and student achievement. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 340–370). New York: Macmillan Library.

Bruijn. E de & Leeman, Y. (2011). Authentic and self-directed learning in vocational education: Challenges to vocational educators. *Teaching and Teacher Education, 27*, 694-702.

Butler, R. (2012). Striving to connect: Extending an achievement goal approach to teacher motivation to include relational goals for teaching. *Journal of Educational Psychology 104* (3), 726-742.

Buzzelli, C, & Johnston, B. (2001). Authority, power, and morality in classroom discourse. *Teaching and Teacher Education, 17*, 873–884.

Cabaroglu, N. & Tillema, H. H. (2011): Teacher educator dilemmas: A concept to study pedagogy. *Teachers and Teaching: Theory and Practice*, *17*, 559-573.

Callahan, C. M. (2000). Intelligence and Giftedness. In R. J. Sternberg (Ed.), *Handbook of intelligence (pp.* 159-175) Cambridge: Cambridge University Press.

Carnicom, S. (2011). Honors education: Innovation or conservation? *Journal of the National Collegiate Honors Council,* Online Archive.Paper 321. http://digitalcommons.unl.edu/nchcjournal/321.

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

Chamorro-Premuzic, T., & Arteche, A. (2008). Intellectual competence and academic performance: Preliminary validation of a model. *Intelligence, 36*, 564-573.

Chamorro-Premuzic, T., & Furnham, A. (2005). *Personality and intellectual competence*. New York: Lawrence Erlbaum.

Chi, M. T. H. (2006). Two approaches to the study of expert' characteristics. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.) *The Cambridge handbook of expertise and expert performance (*pp. 683-703). New York: Cambridge University Press.

Chi, M.T.H. (1997): Quantifying qualitative analyses of verbal data: A practical guide. *Journal of the Learning Sciences*, *6*, 271-315.

Clark, L. (2000). A review of the research on personality characteristics of academically talented college students. In L. Clark , & C. Fuiks (Eds.), *Teaching and learning in honors* (pp. 7-20). Radford, VA: The National Collegiate Honors Council.

Clinkenbeard, P. (1994). Motivation and highly able students: resolving paradoxes. In J. F. Feldhusen, J. B., Hansen, & S. M. Hoover (Eds.), *Talent development: Theories and practice (pp*. 187–202). Dubuque, Iowa: Kendall/Hunt.

Clough, M. P., Berg, C. A., & Olson, J. K. (2009). Promoting effective science teacher education and science teaching: A framework for teacher decision-making. *International Journal of Science and Mathematics*, *7*, 821-847.

Coaldrake, P. (2001) Responding to changing student expectations. Higher Education Management, *13*(2), 75-92.

Cohen, P. A. (1985). College grades and adult achievement: A research synthesis. *Research in Higher Education, 20*(3), 281-293.

Collins, M. A. & Amabile, T. M. (1999). Motivation and creativity. In R.J. Sternberg (Ed.), *Handbook of creativity* (pp. 297-312). New York: Cambridge University Press.

Colnerud, G. (1997*)*. Ethical conflicts in teaching. *Teaching and Teacher Education*, *13*, 627–635.

Colnerud, G. (2006): Teacher ethics as a research problem: syntheses achieved and new issues. *Teachers and Teaching: Theory and Practice*, *12*, 365-385.

Cortina, J. M. (1993). What is coefficient Alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*, 98-104.

Costa, P. T. & McCrae, R. R. (1992). *NEO personality Inventory professional manual*. Odessa, FL: Psychological Assessment Resources.

Cox, C. M. (1926). *Genetic studies of genius; the early mental traits of three hundred geniuses*. Stanford: Stanford University Press.

Crossley, M. (2007). Narrative analysis. In E. Lyons & A. Coyle, *Analysing qualitative data in psychology* (pp. 131-144). Lonson: Sage.

Csikszentmihalyi, M. C. (1992). Motivation and creativity. In R. S. Albert (Ed.), *Genius & Eminence* (2nd edn., pp. 19-33). Pergamon Press, Oxford.

Csikszentmihalyi, M. (1975). *Beyond boredom and anxiety: Experiencing flow in work and play*. San Francisco: Jossey-Bass.

Csikszentmihalyi, M., Abuhamdeh, S., & Nakamura, J. (2005). Flow. In A. Elliot *(*Ed.), *Handbook of competence and motivation* (pp. 598–698). New York: The Guilford Press.

Csikszentmihalyi, M., Rathunde, K., & Whalen, S, (1993). *Talented teenagers: The roots of success and failure*. New York: Cambridge University Press.

Cushman, F. & Young, L. (2009). The psychology of dilemmas and the philosophy of morality. *Ethical Theory and Moral Practice 12*, 9-24.

Dai, D. Y., Moon, S. M., & Feldhusen, J. F. (1998). Achievement motivation and gifted students: A social cognitive perspective. *Educational Psychologist, 33*, 45-63.

Darling-Hammond, L. (1999). *Teacher quality and student achievement: A review of state policy evidence*. Seattle: Center for the Study of Teaching and Policy, University of Washington.

Davis, Ray J., & Montgomery, Soncerey L. (2011). Honors Education at HBCUs: Core Values, Best Practices, and Select Challenges. *Journal of the National Collegiate Honors Council --Online Archive*. Paper 300. http://digitalcommons.unl.edu/nchcjournal/300

De Boer, G. C., Minnaert, A. E. M. G. & Kamphof, G. (2013). Gifted education in the Netherlands. *Journal for the Education of the Gifted, 36*(1), 133-150.

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.

Deci, E. L., Koestner, R. & Ryan, R.M. (1999). A meta-analytic review of experiments examining the effects of extrinsic rewards on intrinsic motivation. *Psychological Bulletin, 125* (6), 627-668.

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry: An International Journal for the Advancement of Psychological Theory, 11*(4), 227-268.

Denicolo, P. (1996). Productively confronting dilemmas in educational practice and research. In M. Kompf, R. Bond, D. Dworet, T. Boak (Eds.), *Changing research and practice: Teachers' professionalism, identities and knowledge* (pp. 56–65). London: Falmer Press.

Detterman, D. K., & Ruthsatz, J. (1999). Toward a more comprehensive theory of exceptional abilities. *Journal for the Education of the Gifted*, *22*(2), 148–158.

Dewey, J. (1929). *The sources of a science of education*. New York: Horace Liveright

Dweck, C. S. & Leggett, E. L. (1988) A social-cognitive approach to motivation and personality. *Psychological Review, 95*, 256-273.

Eddles-Hirsch, K., Vialle, W., Rogers, K. B., & McCormick, J. (2010). "Just challenge those high-ability learners and they'll be all right!". The impact of social context and challenging instruction on the affective development of high-ability students. *Journal of Advanced Academics, 22*, 106-128.

Edmondson, A. (1999). Psychological safety and learning behavior in work teams. *Administrative Science Quarterly*, *44*, 350-383.

Elliot, A., & H. McGregor (2001). A 2 x 2 achievement goal framework. *Journal of Personality and Social Psychology, 80*, 501-519.

Engeser, S., & Rheinberg, F. (2008). Flow, performance and moderators of challenge-skill balance. *Motivation and Emotio*n, *32*, 158–172.

Engestrom, Y. & A. Sannina (2011). Discursive manifestations of contradictions in organizational change efforts: A methodological framework. *Journal of Organizational Change Management, 24*(3), 368-387.

Enyedy, N., Goldberg, J., & Welsh, K. (2006). Complex dilemmas of identity and practice. *Science Education, 90*(1), 68-93.

Ericsson, K. A. (2006). The influence of experience and deliberate practice on the development of superior expert practice. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge handbook of expertise and expert performance* (pp. .683-703). New York: Cambridge University Press.

Ericsson, K. A., Roring, R. W., & Nandagopal, K. (2007). Giftedness and evidence for reproducibly superior performance: An account based on the expert performance framework. *High-ability Studies*, *18*(1), 3-56.

Feist, G. J. (1998). A meta-analysis of personality in scientific and artistic creativity. *Personality & Social Psychology Review,* 2(4), 290-309.

Field, Andy (2009). *Discovering statistics using SPSS*. London: Sage Publications Ltd.

Flyvbjerg, B. (2006). Five misunderstandings about case-study research. *Qualitative Inquiry*, *12* (2), 219-245.

Fransson, G. & Grannäs, J. (2013). Dilemmatic spaces in educational contexts – towards a conceptual framework for dilemmas in teachers work. *Teachers and Teaching, 19*, 4-17.

Freeman, J. (1998). *Educating the very able: Current international research*. The Stationery Office: London.

Freeman, J. (1990). The intellectually gifted adolescent. In M. J. A. Howe (Ed.), *Encouraging the development of exceptional skills and talents* (pp. 89–108). Leicester: British Psychological Society.

Frenzel, A. C., Goetz, T., Lüdtke, O., Pekrun, R., & Sutton, R. (2009). Emotional transmission in the classroom: Exploring the relationship between teacher and student enjoyment. *Journal of Educational Psychology*, *101*, 705–716.

Friedman-Nimz, R. & Skyba, O. (2009). Personality qualities that help or hinder gifted and talented individuals. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 421-435). The Netherlands: Springer Science and Business Media B. V.

Gagné, F. (1993) Constructs and models pertaining to exceptional human abilities, in: K. A. Heller, F. J. Mönks & A. H. Passow (Eds), *International handbook of research and development of giftedness and talent* (pp. 69–88). Oxford, Pergamon Press.

Gagné, F. (1995). From giftedness to talent: a developmental model and its impact on the language of the field. *Roeper Review, 18*(2), 103-111.

Gallagher, J., Harradine, C. C., & Coleman, M. R. (1997). Challenge or boredom? Gifted students' views on their schooling. *Roeper Review, 19*(3), 132-136.

Gentry, M., Gable, R. K., & Springer, P. (2000). Gifted and nongifted middle school students: Are their attitudes toward school different as measured by the new affective instrument, My Class Activities ...? *Journal for the Education of the Gifted, 24*(1), 74-96.

