Mathematics bridging education using an online, adaptive e-tutorial: preparing international students for higher education

To be published in:
Teaching Mathematics Online: Emergent Technologies and Methodologies, Chapter 30
Editors: Angel A. Juan; Maria A.; Sven Trenholm; Cristina Steegmann
IGI Global

Dirk T. Tempelaar
Maastricht University School of Business & Economics, the Netherlands

Bart Rienties
Centre for Educational and Academic Development, University of Surrey, UK

Wolter Kaper
Universiteit van Amsterdam, the Netherlands

Bas Giesbers
Maastricht University School of Business & Economics, the Netherlands

Sybrand Schim van der Loeff
Maastricht University School of Business & Economics, the Netherlands

Leendert van Gastel
Universiteit van Amsterdam, the Netherlands

Evert van de Vrie
Open Universiteit Nederland, the Netherlands

Henk van der Kooij
Freudenthal Institute, Universiteit Utrecht, the Netherlands

Hans Cuypers
Technische Universiteit Eindhoven, the Netherlands

ABSTRACT

This contribution describes and evaluates a postsecondary remediation program in mathematics, aiming to ease the transition from high school to university and to improve the success rates in the first year of bachelor studies. The remediation program consists of the administration of an entry test and the organisation of voluntary bridging education in the format of an online summer course, using the adaptive e-tutorial ALEKS. Participants are prospective students of the university programs business and economics of Maastricht University, and are mostly students with an international background. Effect analysis suggests a strong treatment effect of successful participation in the summer course. However, given the quasi-experimental setup of this study, with non-equivalent groups, selection effects may be responsible for part of that effect. Correction of the treatment effect by applying the propensity score method indicates that indeed a selection effect is present, but that a substantial treatment effect remains, of about 50% the size of the effect of being educated in advanced math versus basic math, in high school.

INTRODUCTION

This contribution focuses on a type of education that is referred to in different ways: bridging education, developmental education, or remedial education. Whatever the label, the type of education subject of this contribution is education directed to ease the transition from high school to college and to improve the success rates in the first year of bachelor studies. In the Netherlands, the main advising council for educational affairs, the Educational Council of the Netherlands, has stressed the importance of bridging education in a range of studies and recommendations (Onderwijsraad 2006, 2007, 2008). The dating of these advices makes evident that Dutch interest in bridging education beyond the institute of open education is recent. Nation wide projects, supported by SURF, the Dutch collaborative organisation for higher education institutions and research institutes aimed at innovations in ICT, run from 2004 onwards. Some of these Dutch initiatives have acted as pioneer for European projects, indicating that interest in (continental) Europe is also of very recent date. EU projects M.A.S.T.E.R., S.T.E.P. and MathBridge collect experiences with bridging education with a specific European focus: that of internationalisation of European higher education. This internationalization development is going very fast; for example, some Dutch universities located at short distance of country borders, like the case elaborated in this article, the share of international students in the inflow of new bachelor students has risen to 75% (mostly from continental Europe). Although most of these students are not very international in terms of the geographical distance they have to bridge, there is certainly a huge diversity with respect to high school education they have received. Secondary school systems, even in neighbouring countries as Netherlands, Germany and Belgium, are very different, producing strong heterogeneity in knowledge and skills of prospective students. That heterogeneity brings about a strong need for bridging education in the transfer from secondary to university education, which adds to the more national focused needs for bridging education that have existed for some time: to bridge knowledge and skills deficiencies in areas that are part of the national secondary school program, but are not sufficiently mastered by students transferring to university.

US context

The longest tradition of bridging education is without doubt to be found in the Anglo-Saxon education system, and specifically, in the US. Developmental education for underprepared students, as it is generally labelled, is in the US quite often organized state-wise, and has achieved an enormous reach: estimates of participation of undergraduate students in developmental education in any format offered by community colleges and universities ranges between 40% and 58% of first year students (Attewell et al., 2006; Bailey, 2009; Kozeracki, 2005). Most recent discussions in the US on the topic of developmental education is focusing on the question if there is any way back: the opinion that too large a share of public funding of education is finding its way into developmental education is shared by many, opening the debate how to improve regular education to diminish the need of developmental education (see e.g. the special edition of New Directions for Community Colleges, 2008). No surprise therefore that by far the most empirical studies into the effect of bridging education refer to the US context: Bahr (2008), Bettinger and Long (2008), Calcagno and Long (2008), Jamelske (2009). The specific US context of these studies steers to a large extent the way the research question of the impact of bridging education is approached: the US higher education system is strongly based on selection, and part of most selection procedures is that prospective students participate in a placement or entry test and, in the case that they score less than a certain cut-off point, are required to take developmental education. In such a typical US context, impact studies compare the academic success of students scoring just below the cut-off score (who are obliged to participate in the bridging education) with that of students who score just above the cut-off score (and who are excluded from bridging education), using so-called regression-discontinuity models.