George, D., & Mallery, P. (2003). *SPSS for Windows step by step: A simple guide and reference. 11.0 update* (4th edn.). Boston: Allyn & Bacon.

Gerris, J. R. M., Houtmans, M. J. M., Kwaaitaal-Roosen, E. M. G., de Schipper, J. C., Vemulst, A. A., & Janssens, J. M. A. M. (1998). *Parents, adolescents, and young adults in Dutch families: a longitudinal study*. Nijmegen: Institute of Family Studies, University of Nijmegen.

Ginkel, J. R. v. & Ark, L. A. van der (2008). *SPSS Syntax for two way imputation of missing test data.* Tilburg University.

Goertzel, V., Goertzel, M., Goertzel, T. & Hansen, A. (2004). *Cradles of eminence: Childhoods of more than 700 famous men and women* (2nd edn.). Scottsdale: Great Potential Press.

Goldberg, L. R. (1999). A broad-bandwidth, public-domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, , I. Deary, , F. DeFruyt, & Ostendorf, F. (Eds.), *Personality psychology in Europe* (pp. 7, 7–28). The Netherlands: Tilburg University Press.

Goncalo, J. A., Polman, E., & Maslach, C. (2010). Can confidence come too soon? Collective efficacy, conflict and group performance over time. *Organizational Behavior and Human Decision Processes, 113*, 13-24.

Gottfredson, L. S. (1997). Why g matters: the complexity of everyday life. *Intelligence, 24*(1), 79-132.

Gottfried, A. W., Eskeles Gottfried, A., Cook, C. R., & Morris, Ph. E. (2005). Educational characteristics of adolescents with gifted academic intrinsic motivation: A longitudinal investigation from school entry through early adulthood. *Gifted Child Quarterly, 49*(2), 172-186.

Gruber, H., & Mandl, H. (2000). Instructional psychology and the gifted. In K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds.*), International handbook of giftedness* (pp. 383–96). Oxford: Elsevier Science.

Haar, S. van der, Segers, M., & Jehn, K. A. (2013). Towards a contextualized model of team learning processes and outcomes. *Educational Research Review*, *10*, 1-12.

Haidt, J. (2001). The emotional dog and its rational tail: A social intuitionist approach to moral judgment. *Psychological Review*, *108*(4), 814–834.

Hamberg, K., Johannson, E., Lindgren, G., & Westman, G. (1994). Scientific rigour in qualitative research – Examples from a study of womens health in family practice, *Family Practice*, *11*(2), 176-181.

Harackiewicz, J. M., Barron, K. E., Tauer, J. M., Carter, S. M., & Elliot, A. J. (2000). Short-term and long-term consequences of achievement goals: Predicting interest and performance over time. *Journal of Educational Psychology, 92*(2), 216-230.

Hargreaves, A. (1994). *Changing teachers, changing times: Teachers' work and culture in the postmodern age*. London: Cassell.

Hattie, J C. (2009). *Visible learning: A synthesis of over 800 meta-analyses relating to achievement*. London and New York: Routledge.

Hattie, J C. (2012). *Visible learning for teachers; maximizing impact on learning*. London and New York: Routledge.

Hébert, T. P. & McBee, M. T. (2007). The Impact of an undergraduate honors program on gifted university students. *Gifted Child Quarterly, 51*(2), 136-151.

Heilbronner, N. N., Connel, E. E., & Reis, S. W. (2010). The "stepping stone phenomenon": Exploring the role of positive attrition at an early college entrance program. *Journal of Advanced Academics, 21*(3), 392-425.

Heller, K. A. (1999). Individual (learning and motivational) needs versus instructional conditions of gifted education [1]. *High-ability Studies, 10*(1), 9-21.

Heller, K. A. (2005). Education and counseling of the gifted and talented in Germany. *International Journal for the Advancement of Counselling*, *27*(2), 191-210.

Heller, K. A. (2007). Individual (learning and motivational) needs versus instructional conditions of gifted education. *High-ability Studies*, *10*(1), 9-21.

Heller, K. A., Mönks, F. J., & Passow, A. H. (Eds.), (1993). *International Handbook of Research and* Development. Oxford: Perganmon.

Heller, K. A., Perleth, C., & Lim, T. K. (2005). The Munich model of giftedness designed to identify and promote gifted students. In R. J. Sternberg, & J. W.Davidson (Eds.), *Conceptions of giftedness (pp*. 147-170). New York, Cambridge University Press.

Helsing, D. (2007) Regarding uncertainty in teachers and teaching. *Teaching and Teacher Education* 23, 1317–1333.

Hembree, R. (1988). Correlates, causes, effects, and treatment of test anxiety. *Review of Educational Research*, 58(1), 47-77.

Hennessey, B. A., & Amabile, T.M. (2010). Creativity. *The Annual Review of Psychology,* 61, 569-598.

Henze, I., van Driel, J. H. , & Verloop, N. (2009). Experienced science teachers' learning in the context of educational innovation. *Journal of Teacher Education,* 60, 184-199.

Hertberg-Davis, H., & Callahan, C. M. (2008). A narrow escape: Gifted students' perceptions of advanced placement and international baccalaureate programs. *Gifted Child Quarterly*, 52, 199-216.

Hoekman, K., McCormick, J., & Gross, M.U.M. (1999). The optimal context for gifted students: A preliminary exploration of motivational and affective considerations. *Gifted Child Quarterly, 43*, 170-193.

Howe, M. J. A. (1999). *Genius explained*. Cambridge: Cambridge University Press.

Hoyt, D. P. (1966). College grades and adult accomplishment: A review of research. *The Educational Record, 47*, 70-75.

Jackson, S. A., & Roberts, G. C. (1992). Positive performance states of athletes: Toward a conceptual understanding of peak performance. *The Sport Psychologi*st, *6*, 156–171.

Järvela, S., Järvenoja, H., & Veermans, M. (2008). Understanding the dynamics of motivation in socially shared learning. *International Journal of Educational Research*, *47*, 122–135.

Kaczvinsky, D. P. (2007). What is an honors student? A Noel-Levitz survey. *Journal of the National Collegiate Honors Council, 8*(2), 87-95.

Kahn, W. A. (1990). Psychological conditions of personal engagement and disengagement at work. *Academy of Management Journal*, *33*, 692–724.

Kahneman, D. (2011). *Thinking fast and slow*. Great Britain: Allen Lane.

Kanevsky, L., & Keighly, T. (2003). To produce or not to produce? Understanding boredom and the honor in underachievement. *Roeper Review, 26*(1), 20-28.

Kelchtermans, G. (2005). Teachers' emotions in educational reforms: self-understanding, vulnerable commitment and micropolitical literacy. *Teaching and teacher education, 21* (8), 995-1006.

Kelchtermans, G. (2009). Who I am in how I teach is the message. Self-understanding, vulnerability and reflection. *Teachers and Teaching: Theory and Practice*, *15*, 257–272.

Kelchtermans, G., & Ballet, K. (2002). The micropolitics of teacher induction: A narrative-biographical study on teacher socialisation. *Teaching and Teacher Education*, *18*, 105–120.

Keller, Johannes, & Bless, Herbert (2008). Flow and regulatory compatibility: An experimental approach to the flow model of intrinsic motivation. *Personality and Social psychology Bulletin, 34*, 196-209.

Kelly, K. E. (2004). A brief measure of creativity among college students. *College Student Journal, 38*, 594-596.

Klein, B., Rossin, D., Guo, Y., & Ro, Y. (2010). An examination of the effects of flow in learning in a graduate-level introductory operations management course. *Journal of Education for Business, 85*, 292-298.

Kolb, K. J., & Jussim, L. (1994). Teacher expectations and underachieving gifted children. *Roeper Review,* 17(1): 26-30.

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*(1), 148-161.

Kuncel, N.R., Ones, D.S., & Sackett, P. R. (2010). Individual differences as predictors of work, educational, and broad life outcomes. *Personality and Individual Differences*, *49*, 331–336

Lampert, M. (1985). How do teachers manage to teach? Perspectives on problems in practice. *Harvard Educational Review*, *55*, 178– 194.

Landis, J.R., & Koch, G.G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Langford, B. E., Schoenfeld, G., & Izzo, G. (2002). Nominal grouping sessions vs focus groups. *Qualitative Market Research: An International Journal, 5*(1), 58-70.

Larson, R. (1988). Flow and writing. In M. Csikszentmihalyi, and S. I. Csickszentmyhalyi (Eds.), *Optimal experience; psychological studies of flow in consciousness* (pp. 150-17). Cambridge: Cambridge University Press.

Lens, W., & Rand, P. (2000). Motivation and cognition: Their role in the development of giftedness. In K. A. Heller, F. J. Mönks, R. J. Sternberg, and R. F. Subotnik (Eds.), ,*International Handbook of Giftedness and Talent* (2nd edn., pp. 193-202). Oxford: Elsevier Science Ltd.

Lepper, M. R., & Henderlong, J. (2000). Turning "play" into "work" and "work" into "play": 25 years of research on intrinsic versus extrinsic motivation. In C. Sansone, & J. M. Harackiewicz (Eds.), *Intrinsic and extrinsic motivation: The search for optimal motivation and performance* (pp. 257-307). San Diego, CA: Academic.

Lincoln, Y.S., & Guba, E.G. (1985). Naturalistic inquiry, London: Sage publications Inc.

Linnenbrink, E. A. (2006). Emotion research in education: Theoretical and methodological perspectives on the integration of affect, motivation, and cognition. *Educational Psychology Review*, *18*, 307–314.

Linnenbrink, E. A., & Pintrich, P. R. (2002). Achievement goal theory and affect: An asymmetrical bidirectional model. *Educational Psychologist,* 37(2), 69-78.

Long, A., & Mullins, D. (2012). Honors around the globe. *Journal of the National Collegiate Honors Council, 13*(2).

Long, E. C. J. & Lange, S. (2002). An exploratory study: a comparison of honors and nonhonors students. *The National Honors Report, 23*(1), 20-30.

Long, J. F., & Hoy, A. W. (2006). Interested instructors: A composite portrait of individual differences and effectiveness. *Teaching and Teacher Education*, *22*, 303–314.

Lubinski, D. (2004). Introduction to the special section on cognitive abilities: 100 years after spearman's (1904) "'General intelligence,' objectively determined and measured". *Journal of Personality and Social Psychology* , *86*(1), 96-111.

Lubinski, D., & Benbow, C. P. (2000). States of excellence. *American Psychologist, 55*(1), 137-150.

Lubinski, D., Benbow, C. P., Webb, R. M., & Bleske-Reschek, A. (2006). Tracking exceptional human capital over two decades. *Psychological Science, 17*(3), 193-199.

Lyons, N. (1990). Dilemmas of knowing: ethical and epistemological dimensions of teacher's work and development. *Harvard Educational Review*, *60*, 159- 181.

Lyons, L. L., Freitag, P. K., & Hewson, P. W. (1997). Dichotomy in thinking, dilemma in actions: Researcher and teacher perspectives on a chemistry teaching practice. *Journal of Research in Science Teaching*, *34*, 239–254.

Mabe, P. A., & West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of Applied Psychology*, *67*(3), 280-296.

MacLure, M. (1993). Arguing for yourself: Identity as an organising principle in teachers' jobs and lives. *British Educational Research Journal*, *19*(4), 311–322.

Malpass, John R., O'Neil, Harold F., & Hocevar Jr., Dennis (1999). Self-regulation, goal orientation, self-efficacy, worry, and high-stakes math achievement for mathematically gifted high school students. *Roeper Review, 21*(4), 281-288.

Marra, R., & Palmer, B. (2004). Encouraging intellectual growth: Senior college student profiles. *Journal of Adult Development, 11(2),111-122*.

Marsh, H. W. (1987). The big-fish-little-pond effect on academic self-concept. *Journal of educational Psychology, 79*(3), 148-161.

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

McKeachie, W. J. (1986). *Teaching Tips* (8th edn.). Lexington, Mass.: Heath.

Mehan, H. (1979). *Learning lessons: Social organization in the classroom*. Cambridge, MA: Harvard University Press.

Meijer, P. C., de Graaf, G., & Meirink, J. (2011). Key experiences in student teachers' development. *Teachers and Teaching: theory and practice, 17*(1), 115–129.

Meyer, D. K. & Turner, J. C. (2002). Discovering emotion in classroom motivation research. *Educational Psychologist*, *37*(2), 107–114.

Meyer, D. K., & Turner, J. C. (2006). Re-conceptualizing emotion and motivation to learn in classroom contexts. *Educational Psychological Review, 18*, 377–390.

Milgram, R. M., & Hong, E. (1993). Creative thinking and creative performance in adolescents as predictors of creative attainments. *Roeper Review, 15*(3), 135.

Ministerie van Onderwijs, Cultuur en Wetenschap (OCW), Platform Bèta Techniek. (2008). *Sirius Program: Promoting Excellence in Dutch Higher Education*. Retrieved from http://www.siriusprogramma.nl/?pid=184.

Mönks, F. J. & Katzko, M. W. (2005). Giftedness and Gifted Education. In R. J. Sternberg and J. W. Davidson (Eds.), *Conceptions of Giftedness* (pp.187-200). New York: Cambridge University Press.

Mönks, F. J. & Mason, E. J. (2000) Developmental psychology and giftedness: theories and research. In: K. A. Heller, F. J. Mönks, R. J. Sternberg & R. F. Subotnik (Eds), *International handbook of giftedness and talent* (pp. 587–594). Oxford: Elsevier Science.

Nakamura, J., & Csikszentmihalyi, M. (2005). The concept of flow. In C. R. Snyder and S. J. Lopez (Eds.), *Handbook of positive psychology (*pp. 89–105). Oxford: Oxford University Press.

National Collegiate Honors Council (2012). Honors around the globe [special issue]. *Journal of the National Collegiate Honors Council*, *3*(2).

Nicol, D. J., & Macfarlane,-D. D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education,* 31(2), 199-218.

Noble, K. D., & Childers, S. (2008). A passion for learning: The theory and practice of optimal match at the University of Washington. *Journal of Advanced Academics*, *19*, 236-270.

Noddings, N. (1992). *The challenge to care in schools: An alternative approach to education*. New York: Teachers College Press.

Nunnaly, J. (1978). *Psychometric theory*. New York: McGraw-Hill.

Orland, L. (2000). What's in a Line? Exploration of a research and reflection tool. *Teachers and Teaching: Theory and Practice, 6*(2), 198-213.

Osborne, M. D. (1997). Balancing individual and the group: A dilemma for the constructivist teacher. *Curriculum Studies*, 29, 183-196.

Oser, F. K. (1994). Moral perspectives on teaching. *Review of Research in Education*, *20*, 58-127.

Pajares, F. (1996). Self-efficacy beliefs in achievement settings. *Review of Educational Research*, *66*, 543-578.

Pajares, F., and Urdan, T. (1996). Exploratory factor analysis of the Mathematics Anxiety Scale. *Measurement and Evaluation in Counseling and Development, 29*, 35-47.

Pareja Roblin, N. & Margalef, L. (2012). Learning from dilemmas: Teacher professional development through collaborative action and reflection. *Teachers and Teaching: Theory and Practice*, DOI:10.1080/13540602.2013.744196.

Park, G, Lubinski, D. & Benbow, C. P. (2008). Ability differences among people who have commensurate degrees matter for scientific creativity. *Psychological Science,* 19(10), 975-962.

Pascarella, E. & Terenzini, P. (1991). *How college affects students: Findings and insights from twenty years of research*. San Francisco: Jossey-Bass.

Pekrun, R. (2005). Progress and open problems in educational emotion research. *Learning and Instruction, 15*(5), 497-506.

Pekrun, R. (2006). The control-value theory of achievement emotions: Assumptions, corollaries, and implications for educational research and practice. *Educational Psychological Review, 18*, 315-341.

Pekrun, R., Goetz, T., & Titz, W. (2002). Academic emotions in students' self-regulated learning and achievement: A program of qualitative and quantitative research. *Educational Psychologist, 37*(2), 91–105.

Pelletier, L. G., Seguin-Levesque, C., & Legault, L. (2002). Pressure from above and pressure from below as determinants of teachers' motivation and teaching behaviors. *Journal of Educational Psychology*, *94,* 186-196.

Peterson, C., & Seligman, M. E. P. (2004). *Character strengths and virtues: A handbook and classification*. New York: Oxford University Press/Washington, DC: American Psychological Association.

Phillips, N., & Lindsay, G. (2006). Motivation in gifted students. *High-ability Studies ,1*, 57-73.

Pintrich, P. (1999). The role of motivation in promoting sustained self-regulated learning. *International Journal of Educational Research, 31*(6), 459-470.

Pintrich, P. R., & Zusho, A. (2007). Student motivation and self-regulated learning in the college classroom. In P. P. Raymond, & J. C. Smart (Eds), *The scholarship of teaching and learning in Higher Education: An evidence-based perspectiv*e (pp. 731-81). The Netherlands: Springer.

Pope, N., Green, S.K., Johnson, R.L., & Mitchelle, M. (2009). Examining teacher ethical dilemmas in classroom assessment. *Teaching and Teacher Education*, *25*, 778–782.

Pyryt, M. C., Masharov, Y. P., & Feng, C. (1993). Programs and strategies for nurturing talents/gifts in science and technology. In K.A. Heller, F.J. Monks, and A.H. Passow (Eds.), *International Handbook of Research and Development of Giftedness and Talent (*pp. 453-471). New York: Pergamon.

Reeve, J., Bolt, E., & Cai, Y. (1999). Autonomy-supportive teachers: How they teach and motivate students. *Journal of Educational Psychology*, *91*(3), 537-548.

Reeve, J., & Tseng. C.M. (2011). Agency as a fourth aspect of students' engagement during learning activities. *Contemporary Educational Psychology, 36*, 257–267.

Reis, S. M., & Renzulli, J. S. (2010). Is there still a need for gifted education? An examination of current research. *Learning and Individual Difference, 20*: 308-317.

Renninger, K. A. (2009) Interest and identity development in instruction: An inductive model. *Educational Psychologist*, 44(2), 105–118.

Renninger, K. A., & Wozniak, R. H. (1985). Effect of interest on attentional shift, recognition, and recall in young children. *Developmental Psychology*, *21*, 624-632.

Renzulli, J. S. (1986). The three-ring conception of giftedness. In R. J. Sternberg, & J. E. Davidson (Eds.), *Conceptions of giftedness* (pp. 53–92). New York: Cambridge University Press.

Renzulli, J. S. (1998). A rising tide lifts all ships: Developing the gifts and talents of all students. *Phi Delta Kappan, 80*(2),104.

Renzulli, J. S. (2003). The three-ring conception of giftedness: Its implications for understanding the nature of innovation. In L. V. Shavinina (Ed.) *The international handbook on innovation* (pp. 79-95). Oxford: Elsevier.

Reumer, C., & van der Wende, M. (2010). Excellence and diversity: The emergence of selective admission policies in Dutch higher education -A case study on Amsterdam University College. *Research & Occasional Paper Series*: *CSHE.15.10*

Rimm, S. (2002). Peer pressures and social acceptance of gifted students. In M. Neihart, S. M. Reis, N. M. Robinson, & S. M. Moon (Eds.). *The social and emotional development of gifted children: What do we know?* (pp. 13–18). Waco, TX: Prufrock Press.

Rindermann, H., & Neubauer. A.C. (2004). Processing speed, intelligence, creativity, and school performance: Testing of causal hypotheses using structural equation models. *Intelligence*, *32*(6), 573-589.

Rinn, A. N. (2007). Effects of programmatic selectivity on the academic achievement, academic self-concepts, and aspirations of gifted college students. *Gifted Child Quarterly, 51*(3), 232-245.