European context

In the (continental) European context, such an approach cannot be followed: in most cases, no selection takes place upon entering university education, so the option is missing to obligate some, and to exclude other, students from bridging education (Brants & Struyven, 2009; Rienties et al., in press; Tempelaar & Rienties, 2008). In countries where no selection takes place, as e.g. the Netherlands, the legal basis is lacking to require prospective students to participate in an entry exam, and/or remedial program. Another difference between the US and European case is that where in the US placement tests are used to place students in different levels of their first math course, with the lowest level being a remedial course that usually does not earn credits toward a degree, in the typical European case all students are placed in one math course, with remedial educational playing the role of bridging toward that single course, again without earning credits. But although contexts are rather different, from a methodological point of view both European and US contexts share important characteristics: in the investigation of the impact of bridging education, one cannot use the experimental design, since participation in bridging education does not take place on the basis of randomized assignment, but on the basis of the outcome of a placement test (US), or self-selection (Europe). A direct comparison of academic success of participants and non-participants of bridging education is therefore not a proper way to find a treatment effect, since the composition of the two groups of prospective students will, in general, be different. The relevant research design is that of the quasi-experiment with non-equivalent groups, that requires a correction of the differences observed between experimental and control group on the basis of differences in background statistics of students in both groups (the covariates). In the US-based empirical studies, it is one single background factor, the score on the placement test that distinguishes the students in the treatment group from students in the control group, and so allows the use of regression discontinuity methods. The typical European case lacks such a discontinuity, and directs the investigator to methods recently developed for

the quasi-experimental setup without pre-test and with non-equivalent groups: propensity score-based methods (Fraas, 2007; Shadish, Cook, & Campbell, 2002; Yanovitzky, Zanutto, Hornik, 2005).
The effect analysis presented in this contribution makes use of experiences achieved in the bridging courses mathematics for prospective bachelor students of the Maastricht University School of Business and Economics. Those courses are designed as optional summer courses that take place in the summer before the start of the regular bachelor program. In the European context, it is one of the longest lasting cycles of bridging education: from summer 2003 on, these summer courses have been offered without major changes, and in seven consecutive runs, 750 prospective students have participated. The bridging courses focus on the international students, entering the bachelor study with a non-Dutch prior education, and indeed 90% of the participants are of international background. Is an optional summer course an effective instrument to help international students bridge math deficiencies caused by differences in national secondary school systems? This is the central question of this contribution, against the background of cumulating evidence in the US context that developmental education certainly is expensive, but doubtful in its effects.

THE UM SUMMER COURSE MATHEMATICS

The characteristics of students flowing in into the programs in Business and Economics, combined with the outcomes of the entry assessments to be discussed in more detail in the next section, have been conclusive with regard to major design choices of the bridging education, including the preference for a summer school format. Some of the major considerations at play were:

- The large differences in prior math mastery require a bridging course of considerable size: up to a workload of approximately 100 hours for students with the most basic forms of prior math schooling. This size is incomparable with that of most of the existing national bridging courses, which are quite often scheduled in a couple of days of intensive teaching.
- For a bridging course of this size and the strong heterogeneity of students, it is crucial that the course is tailor-made: adapting to the students' mastery. Each student should be able to enter the course at the appropriate level.
- To achieve this adaptive feature, (repeated) diagnostic testing is crucial, as well as the ability to adapt learning materials to the outcomes of individual, diagnostic tests.
- The size of the bridging course, and the large variation in work load for students depending on their prior mastery, prevents offering such a bridging course 'in the gate' (that is: intra-curricular, during the first few weeks of the regular program), but forces it to be offered 'before the gate' (that is: extra-curricular, during the summer that precedes the start of the regular program).
- Since participants of the bridging courses are -predominantly- international students, the bridging course cannot be offered on site, but should be offered according the model of distance e-learning.
- Since the period in which the summer course is offered is also occupied by holidays, jobs, and practical work, the format of the summer course should be very flexible: the summer course should be available over a relative long period (June, July, August), with a maximum of freedom for students to schedule their individual learning around other activities in that summer.

Based on all these grounds, it was concluded that face-to face education could not meet several of the above requirements, making the decision to organize the bridging course around an existing adaptive, electronic tutorial inevitable: ALEKS (Assessment and LEarning in Knowledge Spaces) College Algebra module. The tool makes use of server based computing, and can be characterised as supporting individual, distance learning. The ALEKS system (see also Doignon & Falmagne, 1999; Falmange et al., 2004; Tempelaar et al., 2006) combines adaptive, diagnostic testing with an electronic learning and practice tutorial in several domains relevant for higher education. In addition, it provides lecturers an instructor module where students' progress can be monitored, both in the learning and assessment modes.

The ALEKS assessment module starts with an entry assessment in order to evaluate a student's knowledge state for the domain. Following this assessment, ALEKS delivers a graphic report analyzing the student's knowledge within all curricular areas for the course. The report also recommends concepts on which the student can begin working; by clicking on any of these concepts or items the student gains immediate access to the learning module.