Rinn, A. N. & Plucker, J. A. (2004). We recruit them, but then what? The educational and psychological experiences of academically talented undergraduates. *Gifted Child Quarterly, 48*(1), 54-67.

Robinson, A., & Clinkenbeard, P. R. (1998). Giftedness, exceptionality examined. *Annual Review of Psychology, 48*, 117-139.

Robinson, N. M. (1997). The role of universities and colleges in educating gifted undergraduates. *Peabody Journal of Education, 72*, 217-236.

Robinson, N. M. (2003). Two wrongs do not make a right: Sacrificing the needs of gifted students does not solve society's unsolved problems. *Journal for the Education of the Gifted, 26*, 251-273.

Roe, A. (1953). *The making of a scientist*. New York: Dodd, Mead.

Rogers, K. B. (2007). Lessons learned about educating the gifted and talented: A synthesis of the research on educational practice. *Gifted Child Quarterly*, *51*(4), 382-396.

Rubie-Davies, C. M., Hattie, J. A. C., Townsend, M. A. R., & Hamilton, R. J. (2007). Aiming high: Teachers and their students. In N. Galwye (Ed.), *Progress in educational psychology research* (pp. 65-91). Hauppauge, NY: Nova Publishers.

Rubie-Davies, C., Peterson, E., Irving, E., Widdowson, D., & Dixon, R. (2010). Expectations of achievement student, teacher and parent perceptions. *Research in Education*, *83*, 36-53.

Ryan, R. M., & Deci, E. L. (2006). Self-Regulation and the problem of human autonomy: Does psychology need choice, self-determination, and will? *Journal of Personality, 74*, 1557-1585.

Ryan, R. M., & Deci, E. L. (1987). The support of autonomy and the control of behavior. *Journal of Personality and Social Psychology, 53*(6), 1024-1037.

Ryan, R. M., Koestner, E. L., & Deci, E. L. (1991). Ego-involved persistence: When free-choice behavior is not intrinsically motivated. *Motivation and Emotion, 15*(3), 185-205.

Ryan, R. M., & Stiller, J. (1991). The social contexts of internalization: Parent and teacher influences on autonomy, motivation and learning. In P. R. Pintrich & M. L. Maehr (Eds.), *Advances in motivation and achievement ,7* (pp. 115–149. Greenwich, CT: JAI Press.

Sadler, D.R. (1989). Formative assessment and the design of instructional systems. *Instructional Science, 18*, 119-144.

Salas, E., Rosen, M. A., & DiazGranados, D. (2010). Expertise-based intuition and decision making in organizations. *Journal of Management*, *36*, 941-973.

Sansone, C., & Thoman, D. B. (2005). Interest as the missing motivator in self-regulation. *European Psychologist*, *10*, 175-186.

Sayler, M. F. (2009). Gifted and thriving: A deeper understanding of the meaning of GT. In L. Shavinina (Ed), *The international handbook on giftedness* (pp. 215–230). Amsterdam: Springer Science & Business Media.

Scager, K., Akkerman, S. F., Pilot, A., & Wubbels, T. (2012). Challenging high-ability students. *Studies in Higher Education*, DOI:10.1080/03075079.2012.743117.

Scager, K., Akkerman, S. F., Pilot, A., & Wubbels, T. (2013). How to persuade honours students to go the extra mile; Creating a challenging learning environment. *High Ability Studies,* DOI: 10.1080/13598139.2013.941092.

Schick,H., & Phillipson, S. N. (2009). Learning motivation and performance excellence in adolescents with high intellectual potential: What really matters? *High-ability Studies, 20*(1), 15- 37.

Schweizer, K., Stefan J. Troche, S. J., & Rammsayer, T. H. (2011). On the special relationship between fluid and general intelligence: New evidence obtained by considering the position effect. *Personality and Individual Differences, 50*, 1249-1254.

Sederberg, P. C. (2008). Introduction. In C. Sederberg Peter (Ed.), *The honors college phenomenon*. NCHC monographs series. Lincoln: The University of Nebraska Press.

Selby, E. C., Shaw, E. J., & Houtz, J.C. (2005). The creative personality. *Gifted Child Quarterly, 49*(4), 300-314.

Shapira-Lishchinsky, O. (2011). Teachers' critical incidents: Ethical dilemmas in teaching practice. *Teaching and Teacher Education*, *27*, 648–656.

Shavinina, L. V. (2009). Innovation education for the gifted: A new direction in gifted education. In Shavinina, L. V. (Ed.). *International handbook on giftedness* (pp. 1257-1267). New York: Springer.

Sheridan, J., Chamberlain, K. & Dupuis, A. (2011). Timelining: visualizing experience. *Qualitative Resear*ch *11*(5), 552–569.

Shernoff, D. J., Csikszentmihalyi, M., Schneider, B., & Shernoff, E.S. (2003). Student engagement in high school classrooms from the perspective of flow theory. *School Psychology Quarterly, 18*, 158–176.

Shore, B. M., & Kanevsky, L. S. (1993). Thinking processes: Being and becoming gifted. In H. A. Heller, F. J. Mönks & A. H. Passow (Eds.), *International handbook of research and development of giftedness and talent* (pp. 133–47). New York: Pergamon.

Seifert, T. A., Pascarella, E. T., Colangelo, N., & Assouline, S. G. (2007). The effects of honors program participation on experiences of good practices and learning outcomes, *Journal of College Student Development*, *48*(1), 57-74.

Silverman, D. (2013). Doing qualitative research. London: Sage.

Simonton, D. K. (1988). *Scientific genius: A psychology of science*. New York: Cambridge University Press.

Simonton, D. K. (1994). *Greatness, who makes history and why?* New York: The Guilford Press.

Simonton, D. K. (1997). Creative productivity: A predictive and explanatory model of career trajectories and landmarks. *Psychological Review, 104* (1), 66-89.

Simonton, D. K. (1999). Significant samples: The psychological study of eminent individuals. *Psychological Methods, 4*(4), 425-51.

Simonton, D. K. (2000). Creativity: Cognitive, personal, developmental, and social aspects. *American Psychologist, 55*(1), 151-158.

Snow, R. E. (1986). Individual differences and the design of educational programs. *American Psychologist, 41*(10), 1029-1039.

Snow, R. E., & Swanson, J. (1992). Instructional psychology: Aptitude, adaptation, and assessment. *Annual Review of Psychology*, *43*, 583-626.

Sontag, C., Stoeger, H., & Harder, B. (2012). The relationship between intelligence and the preference for self-regulated learning: A longitudinal study with fourth-graders. *Talent Development & Excellence*, *4*(1), 1-22.

Stake, J. E., & Mares, K. R. (2001). Science enrichment programs for gifted high school girls and boys: Predictors of program impact on science confidence and motivation. *Journal of Research in Science Teaching, 38*, 1065–1088.

Sternberg, R. J. (2001). Giftedness as developing expertise: A theory of the interface between high abilities and achieved excellence. *High-ability Studies, 12*(2), 159-179.

Sternberg, R. J. (2003). WICS as a model of giftedness. *High-ability Studies, 14*(2), 109-137.

Sternberg, R. J. (2005). The WICS model of giftedness. In R. J. Sternberg, & J. E. Davidson, (Eds.) Conceptions of giftedness (2nd edn., pp., 327–342). New York: Cambridge University Press.

Sternberg, R. J., & Lubart, T. I. (1993). Creative giftedness: A multivariate investment approach. *Gifted Child Quarterly*, *37*(1), 7-15.

Strauss, A. (1987). *Qualitative analysis for social scientists*. New York: Cambridge University press.

Sumsion, J. (2000). Caring and Empowerment: A teacher educator's reflection on an ethical dilemma. *Teaching in Higher Education, 5*, 167-179.

Sunley, R. & Locke, R. (2012) Educational professionals' values: voices from secondary schools in England. *Educational Research*, *54*(3), 285-307.

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics* (5th edn.). Boston: Allyn & Bacon.

Tannenbaum, A. J. (1983). *Gifted children: Psychological and educational perspectives*. New York: Macmillan.

Tannenbaum, A. J. (1996). The IQ controversy and the gifted. In C. P. Benbow, & D. Lubinski (Eds.) *Intellectual talent* (pp. 44-77). Baltimore, The John Hopkins University Press.

Tassel-Baska, J. V., & MacFarlane, B. (2009). Enhancing creativity in curriculum. In L. V. Shavinina (Ed.), *International handbook on giftedness* (pp. 1062-1083). New York: Springer.

Taylor, C. W., Albo, D., Holland, J. & Brandt, G. (1985). Attributes of excellence in various professions: Their relevance to the selection of gifted/talented persons. *Gifted Child Quarterly, 29*(1), 29-34.

Terman, L. M. (1954). The discovery and encouragement of exceptional talent. *American Psychologist, 9*(6), 221-230.

Thompson, Lee Anne, & Oehlert, Jeremy (2011). The etiology of giftedness. *Learning and individual Differences, 20*, 298–307.

Tillema,H. H. (2004). The dilemma of teacher educators: Building actual teaching on conceptions of learning to teach. *Teaching Education*, *15*, 277-291.

Tomlinson, C. A. (2005). Quality curriculum and instruction for highly able students. *Theory Into Practice, 44*(2): 160-166.

Treffinger, D. J., Young, G. C., Selby, E. C., & Stephardson, C. (2002). *Assessing creativity: A guide for educators*. [Research monograph series]. Washington DC: National Research Center on the Gifted and Talented, Storrs, CT.

Troman, G., Jeffrey, J., & Raggl, A. (2007). Creativity and performativity policies in primary school cultures. *Journal of Education Policy*, *22*(5),549 –572.

Trost, G. (2000). Prediction of excellence in school, university and work. In K.A. Heller, F. J. Mönks, R. J. Sternberg, & R. F. Subotnik (Eds.) *International handbook of giftedness and talent* (2nd edn., pp. 317-327). Oxford: Elsevier.

Tuttle, F. B., & Becker, L. A. (1983). *Characteristics and identification of gifted and talented students* (2nd edn.). Washington, DC: National Education Association.

Vallerand, R. J. (1997). Toward a hierarchical model of intrinsic and extrinsic motivation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (pp. 271-360). New York: Academic Press.

Van Bijsterveldt-Vliegenthart, M., & Zijlstra, H. (2011). *Aanbieding actieplannen primair onderwijs, voortgezet onderwijs en leraren* [Offering action plans primary education, secondary education, and teachers]. Den Haag, Netherlands. Available from www.rijksoverheid.nl.

VandenBos, G. R. (Ed.) (2007). *APA Dictionary of Psychology*. Washington, DC: American Psychological Association.

Van Tassel-Baska, J. (1994). Science curriculum for the gifted. In J. Van Tassel-Baska (Ed.), *Comprehensive curriculum for gifted learners* (2nd edn., pp. 231-261). Needham Heights: MA: Allyn and Bacon.

Van Tassel-Baska, J., and Brown, E. F. (2007). Toward the best practice: An analysis of the efficacy of curriculum models in gifted education. *Gifted Child Quarterly, 51*(4): 342-358.

Vollmeyer, R., & Rheinberg, F. (2006). Motivational Effects on Self-Regulated Learning with Different Tasks. *Educational Psychology Review, 18*, 239–253.

Walberg, H. J. & Stariha, W. E. (1992). Productive human capital: learning. creativity and eminence. *Creativity Research Journal, 5*(4), 323-340.

Walberg, H. J., Shiow-Ling Tsai, S-L., Weinstein, T., Gabriel, C. L., Pinzur Rasher, S., Rosecrans, T., Rovai, E., Ide, J., Trujillo, M., & Vukosavich, P. (1981). Childhood traits and environmental conditions of highly eminent adults. *Gifted Child Quarterly, 25*(2), 103-107.

Walberg, H., Williams, D. B., & Zeiser, S. (2003). Talent, accomplishment, and eminence. In N. E. Colangelo & G. A. Davis (Eds.), *Handbook of gifted education* (3rd edn., pp. 350-357). Boston MA: Allyn & Bacon.

Wallace, B. (2000). *Teaching the very able child*. London, England: David Fulton Publishers.

Wang, D. (2011). The dilemma of time: Student-centered teaching in the rural classroom in China. *Teaching and teacher education*, *27*, 157-164.

Weinstein, R. S. (2002). *Reaching higher: The power of expectations in schooling*. Cambridge, MA: Harvard University Press.

Wentzel, K. R., Battle, A., Russel, S., & Looney, L. (2010). Social supports from teachers and peers as predictors of academic and social motivation. *Contemporary Educational Psychology*, *35*, 193-202.

Wiegant, Fred, Scager, Karin, & Boonstra, Johannes (2011). An Undergraduate Course to Bridge the Gap between Textbooks and Scientific Research. *CBE—Life Sciences Education, 10*, 83–94.

Williams, C., & Wilson, S. (2012): Achieving both intellectual quality and relevance in classroom learning experiences: some conceptual challenges. *Discourse: Studies in the Cultural Politics of Education*, *33,* 4, 471-484.

Windschitl, M. (2002). Framing constructivism in practice as the negotiation of dilemmas: An analysis of the conceptual, pedagogical, cultural, and political challenges facing teachers. *Review of Educational Research*, 72,131-175.

Winheller, S., Hattie, J. A., & Brown, G. T. (2013). Factors influencing early adolescents' mathematics achievement: High-quality teaching rather than relationships. *Learning Environments Research*, *16*, 49–69.

Witcher, A. E., Onwuegbuzie, A. J., & Minor, L. C. (2001). Characteristics of effective teachers: Perceptions of preservice teachers. *Research in the Schools*, *8*, 45–57.

Wolffensberger, M. V. C. (2012). *Teaching for Excellence; honors pedagogies revealed* (doctoral dissertation). Münster: Waxmann.

Wubbels, T. & Brekelmans, M. (2005). Two decades of research on teacher-student relationships in class. *International Journal of Educational Research, 43*, 6-24.

Wubbels, T., Brekelmans, M., den Brok, P., & van Tartwijk, J. (2006). An interpersonal perspective on classroom management in secondary classrooms in the Netherlands. In C. Evertson & C. Weinstein (Eds.), *Handbook of classroom management: Research practice and contemporary issues* (pp. 1161–1191). New York: Lawrence Erlbaum Associates.

Yoon, Hye-Gyoung & Kim, Mijung (2010). Collaborative reflection through dilemma cases of science practical work during practicum. *International Journal of Science Education*, *32*, 283-301.

Yorke, M. (2003). Formative assessment in higher education: Moves towards theory and the enhancement of pedagogic practice. *Higher Education, 45*, 477-501.

Zenasni, F., Besançon, M., & Lubart, T.(2008). Creativity and tolerance of ambiguity: An empirical study. *The Journal of Creative Behavior*, *42*(1), 61-73.

Ziegler, A., & K. A. Heller (2000). Conceptions of giftedness from a meta-theoretical perspective. In Heller, K. A. Mönks, F. J. Sternberg R. J., & Subotnik, R. F. *International handbook of giftedness and talent* (2nd edn., pp. 3-21). Oxford: Elsevier.

Ziegler, A., & Phillipson, S. N. (2012). Towards a systemic theory of gifted education. *High-ability Studies*, *23*(1), 3-30.

Appendix 1

Questionnaire scales and items

Answering scale: Not at all true of me 1 2 3 4 5 6 7 Very true of me

Intelligence

1. I would prefer complex to simple problems.
2. I'm considered exceptionally or unusually intelligent.
3. I use difficult words.
4. I have a rich vocabulary.
5. I am quick to understand things.
6. I have difficulty at understanding abstract ideas.

Creative thinking

7. I am often able to see the "big picture" where others can't.
8. I am flexible in my thinking.
9. I'm able to grasp new ideas with ease.
10. When someone asks me to solve a difficult problem, I can usually find creative solutions.
11. I like studying various subjects from different perspectives.
12. I enjoy putting forward unconventional ideas.

Openness to experience

13. I am imaginative.
14. I am innovative.
15. I show initiative
16. I am full of ideas.

Desire to learn

17. I am thrilled when I learn something new.
18. I look forward to the opportunity to learn and grow.
19. I want to learn as much as possible.
20. I desire to completely master the subject of my study.
21. I am a true life-long learner.
22. I don't like to learn new things.

Drive to excel

23. It's important for me to do better than other students.
24. It is important to me to do well, compared to other students.
25. My goal is to get a better grade than most of the other students.

Persistence

26. I don't quit a task before it is finished.
27. I finish things despite obstacles in the way.
28. I am a hard worker.
29. I am a goal-oriented person.
30. I do not tend to stick with what I decide to do.
31. I don't finish what I start.

Samenvatting

De ontwikkeling van honors programma's in het hoger onderwijs heeft in de afgelopen twee decennia een grote vlucht genomen in en buiten Europa. In Nederland is deze ontwikkeling onder meer gestimuleerd door het 'Sirius programma: Excellentie in het hoger onderwijs'. Dit programma, gefinancierd door het Ministerie van OCW, stelt hogescholen en universiteiten in staat de beste studenten uit te dagen het beste uit zichzelf te halen. Het doel van honors programma's, zo wordt gesteld op de website van het Siriusprogramma, is tweeledig. Enerzijds moet studenten de mogelijkheid worden geboden om onderwijs te krijgen dat past bij hun kwaliteiten en hen uitdaagt op hun niveau, en anderzijds is het stimuleren van talent belangrijk voor de (economische) ontwikkeling van ons land. Ondanks de wijd verbreide groei van honors programma's is er nog niet veel bekend over hoe juist deze studenten kunnen worden uitgedaagd om het beste uit zichzelf te halen. Het meeste onderzoek naar onderwijs voor excellente studenten is gedaan in het basis- en voortgezet onderwijs. Het belangrijkste doel van deze dissertatie was dan ook het onderzoeken welke didactische methoden geschikt zijn voor het uitdagen van honors studenten.

Dit centrale doel was gebaseerd op de veronderstelling dat honors studenten in bepaalde kwaliteiten significant verschillen van reguliere studenten (hier non-honors studenten genoemd), en daarom andere didactische methoden nodig hebben om hun leerproces optimaal te stimuleren. Honors studenten worden vooral geselecteerd op basis van hun motivatie en prestaties (eindexamencijfers). Uit onderzoek blijkt echter dat examencijfers wel een voorspellende waarde hebben voor cijfers die studenten halen in hun studie, maar nauwelijks voor excellente prestaties in het beroepsleven. Het opleiden van mensen die later een significante bijdrage kunnen leveren in hun vakgebied is, zoals hierboven gesteld, wel een van de doelen van honors onderwijs. Omdat de gebruikte selectiemethoden geen garantie geven wilden we eerst onderzoeken in hoeverre honors studenten inderdaad meer potentie hebben voor excellentie op latere leeftijd. De twee volgende studies richtten zich op de wijze waarop studenten kunnen worden uitgedaagd, en als laatste exploreerden we ervaringen van docenten en hun reflecties op het uitdagen van studenten. Drie vragen staan centraal in dit proefschrift:

1. In hoeverre verschillen honors studenten van non-honors student als het gaat om de kwaliteiten die voorspellend zijn voor excellente prestaties in het beroepsleven?
2. Hoe kunnen excellente studenten in het hoger onderwijs het best worden uitgedaagd?, en
3. Welke dilemma's ervaren docenten bij het uitdagen van excellente studenten?