Some key features of the assessment module are (see Figure 1 for a sample):

- All problems require that the student produce authentic input.
- All problems are algorithmically generated.
- Assessment questions are generated from a carefully-designed repertoire of items ensuring comprehensive coverage of the domain.

The assessment is adaptive: the choice of each new question is based on the aggregate of responses to all previous questions. As a result, the student's knowledge state can be found by asking only a small subset of the possible questions (typically 15-25).

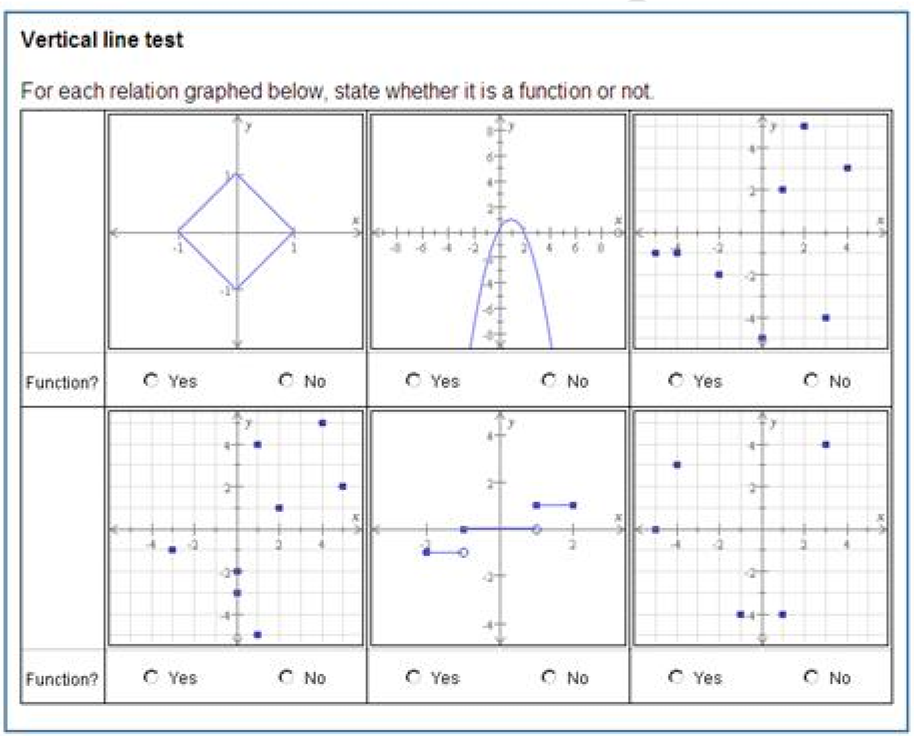

Figure 1: Sample of an ALEKS assessment item.

The learning report, of which Figure 2 shows part of, provides a detailed, graphic representation of the student's knowledge state by means of a pie-chart divided into slices, each of which corresponds to an

area of the syllabus. In the ALEKS system, the student's progress is shown by the proportion of the slice that is filled in by solid colour: if the slice is entirely filled in, the student has mastered that area. Also, as the mouse is held over a given slice, a list is displayed of items within that area that the student is currently "ready to learn," as determined by the assessment. Clicking on any of these items gives access to the learning mode (beginning with the item chosen).
At the conclusion of the assessment ALEKS determines the concepts that the student is currently ready to learn, based on that student's current knowledge state. These new concepts are listed in the report, and the learning mode is initiated by clicking on any highlighted phrase representing a concept in the list. The focus of the learning mode is a sequence of problems to be solved by the student, representing a series of concepts to be mastered. The facilities offered by the learning mode are as follows:

- Practice (that is, the problems themselves);
- Explanations of concepts and procedures;
- Dictionary of technical terms;
- Calculator (adapted to the topic studied, e.g. in statistical items, a special "statistics calculator" is provided).

For example, a student working on a particular problem may "ask for" an explanation of that problem (by clicking on the button marked "Explain"). The explanation typically provides a short solution of the problem, with commentary. After reading the explanation(s), the student may return to "Practice", where she or he will be presented with another problem exemplifying the item or concept just illustrated. If the student is successful in solving the problem, the system will offer (usually) two or three more instances of the same item to make sure the student has mastered it. In the text of problems and explanations, certain technical terms such as "addition", "factor" and "square root" are highlighted. Clicking on any highlighted word or phrase will open the dictionary to a definition of the corresponding concept. The dictionary can also be used independently of the current problem to look up any term the student may be curious about. A graphing calculator is available for computing and displaying geometrical figures in analytical geometry and calculus. Other, related features of the learning mode are Feedback, Progress Monitoring, and Practice. Whenever the student attempts to solve a problem in the learning mode, the system responds to the input by saying whether or not the answer is correct and, if it is incorrect, what the student's error might have been. More generally, ALEKS follows the student's progress during each learning sequence, and will at times offer advice. For example, if a student has read the explanation of a problem a couple of times and yet continues to provide incorrect responses, ALEKS may suggest -- depending on the circumstances -- that the student looks up the definition of a certain word in the dictionary. ALEKS may also propose that the student temporarily abandon the problem at hand and work instead on a related, but easier, problem. When a student has demonstrated mastery of a particular item by repeatedly solving problems, ALEKS will encourage the student to proceed to a new item.