Verschillen tussen honors en– non-honors studenten **(hoofdstuk 2)**

Uit onderzoek naar de kenmerken die leiden tot excellente prestaties op volwassen leeftijd blijkt dat een hoge intelligentie alleen niet voldoende is. Daarnaast is passie, gedrevenheid, werklust en doorzettingsvermogen nodig om talent volledig tot bloei te brengen (Cox 1926; Snow 1986; Trost 1993). In de literatuur wordt talent gezien als een multidimensioneel

concept (Renzulli 1978; Gagné 1995; Gardner 2002; Sternberg 2003; Heller, Perleth et al. 2005). Eén van de bekendste theorieën is het drie-ringen-model van Renzulli (1978). Talent, zo stelt Renzulli, berust op de interactie tussen bovengemiddelde 'intelligentie', 'creativiteit' en 'motivatie'. Alle drie de componenten in dit model zijn noodzakelijk, en geen van de kwaliteiten afzonderlijk zijn voldoende voor creatieve prestaties op volwassen leeftijd. In dit onderzoek gebruikten we Renzulli's model als basis voor een vragenlijst die door 1122 studenten is ingevuld (666 non-honors studenten en 456 honorsstudenten). Uit de resultaten van dit onderzoek blijkt dat honors studenten zichzelf op deze drie dimensies en de 6 subschalen, significant hoger in schatten dan 'non-honors' studenten, behalve op 'doorzettingsvermogen'. Tevens bleken de resultaten te variëren tussen de opleidingen. Bij rechtswetenschappen en geesteswetenschappen was het verschil in prestatiemotivatie tussen honors en non-honors studenten het sterkst, honors studenten bij natuurkunde waren vooral meer leergierig, terwijl de studenten van het University College zich vooral onderscheidden van non-honors studenten in hun creativiteit. Een mogelijke verklaring voor deze variaties kan worden gevonden in de verschillen in selectie criteria- en methoden die opleidingen hanteren voor honorsstudenten. Daarnaast kan een verschil in academische cultuur tussen de disciplines een verklaring vormen voor deze variatie (Becher & Trowler, 2001). De waarden, normen en regels binnen een opleiding kunnen als referentiepunt hebben gediend bij het beoordelen van studenten van hun eigen kwaliteiten.

Het feit dat er significante verschillen tussen honors en non-honors groepen werden gevonden op deze kwaliteiten die als voorspellend worden gezien voor excellentie in het latere beroep, biedt een rechtvaardiging voor de speciale aandacht voor honorsstudenten. De effectiviteit van onderwijs hangt tenslotte in belangrijke mate af van de aansluiting van de leeromgeving bij de vermogens, interesses en motivatie van de studenten (McKeachie 1986; Snow 1986; Pascarella and Terenzini 1991).

Uitdaging voor studenten

Uitdaging is een centraal begrip in de literatuur over passend onderwijs voor excellente studenten. Excellente studenten leren sneller en weten meer, en zouden zich vervelen in onderwijs dat is afgestemd op de gemiddelde student (Shore & Kannevski, 1993). In Csikszentmihalyi's (1975) flow-model wordt uitdaging gekoppeld aan moeilijkheidsgraad. Dit model gaat uit van een evenwicht tussen de mate van uitdaging en de (door de student ervaren) competentie. Als deze twee in balans zijn, is de intrinsieke motivatie optimaal en kan een ervaring van flow optreden: een toestand van optimale concentratie en focus. Empirisch onderzoek laat zien dat studenten in de flow zone meer leren (Volmeijer & Rheinberg, 2006). Een goed evenwicht tussen uitdaging en competentie zou dus leiden tot optimale motivatie en leren. Volgens Csikszentmihalyi (1975) moet de mate van uitdaging echter in balans zijn met de competenties: teveel uitdaging zou leiden tot stress en frustratie, terwijl te weinig uitdaging tot verveling en negatieve emoties zou leiden die de motivatie en het leren van studenten zouden schaden.

In een eerste case studie (**hoofdstuk 3**) onderzochten we op basis van Csikszentmihalyi's (1975) flow-model, 1) hoe studenten de balans tussen ervaren uitdaging en competentie ervoeren, en 2) welke kenmerken van de leeromgeving bepalend waren voor de uitdaging. Deze vraag is onderzocht in de setting van een bestaande cursus die er goed in slaagt studenten uit te dagen het uiterste uit zichzelf te halen: de cursus 'Advanced cell biology' van het Utrecht University College (UCU). Uit evaluaties van voorgaande jaren bleek, dat studenten deze cursus ervoeren als één van de meest leerzame maar ook meest veeleisende die ze op het UCU hadden gevolgd. Het UCU is het honors college van de Universiteit Utrecht, waar de bachelorfase volgens het Liberal Arts & Science filosofie wordt gegeven aan een groep van ongeveer 700 internationale studenten. De dataverzameling bestond uit interviews met de 13 studenten die deze cursus volgden, en uit observaties tijdens een deel van de bijeenkomsten.

Een interessante uitkomst was dat deze excellente studenten veel uitdaging aankonden; studenten leerden juist het meest in de periode waarin de uitdaging hun competentie het meest oversteeg. Volgens Csikszentmihalyi (1975) raken mensen gefrustreerd en bezorgd als de uitdaging de competentie overstijgt, omdat negatieve emoties het gevoel van flow en de intrinsieke motivatie doen dalen. In de hier onderzochte groep liep de spanning tussen uitdaging en het vertrouwen in de vermogens van de groep om het project te volbrengen geregeld hoog op, waarbij de emoties tijdens het proces lang niet altijd positief waren. Dit had zoals verwacht een negatief effect op de motivatie, op de zin om aan het project te werken, maar interessant genoeg niet op het gedrag van de studenten: ze bleven even hard werken of versnelden hun tempo zelfs. Interessant is, dat de scores voor 'leren' het hoogst zijn op het moment dat de uitdaging het hoogst is. Het tijdelijke verlies aan motivatie (plezier in de taak) en de negatieve emoties hadden kennelijk geen negatieve gevolgen voor het leren. De steun en stimulans die studenten van elkaar ondervonden kan wellicht verklaren waarom studenten hard bleven werken ondanks het tijdelijke verlies aan intrinsieke motivatie.

Uit de analyse van de interviews kwamen drie factoren naar voren die de cursus uitdagend maakten: 1) de complexiteit, 2) de autonomie (het gebrek aan docentsturing) en 3) de hoge verwachtingen die de docenten van studenten hadden. De combinatie van deze drie factoren hebben onzekerheid en stress veroorzaakt tijdens de cursus, maar achteraf gezien hebben juist deze factoren ervoor gezorgd dat studenten het maximale uit zichzelf hebben gehaald.

In de daaropvolgende studie (**hoofdstuk 4**) onderzochten we in hoeverre de drie uitdagings-factoren die we vonden in de eerste casestudy generaliseerbaar waren naar andere honorscursussen. We onderzochten in zes verschillende cursussen welke factoren studenten uitdaagden, en hoe zich die factoren manifesteerden in de leeromgeving. De drie factoren uit de vorige casestudy werden ook in deze zes cursussen gevonden. De derde factor, hoge verwachtingen van de docent, bleek cruciaal. Tijdens cursusactiviteiten die hoog scoorden op complexiteit en autonomie waren studenten weliswaar gemotiveerd, maar ervoeren ook een

zekere vrijblijvendheid. De hoge verwachtingen zijn schijnbaar nodig om studenten uit te dagen om het beste uit zichzelf te halen.

De drie factoren manifesteerden zich op diverse wijzen in de leeromgeving. Nieuwe en abstracte leerstof gecombineerd met een hoog niveau van denken (kritische analyse) waren indicatoren van complexiteit. Autonomie manifesteerde zich op diverse manieren. Een duidelijke indicator van autonomie was het aanbieden van keuzevrijheid, bijvoorbeeld bij het schrijven van een paper. Ook belangrijk voor studenten was het gevoel van controle te hebben over hun resultaten. Het schrijven van papers geeft studenten meer controle dan een tentamen. Andere indicatoren waren het geven van verantwoordelijkheid aan studenten, feedback geven op een niet sturende manier, en de ontvankelijkheid van de docent voor inbreng van studenten. Hoge docent-verwachtingen werden op verschillende manieren waargenomen door studenten. Ten eerste schept de reputatie van de docent al verwachtingen bij studenten. Een tweede indicator was de manier waarop de docent zijn/haar cursus aankondigde in de eerste les en/of in de studiehandleiding; docenten met hoge verwachtingen gaven aan dat studenten door een moeilijke periode heen zouden gaan. Een derde indicator werd gevonden in positionering van de docent ten opzichte van studenten. Studenten ervoeren hoge verwachtingen wanneer de docent hen meer als gelijken beschouwde en ze bijvoorbeeld vroeg om ideeën voor zijn eigen onderzoek. Als laatste was het stellen van veel vragen aan studenten en het lang wachten op een antwoord een manifestatie van hoge verwachtingen.

Uitdaging voor docenten **(hoofdstuk 5)**

De zes cursussen die we onderzochten in de vorige studie verschilden in de mate waarin studenten werden uitgedaagd, wat de vraag opriep waarom docenten in een vergelijkbare context verschillende keuzes maakten voor wat betreft het uitdagen van studenten. Uit de interviews met twaalf docenten bleek, dat het uitdagen van studenten gepaard ging met dilemma's. Dilemma's zijn situaties waarin een docenten een keuze moet maken tussen twee conflicterende alternatieven die beide consequenties hebben. Een dilemma is dus per definitie een situatie waarbij bepaalde waarden of normen minder of niet gerespecteerd zullen worden omdat men prioriteit geef aan andere argumenten. Door de complexe aard van onderwijs zijn dilemma's onlosmakelijk verboden aan het lesgeven; docenten vellen voortdurend oordelen over wat in concrete situaties pedagogisch gezien wenselijk is. Zo kunnen er bijvoorbeeld goede redenen zijn in het belang van de student om hem of haar uitstel te verlenen voor het inleveren van een paper, maar hoe fair is dat ten opzichte van de ander studenten?

Het uitdagen van studenten leverde voor de geïnterviewde docenten zeven verschillende dilemma's op:

1 Maximaliseren van de uitdaging versus de zorg voor het welzijn van de student

Dit dilemma behelst de keuze tussen het uitdagen en de zorg voor andere belangen van studenten, zoals hun gezondheid (overvragen kan te hoge stress veroorzaken) of de vrijheid

van studenten om zelf te bepalen in welke cursussen of activiteiten ze hun extra energie steken.

2. Maximaliseren van de uitdaging versus een veilig leerklimaat

In dit dilemma speelt de keuze tussen uitdagen, bijvoorbeeld door moeilijke vragen te stellen, en het behoud van het gevoel van veiligheid van de student. Terwijl docenten het aan de ene kant belangrijk vonden om studenten uit hun comfortzone te trekken en tot nadenken aan te zetten, was er aan de andere kant het gevaar dat studenten uit angst voor gezichtsverlies minder zouden bijdragen aan de discussies.

3. Maximaliseren van de uitdaging versus het behoud van een goede relatie met de studenten

In dilemma voorzagen docenten dat de veeleisendheid die gepaard gaat met uitdagen de relatie met studenten zou kunnen schaden.

4. Maximaliseren van de uitdaging versus alle studenten binnenboord houden

In dit dilemma staan de verschillen tussen studenten binnen een klas centraal. Ook in honorsgroepen zijn er verschillen in capaciteiten en zelfvertrouwen van studenten. Wat voor de ene student uitdagend is, kan voor de andere student beangstigend zijn, met het mogelijke gevolg dat studenten hun motivatie verliezen.

5. Maximaliseren van de uitdaging versus het behouden van het enthousiasme van studenten

Bij dit dilemma speelde de overweging dat docenten zowel affectieve als cognitieve doelen nastreven met hun cursus. Het enthousiasmeren van studenten voor het vakgebied kon op gespannen voet staan met het stellen van hoge eisen.

6. Maximaliseren van de uitdaging versus trouw blijven aan je (professionele) identiteit

Professionele identiteit verwijst naar het beeld dat de docent heeft van zichzelf als docent, zoals de taakopvatting, motivatie, en zelfvertrouwen (Kelchtermans, 2009). Overwegingen van docenten om studenten wel of juist niet uit te dagen omvatten bijvoorbeeld het idee dat hoge eisen stellen aan docenten ook veel vergt van de inzet van de docent, of dat uitdagender onderwijs meer boeiende discussies oplevert, en daarom ook motiverender voor de docent.

7. Maximaliseren van de uitdaging versus voldoen aan de verwachtingen van anderen

In dit dilemma staat de keuze centraal tussen het uitdagen en het voldoen aan bijvoorbeeld regels en afspraken binnen het instituut. Zo zijn cursussen op het eerste niveau (100-level) bedoeld om basiskennis aan te brengen voor de daaropvolgende cursussen, waardoor uitdaging in de vorm van meer autonomie voor studenten in het gedrang kan komen.

De keuzes die gemaakt werden in deze dilemma's waren ongeveer gelijk verdeeld tussen het wel en niet maximaliseren van de uitdaging. Terwijl docenten vaak dezelfde overwegingen hadden, kon de keuze verschillend zijn, afhankelijk van de prioriteit die werd gegeven aan hun doelen en/of verantwoordelijkheden.

Conclusies en reflectie **(hoofdstuk 6)**

In het slot hoofdstuk (hoofdstuk 6) zijn de resultaten samengevat, en wordt teruggeblikt op de leerpunten, de beperkingen en de implicaties van het onderzoek. In relatie met de drie

centrale vragen in dit proefschrift, hebben we de volgende zes leerpunten beschreven. De eerste drie leerpunten zijn gekoppeld aan onderzoeksvraag 2.

Onderzoeksvraag 2: Hoe kunnen honors studenten het best worden uitgedaagd?

1. Honors studenten worden uitgedaagd om het beste van zichzelf te geven wanneer de uitdaging de competentie overstijgt

Studenten leerden het meest als er een behoorlijke frictie was tussen het niveau van uitdaging en de (ervaren) competentie. Studenten hadden in deze perioden van grote frictie minder plezier in het leren, maar hun inzet verminderde niet. Zoals één van hen zei:

> Het plezier was weg, zeker tijdens dit tweede dal dacht ik 'waarom heb ik hier in hemelsnaam voor gekozen?'. Het hoefde helemaal niet, ik had gewoon geschiedenis kunnen doen weet je, over de middeleeuwen, een beetje achterover zitten en het op je af laten komen. [...] Maar dat waren eigenlijk ook de weken dat ik het hardst heb gewerkt en het meest heb geleerd.

Achteraf vonden studenten deze periodes waardevol omdat de voldoening groot was.

2. Hoge verwachtingen van de docent is een cruciale factor

Autonomie en complexiteit zijn belangrijke factoren voor het uitdagen van studenten. Maar, zonder hoge verwachtingen neigen studenten zich te gedragen naar wat Hattie (2012) het minimax principe noemt: een maximaal cijfer voor een minimum aan inzet. De invloed van hoge verwachtingen van de docent werd verondersteld vooral van invloed te zijn bij jongere kinderen en minder bij studenten (Babad, 1998; Brophy, 1983; Weinstein, 2002). Uit ons onderzoek bleek echter dat deze studenten erg gevoelig waren voor de verwachtingen van hun docenten.

3. Uitdaging zit meer in de combinatie van de drie factoren dan in de aard van de taak

Docenten zoeken uitdaging vaker in de aard van de taak zelf, bijvoorbeeld door project- of onderzoeksopdrachten te geven, terwijl studenten de uitdaging meer vinden in het succesvol uitvoeren van de taak (Innoue, 2007). Voor de uitvoering van de taak is de moeilijkheidsgraad belangrijker dan de aard van de taak, zoals onderzoekstaken op projectopdrachten. In ons onderzoek vonden we dat studenten een sterke uitdaging konden vinden in heel verschillende leertaken, en dat ze bij vergelijkbare (onderzoeks-) opdrachten verschillende mate van uitdaging konden ervaren, afhankelijk van de mate van complexiteit, autonomie, en docentverwachtingen. Specifieke taken, zoals onderzoekstaken kunnen potentieel uitdagend zijn, maar de uitdaging is er alleen als studenten de noodzaak voelen om echt te investeren in de taak (Hattie, 2012).

Onderzoeksvraag 3: welke dilemma's ervaren docenten bij het uitdagen van excellente studenten?

4. Uitdagen van studenten is oncomfortabel voor docenten

Voor docenten bleek het uitdagen van studenten soms net zo oncomfortabel te zijn als voor studenten, omdat de negatieve gevoelens of stress van studenten invloed konden hebben op andere belangrijke zaken zoals de atmosfeer tijdens de lessen of de relatie met studenten. Vooral het hebben van hoge verwachtingen veroorzaakte dilemma's voor docenten, wellicht omdat hoge verwachtingen in directe interactie met studenten worden uitgedragen. Het vereist moed van een docent om hoge verwachtingen te hebben, omdat bijvoorbeeld het aanspreken van studenten op hun gedrag, of het stellen van moeilijke vragen en lang wachten op antwoorden ongemakkelijke stiltes en gevoelens van onveiligheid kunnen oproepen.

5. De overwegingen die docenten hebben bij het maken van keuzes geven inzicht in de complexiteit van het lesgeven

De diversiteit aan overwegingen die docenten naar voren brachten naar aanleiding van het uitdagen van studenten benadrukten de rol die de persoonlijke voorkeuren, mogelijkheden, rolopvattingen en gevoelde verantwoordelijkheden van docenten spelen in het onderwijs. In wezen draait het in het onderwijs om te handelen in het belang van studenten. Dit is echter meer complex dan het lijkt, omdat men verschillend kan denken over wat het beste is voor de student, en bovendien kan wat goed is voor de ene student schadelijk zijn voor de ander.

Onderzoeksvraag 1: In hoeverre verschillen honors studenten van non-honors student als het gaat om de kwaliteiten die voorspellend zijn voor excellente prestaties in het beroepsleven?

6. Honors groepen zijn heterogeen

Honors studenten blijken inderdaad significant hoger te scoren op de kwalieiten die voorspellend zijn voor excellentie in het latere beroep (intelligentie, creativiteit, en motivatie). Uit nadere analyse bleek echter dat de honors groepen bijna even heterogeen zijn als de non-honors groepen, ondanks de selectie. Deze heterogeniteit zou kunnen verklaren waarom docenten zo frequent het dilemma ervoeren met het uitdagen van studenten versus het binnenboord houden van alle studenten.

Dit onderzoek kende ook zijn beperkingen en riep nieuwe vragen op. Zo is de context waarin het onderzoek werd verricht beperkt tot de Universiteit Utrecht, en in het bijzonder het University College. Vervolgonderzoek in andere contexten wordt aanbevolen. Implicaties van dit onderzoek richten zich op het de wijze waarop en de mate waarin uitdaging kan worden geboden voor deze groep studenten en op het belang van het bieden van tegenwicht in de vorm van ondersteuning, om te voorkomen dat studenten teveel gefrustreerd raken en hun motivatie verliezen. Verder wordt geadviseerd om docenten te stimuleren om gemeenschappelijk te reflecteren op de dilemma's die ze ervaren, om elkaar uit te dagen hun vooronderstellingen te onderzoeken, en te helpen nieuwe wegen te zoeken om met deze dilemma's om te gaan.

Publications

Peer reviewed publications

Brekelmans, J.M.G., Brok, P.J. den & Scager, K. (2001). Zorg om de kwaliteit van docenten. *Onderwijskundig lexicon, Katern Kwaliteitszorg, III*, 73-91.

Lecq, R. van der, Scager, K., Graaff, R. de, Nab, J., Lie, O.S.H. & Schreel, L. (2006). Interdisciplinair leren denken. *Onderzoek van Onderwijs, 35*(4), 60-66.

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, T. (2013). How to persuade honours students to go the extra mile; Creating a challenging learning environment, *High Ability Studies,* DOI: 10.1080/13598139.2013.841092.