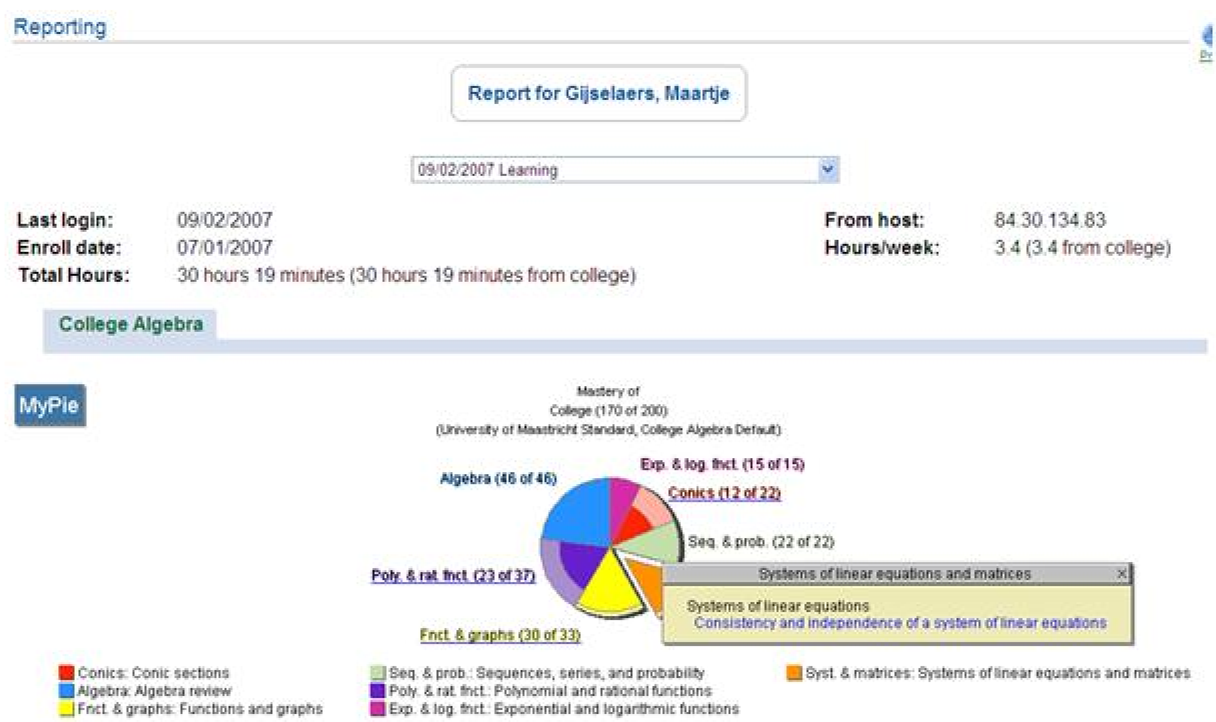

Figure 2: Partial sample of an ALEKS learning report.

The instructor module enables lecturers to monitor individual student progress and achievement, both in learning and assessment mode, and to monitor class progress, again in both modes.

PARTICIPANTS AND NON-PARTICIPANTS

This study is based on the investigation of five cohorts of first year students in the programs business and economics; these are all cohorts for whom full data is available on both the summer course, and relevant students' background characteristics needed for the statistical investigation. In total, these five cohorts contain about 4500 first year students, amongst them 68% are international students. Of these students, 578, or 13%, decide to participate in the voluntary math summer course. That decision to participate follows a chain of information and recruitment activities:

- In the period March to May, prospective students are informed on the option of participating in a free math summer course. Part of the information is a short, digital introductory test that provides the students a global picture of expected math mastery, and their position in this.
- Yearly, about 300-500 people take part in these introductory tests. Serious test takes are selected, receive feedback on their test attempt, and are asked for their interest in participation in the math summer course. Students reacting positively are required to express their willingness to invest at least 80 hours of study efforts in the summer course offered by the university. Yearly, between 150 and 250 prospective students qualify and receive an invitation.
- About half of these invitations, so yearly somewhat more than 100 students, are accepted; these students receive a certificate for the use of ALEKS College Algebra in the summer period.

However, expressing willingness to invest an appropriate amount of time is by far identical to really spending sufficient time: only 52% of all participants manage to achieve a pass for the summer course. To achieve such a pass, students were required to study and master at least 55% of the topics covered in the electronic learning tool. Since this required coverage also includes topics already mastered at the start of the summer course, these passing requirements are rather mild. That is also clear from a comparison of real time investment: total connect time in the e-tutorial of students passing the summer course is on average 52.1 hours, whereas average connect time of students failing the summer course, that is, not achieving the 55% coverage requirement for passing, is only 15.1 hours. Tool connect time is a conservative estimate of total study efforts: it measures how much students study within the tool, but misses study time outside the tool.

After finishing the summer course, end of August, the regular program of the bachelor studies International Business and International Economics starts early September. Both programs begin with two eight-week (half semester) integrated, problem-based learning designed courses, each having a 50% study load. The first course is an introduction into organizational theory and marketing, the other course, called Quantitative Methods I or QM1, an introduction into mathematics and statistics. That second course is of special interest for this study, since the ultimate aim of the summer course is to optimally prepare students for this QM1 course. The very first activity in the QM1 course is to administer an entry test, for several reasons: for longitudinally monitoring the math mastery of prospective students, to provide individual students with diagnostic feedback, and to collect data relevant for the design of both the summer course, and the QM1 course. The coverage of the QM1 course mirrors the circumstance that strong heterogeneity in math mastery, due to students educated in different national systems and at different math levels, necessitates a fair amount of repetition. Most topics covered repeat topics educated in the grades 11 and 12 of Dutch secondary schooling, basic math level (the last two years of high school), with some time devoted to new topics. There is no overlap between QM1 and the content of the summer course, since that content is covering topics of grades 7-10 of secondary schooling (middle school and first year of high school). Effect analysis in this study will focus on student achievements of both participants, and non-participants in the summer course, in this QM1 course. However, outcomes of our study are rather robust with regard to the specific choice of effect variable, due to institutional regulations in Dutch higher education. E.g., both programs are characterized by the presence of a so-called system of binding study advice: students with insufficient academic achievements cannot enrol the program for a second year. Achieving a pass for QM1 is practically a requirement, and in fact the most binding requirement, for achieving a positive binding study advice, implying that academic success in the first year, and that in the QM1 course, do not deviate very much.

The most powerful predictor of academic achievements in QM education is the level of math schooling in high school. In this study, we will distinguish two different levels: basic and advanced. Students who take high school according to the Dutch national system, called VWO (pre-university education), are either taught math at one of two different basic levels (A1 or A1,2), or one of two advanced levels (B1 or B1,2). The lowest level, A1, does not qualify for studies in business or economics, so what remains is one basic, and two different advanced levels. Only a minority of prospective students (32%) is educated within the Dutch national system. Many more, somewhat more than 50%, of the prospective students are educated in a German speaking high school system. That system has again two different levels of math prior education, the advanced level or ‘Leistungskurs’, and the basic level or ‘Grundkurs’. The remaining students either have an International Baccalaureate (IB) diploma, or are educated within a national system outside the Dutch or German speaking part of Europe. IB again allows distinguishing advanced level

(HL) from basic level (SL), whereas for the last category, students were asked to classify their own math prior education either as math major, or as math minor. The binary variable achieved this way is an important predictor of academic achievement. However, it should be realized that it is no more than a very crude classification, given the strong differences between national educational systems. Table 1 contains the decomposition of both participants and non-participants in the math summer course with regard to different types of prior education, and the level of math prior education, of students of which data on prior education are available. With regard to nationality, two different groups are distinguished: Dutch versus International. Students with an IB diploma are regarded as being part of the last group, but can be of any nationality; this implies that International refers to the type of prior education, rather than nationality.

Table 1: Composition of five cohorts of first year students with regard to prior education

Summer course participation	Dutch-prior education		International prior education		Total
	math basic	math advanced	math basic	math advanced	
Participant	44 (4.3%)	10 (2.7%)	394 (18.2%)	104 (12.8%)	552
Non-participant	971 (95.7%)	403 (97.3%)	1769 (81.8%)	708 (87.2%)	3851
Total	1015	413	2163	812	4403

In agreement with the main goal of the bridging course, participation is much stronger amongst international students, than amongst Dutch students, and much stronger amongst students educated at basic level, than amongst students educated at the advanced level. Still, there are relatively many summer course participants amongst the international students with advanced math prior education. Main explanation is the tradition in German speaking countries to halt study in between high school and university for one or several years, either forced by military service, or voluntary, a tradition that is not present in other continental European countries. Most international participants with advanced math prior education have interrupted their study, and regard the summer course as a crucial opportunity to refresh.

STATISTICAL ANALYSES

An important focus of this contribution is the methodology of the effect analysis. Since participation in the summer course is on a voluntary basis, a quasi-experimental setup for the effect analysis is required. Besides, the design contains a post-test, but no pre-test, so that it is best characterized as a quasi-experimental design with non-equivalent groups and post-test only (Shadish e.a., 2002). Such a design embodies the risk of self-selection. In line with recent advices with regard to finding causal effects in observational studies[1], quasi-experimental elements are added to the research design, the most important one being the inclusion of a broad range of students' background factors that may be related to potential self-selection effects. These students' background characteristics are measured both for students participating the voluntary summer course, and for students who have opted not to participate the summer course, and originate from long-term longitudinal research into student related factors explaining

[1] See the AERA 'think tank white paper': Schneider, Carnoy, Kilpatrick, Schmidt & Shavelson (2007).

academic success. These background characteristics refer to: type of secondary education (Dutch or international), level of math prior education (basic or advanced), learning approaches, goal orientations, metacognition, academic motivations, and subject specific achievement motivations.