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, T. (2012). Challenging high-ability students. *Studies in Higher Education,* DOI:10.1080/03075079.2012.743117.

Scager, K., Akkerman, S.F., Keesen, F.J.M., Mainhard, M.T., Pilot, A. & Wubbels, T. (2012). Do honors students have more potential for excellence in their professional lives? *Higher Education, 64*(1), 19-39.

Scager, K., Akkerman, S.F., Boonstra, J., Pilot, A., Wiegant, F.A.C. & Wubbels, T. (2012). Uitdagen van excellente studenten. *Onderzoek van Onderwijs, 41*(1), 16-21.

Scager, K. (2008). Vragen talentvolle studenten ander onderwijs? *Onderzoek van Onderwijs, 37*(4), 66-69.

Wiegant, F.A.C., Scager, K. & Boonstra, J. (2011). An undergraduate course to bridge the gap between textbooks and scientific research. *CBE-Life Sciences Education, 10*, 83-94.

Publications in progress

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, T. (*submitted*). Teacher dilemmas in challenging high-ability students in higher education.

Conference papers

Berge, J.H. ten & Scager, K. (2012). Challenging faculty to challenge their students. Paper presented at the Conference Evoking Excellence in Higher Education and Beyond, 4-5 October: Groningen, the Netherlands.

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, Th. (2012). Challenging gifted university students. Paper presented at the annual conference of the American Educational Research Association (AERA), April 13-17: Vancouver, Canada.

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, Th. (2011). Motiveren van excellente studenten: de spanning tussen uitdaging en vertrouwen. Paper gepresenteerd tijdens de Onderwijs Research Dagen (ORD), 8-10 juni: Maastricht, Nederland.

Scager, K., Akkerman, S.F., Pilot, A. & Wubbels, Th. (2011). Motiveren van excellente studenten: de spanning tussen uitdaging en vertrouwen. Paper gepresenteerd tijdens de Onderwijs Research Dagen (ORD), 8-10 juni: Maastricht, Nederland.

Scager, K., Akkerman, S., Keesen, F.J.M., Mainhard, M.T., Pilot, A. & Wubbels, Th. (2010). Verschillen tussen Honours studenten en reguliere studenten. Paper gepresenteerd op de Onderwijs Research Dagen (23 - 25 juni): Enschede, Nederland.

Scager, K., Boonstra, J., Peeters, A.J.M. & Wiegant, F.A.C. (2012). High-ability students and cooperative learning: What conditions make teams flourish? Paper presented at the Conference Evoking Excellence in Higher Education and Beyond, October 4-5: Groningen, the Netherlands.

Books

Scager, K. (1990). *Instruktie bij het ontwikkelen van een module*. Culemborg: Technipress.

Scager, K. & Thoolen, B.J. (2006). *De docent als coach in het hoger onderwijs*. Groningen/Houten: Wolters-Noordhoff.

DOZ-series, as co-author:

Collegiale consultatie en intervisie (1997)
Didactisch vademecum (1998)
Doelen formuleren in het onderwijs (1994)
Gids voor projectonderwijs (1999)
Handleiding zelfevaluatie (1992)
Kwaliteit, de docent een zorg (1995)
Leerplan en leerplanontwikkeling (1996)
Toetsen en beoordelen (1997)
Studiebegeleiding (1994)
Verkorte leerwegen in het HBO (1999)

Other publications

Bok, Mirjam, Kool, Ada, & Scager, Karin (*submitted*). Excellentie nader onderzocht: studentprofiel, selectiecriteria, didactiek & leeromgeving, opbrengsten en training van docenten.

Gestel, M. van, Rijpkema, P., Berge, H. ten, Eijl, P. van & Scager, K. (2008). Het Utrecht Law College na drie ontwikkeljaren. Onbekend: UU LEG LAW, Economics and Governance.

Keulen, H. van & Scager, K. (2009). *Leren houden van het vraagteken. Onderzoeksgericht onderwijs in bachelor en master*. Utrecht: IVLOS, Universiteit Utrecht.

Lecq, R. van der, Scager, K., Graaff, R. de, Nab, J., Lie, O., en Schreel, L. (2006). Interdisciplinair leren denken. *Onderzoek van Onderwijs*, 35(4), 60-66.

Scager, K. (2010). Wat is talent? In P.J. van Eijl, A. Pilot & M.V.C. Wolfensberger (Eds.), *Talent van morgen. Ontwikkeling van talent in het Hoger Onderwijs* (pp. 17-24). Groningen: Noordhoff.

Scager, K. (2008). *Your guide to successful teaching at University College Utrecht*. Utrecht: Universiteit Utrecht.

Scager, K. (2008). *Casusbeschrijving Honoursprogramma University college Utrecht*. (Mededelingen 85). Onbekend: UU IVLOS.

Scager, K. & Sonneveld, H. (2008). *De kwaliteit van de promotiebegeleiding bij het Onderzoeksinstituut Geschiedenis en Cultuur*. Utrecht: Universiteit Utrecht.

Wiegant, F.A.C., Boonstra, J., Peeters, A.J.M. & Scager, K. (2012). Team-based learning in honors science education: The benefit of complex writing assignments. *Journal of the National Collegiate Honors Council, 13*(2), 197-219.

Curriculum Vitae

Since 1991, Karin Scager is a senior educational consultant and a teacher trainer at Utrecht University in the Netherlands, at the Centre for Teaching and Learning, Faculty of Social and behavioural Sciences. Her particular activities are in the area of training of teachers in didactical and skills, in curriculum design and –development projects, and in quality care strategies and research. Furthermore, she provides workshops, training sessions and lectures at various Universities and Polytechnics in the Netherlands and other countries, on the subject of teaching methods and quality care. She regularly acts as a member of accreditation committees as educational specialist.

Formerly (1985-1991), she was an educational consultant at the Utrecht Polytechnics (Hogeschool van Utrecht). In the department of Social- and Health studies, she worked as project manager of different projects and as interim manager in the departments of Speech Therapy Education and Professional Nursing Education.

Dankwoord

Het zit erop. Het was een leerzame tocht, met mooie vergezichten, mooie ontmoetingen, stijle beklimmingen, en diepe dalen. Soms eenzaam, maar ik was gelukkig niet alleen. Als eerste mee op reis waren mijn trouwe begeleiders, Sanne, Albert en Theo, wat een topteam! Jullie geduld, rust en vertrouwen waren onbegrensd. Sanne, jouw inspiratie en je niet aflatende enthousiasmewaren enorm stimulerend. Je had altijd een luisterend oor voor twijfels en aarzelingen, en met je heldere inzichten ruimde je de problemen die we onderweg tegen kwamen moeiteloos uit de weg. Albert, jouw onuitputtelijke bron van ideeen en connecties brachten steeds weer nieuwe contacten en inspiratie. Jij hebt me gestimuleerd om actief deel te nemen aan de onderzoeksgemeenschap, om samen te werken, en inzichten te delen. En Theo, jij zorgde dat ik koers hield en niet verdwaalde in mijn data. Het was vast niet eenvoudig om van een adviseur een onderzoeker te maken, maar jij hebt daar in elk geval alles aan gedaan. Alle drie waren jullie onvermoeibaar in het lezen en becommentarieren van teksten, met veel aandacht voor de grote lijnen zowel als voor de details. De ruimte die jullie me gaven om mijn eigen weg te volgen, in combinatie met jullie sterke betrokkenheid, was hartverwarmend.

Het gezin was ook mee op stap. Jaap, Anna, Gijs en Niels, jullie grote betrokkenheid bij het proces én het product was fantastisch. Jullie zorgden voor de nodige afleiding, relativering, support, en droegen ook bij aan het proefschrift zelf. Niels, dank voor het bedenken van de titel en je peptalk. Gijs, dank voor het maken van de figuren en de omslag. Anna, dank voor het meedenken over de onderzoeksvragen, je suggesties voor onderzoeksmethoden, en voor het optreden als tweede beoordelaar. En Jaap, hoewel je je twijfels over mijn besluit om deze treis te ondernemen heb je me van begin tot eind gesteund, zowel inhoudelijk als mentaal. Onvermoeibaar dacht je mee, las je mee, gaf je commentaar, en trad je op als tweede beoordelaar. Je bood steun, een luisterend oor, troost, of juist afleiding, alnaargelang mijn gemoedstoestand. En elke dag, 4 jaar lang, kookte je heerlijke maaltijden.

Bijzondere dank ben ik verschuldigd aan de collega's van het University College, voor hun openheid en samenwerking. Rob en Fried, dank voor jullie inspiratie en voor het openzetten van alle deuren. Fred en Johannes, wat heb ik veel van jullie geleerd! De andere twaalf docenten die zo hartelijk bereid waren om zich te laten observeren en interviewen ben ik zeer dankbaar, maar zal ik hier omwille van anonimiteit niet met name noemen. Wat een fantastische docenten zijn jullie en wat een mooie ervaring was het om een kijkje te mogen nemen in jullie topkeukens.

Ook mijn (IVLOS) collega's en leidinggevenden waren onmisbaar op deze reis. De belangstelling van teamgenoten en van niet-teamgenoten, de praatjes bij het kopieerapparaat, en de bereidwilligheid van de onderzoeksgroep om mee te denken en feedback te gevenvoorkwamen isolement. Tijdens de laatste etappe kreeg ik een fijne kamergenoot, die

altijd open stond voor een praatje, bereid was inhoudelijke adviezen te gevenvoor een advies, en bovendien zorgde voor de dagelijkse portie wortejtjes. Dank, Dubravka. Ook de facilitering van het project is belangrijk, en daarin speelde Ineke achter de schermen een cruciale rol. Jij regelde tijd, begeleiding, coordinatie en stroomlijning, en Desirée en Jan hebben de financiele middelen beschikbaar voor het laatste deel van de reis.

En dan natuurlijk mijn lieve familie, vrienden, en Piep en Auke. Zonder jullie was dit lang niet zo'n mooie reis geworden!

Bilthoven,
23 oktober, 2013