Traditional approaches for effect analysis in observational studies determine the treatment effect with a multiple (logistic) regression model or ANCOVA containing as predictor variables, beyond the treatment, also covariates that correct the effect for variation in the effect variable that is not caused by the treatment variable (but is e.g. the outcome of a selection effect). In specific applications, especially when experimental and control group strongly deviate with regard to these background characteristics, this approach has its limitations (Fraas, 2007; Yanovitzky et al., 2005). Therefore, the preferred methodological approach is based on the method of propensity scores, where the treatment effect is corrected (Fraas, 2007; Shadish et al., 2002; Yanovitzky et al., 2005). Basis of that correction are the propensity scores: the conditional probabilities that an individual belongs to the experimental group, or to the control group, given the set of covariates. Propensity scores are generally estimated with logistic regression analysis. The correction of the treatment effect can take place in different ways: using the propensity scores as matching variables, as stratification variables, or as covariates. In this study both of these last approaches will be used.

One background characteristic will not be used in determining the propensity scores, but will be included into the model as a separate factor, together with the propensity score: the level of prior math education. This will allow us to make an explicit comparison of the treatment effect of successfully participating in the summer course, with the effect of being educated at advanced math level in high school.

The covariates: students' background characteristics

In finding relevant covariates, we profited from long term research into study achievements in the first year of study undertaken in our school. The first set of background factors refer to students' approaches to learning, and are investigated in the context of the learning patterns model of Vermunt (Entwistle & Peterson, 2004; Vermunt, 1996) and the instrument based on that model: ILS or Inventory of Learning Styles. Vermunt distinguishes in his model four domains or components of learning: cognitive processing strategies, metacognitive regulation strategies, learning conceptions or mental models of learning, and learning orientations. Next, students' goal orientations are measured with an instrument designed by Grant and Dweck (2003), that classifies goal orientations into six types: intrapersonal outcome goals, intrapersonal ability goals, normative outcome goals, normative ability goals, and two different types of learning goals, that differ in the extent the student is longing for challenge: the learning goal (in the strict sense) and the challenge-mastery focused goal orientation. Metacognitive abilities are measured by the AILI instrument (Elshout-Mohr, Daalen-Kapteijns, & Meijer, 2001; Tempelaar, 2006), that is based on Flavells' three component model of metacognition, which decomposes metacognition into the components knowledge, skills, and attitudes. The Academic Motivation Scale (AMS; Guya, Mageau, & Vallerand, 2003); Ratelle et al., 2007; Vallerand et al., 1992), based upon Ryan and Deci's (2000) model of intrinsic and extrinsic motivation, is applied to achieve motivational profiles of students containing different types of intrinsic, extrinsic, and a-motivation. Lastly, subject achievement motivations based on Eccles' expectancy-value theory (Eccles & Wigfield, 2002) are measured with an instrument derived from the Survey of Attitudes Toward Statistics (SATS) developed by Schau and co-authors (Tempelaar et al., 2007). The SATS instrument measures six aspects of post-secondary students' subject attitudes, amongst which two expectancy factors that deal with students' beliefs about their own ability and perceived task difficulty: Cognitive Competence and Difficulty, and three subjective task-value constructs that encompass students' feelings toward and attitudes about the value of the subject: Affect, Interest and Value. The sixth aspect, Effort, is assumed to be the outcome of the process of weighting expectancy against value.

RESULTS

The national entry test

In the framework of several, consecutive national projects in the Netherlands to improve transition from high school to college, and study success in college, national entry tests for mathematics have been constructed. In three out of the five cohorts in this study, the same version of the national entry test is administered, implying more than 2600 test takes of that test. The short test consists of four categories of math competency, all being part of middle school coverage, or early high school math programs at basic level. Two topics, algebraic skills and e-powers & logarithms, are the most elementary of the topics, the other two, equations and differentiation, slightly more advanced. Figure 3 contains the scores for each of those four topics, and the total score, for four different categories of students: students educated in the Dutch secondary school system or not, crossed with students educated at advanced level, versus basic level, with regard to their math prior education. Scores in the figure are p-values or proportion of correct answers, after correction for guessing (since the entry test contains multiple choice items).

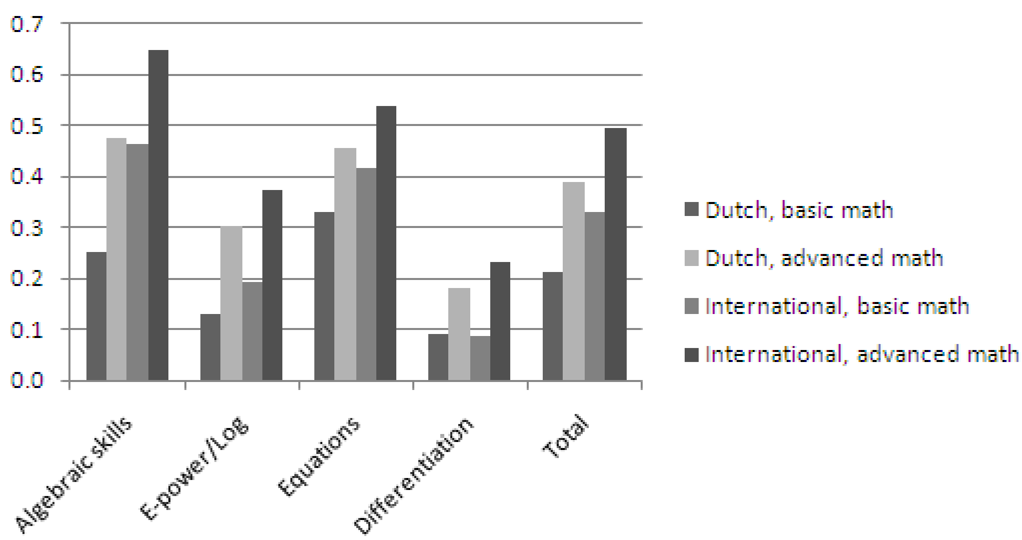

Figure 3: Proportion correct scores entry test and its topics, broken down into prior education.

Scores in the entry test are (disappointing) low. Some low scores have been expected: it was known that more advanced topics as solving equations and differentiation, although being taught in programs for basic math all over Europe, are not fully mastered by prospective students, and are in need of a repetition within the regular program of first year university education. Important outcome of the entry test is however that crucial math deficiencies exist beyond these advanced topics: also topics like algebraic skills and e-powers and logarithms, although firmly rooted in any middle school program, produced no better

than very incomplete mastery. Traditionally, these basic topics are no part of regular university teaching, implying that any deficiency would stay if not addressed in a bridging course.
Another crucial observation from the entry test scores is that whereas the bridging courses have been primarily designed for international students being educated in programs deviating from the typical Dutch program, deficiencies of Dutch prospective students seem to be larger than those of the international students. One should be cautious in generalising the interpretation of such differences in group means, since selection effects are highly probable, but relative to the total score, it is surprising to see how meagre students with a Dutch prior education score in especially the algebraic skills topic. In fact, with regard to this topic students educated at the advanced level do not perform better than international students educated at the basic level. A finding that justifies the large scale national projects in math bridging education within the Netherlands: beyond reasons of internationalization, there are urgent reasons related to the national state of affairs of math education for providing remedial education.

Descriptive analyses

Figures 4 and 5 exhibit the non-corrected treatment effects of successful participation in the summer course of students being educated at basic math level versus advanced math level in high school, respectively for the total score in the course, QM1 total score in Figure 4, and for the QM1 passing rate, in Figure 5. The effect of prior education at advanced level, compared to basic level, is 4.6 points in the QM1 total score (or expressed as effect size, 0.64 standard deviations), against 23% in the passing rate (0.53 standard deviations). The effect of successful participation in the summer course, with no participation as reference, equals 5.3 points in total score (0.76 standard deviations), respectively 28% in passing rate (0.56 standard deviations) for students educated at basic level, and 3.8 points (0.54 standard deviations), respectively 13% (0.31 standard deviations) for students educated at advanced level. As to be expected, the treatment effect is much larger for students educated at basic math level, than for students educated at advanced math level. For the principal target group of students in the summer course, those with a basic math prior education, the non-corrected effect of successful participation in the summer course is even that large, that they outperform students with an advanced prior education who do not participate the summer course, both with regard to QM1 total score, and with regard to QM1 passing rate.

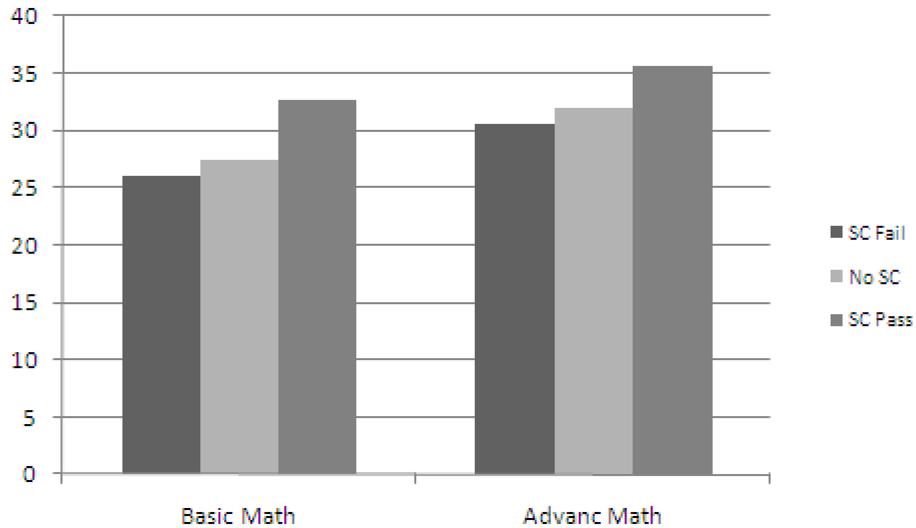

Figure 4: Non-corrected treatment effect of successful participation in summer course on QM1 total score.

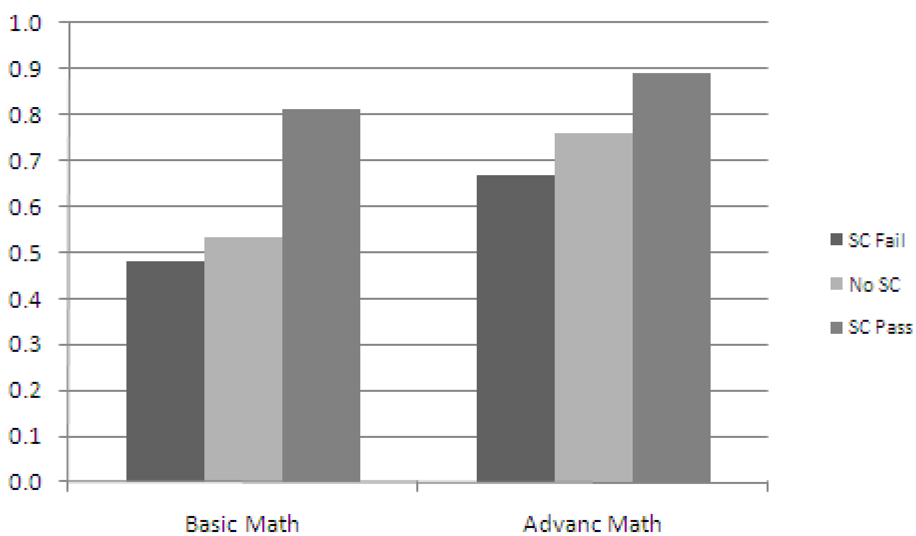

Figure 5: Non-corrected treatment effect of successful participation in summer course on QM1 passing rate.

Propensity scores

The selected instruments of self-perception surveys relevant for learning processes appear to be appropriate resources for potential selection effects. Out of the 42 learning related scales, 30 demonstrate statistically significant differences in means when contrasting summer course participants with non-participants, always in the direction that participants in the summer course achieve on average more favourable scores than non-participants. Only 12 scales do not demonstrate significant differences. Propensity scores or conditional probabilities to participate in the math summer course have been estimated for all 3240 students for which a full data record of background characteristics is available,

using binary logistic regression. Since most of these learning characteristics are associated, in the logistic regression determining the propensity scores, only six of the 42 students' background characteristics appear to be a statistically significant predictor of summer course participation (beyond prior math education). By far the strongest predictor is, in agreement with the design aims of the summer course, the indicator variable distinguishing international students from students with a Dutch prior education. Next, in the order of decreasing impact, the vocational learning orientation, self-perception of cognitive competence (negative), metacognitive knowledge, the constructivist learning conception, and amotivation (negative). The outcomes of the logistic regression, both in terms of statistical significance of predictors and the sign of the regression coefficients, are intuitive: international background, the conception that learning takes place through self-construction of knowledge, and good metacognitive skills strengthen the probability to participate the summer course, whereas lack of learning motivation, and one's perception to be already rather competent in the area of quantitative methods, weaken that probability.
In agreement to procedures advised in the literature (Fraas, 2007; Shadish e.a., 2002; Yanovitzky e.a, 2005), propensity scores are estimated on the basis of the full model, that is all covariates included, both those being statistically significant and those being non-significant.

Propensity score as covariate

After the estimation of propensity scores, the effect analysis is repeated, with the propensity score added as extra predictor, next to the indicator variable of math prior education, and the treatment variable of summer course participation. Depending upon the choice of the effect variable, QM1 total score versus QM1 passing rate, the proper method for doing effect analysis is that of multiple regression in case of the score variable, respectively logistic regression in case of passing rate. Table 2 contains the outcomes of multiple regression of QM1 total score on de predictor variables propensity score and three indicator variables (dummies): math prior education at advanced level, successful participation in the summer course, and non-successful participation in the summer course (this choice of indicator variables implies that math prior education at basic level, and no participation in the summer course, serve as reference groups). Propensity scores and the three indicator variables together explain 11.2% of the variation in total score.

Table 2: Outcomes of effect analysis of summer course participation on QM1 total score with propensity score as covariate

	beta	*t*-value	significance
Propensity score	0.072	4.116	0.000
Advanced math dummy	0.271	15.978	0.000
Successful participation summer course dummy	0.154	8.899	0.000
Non-successful participation summer course dummy	-0.086	-5.002	0.000

Table 2 confirms the picture sketched in the last results section: participants of the summer course stand out from non-participants in terms of background characteristics that have a positive impact on learning. The consequence of this is that in the corrected calculation of the effect of summer course participation, part of explanation of academic success by successful summer course participation is absorbed by the propensity score, as compared to the non-corrected model. The obvious implication of this is that the contribution to explained variation by summer course participation becomes smaller, and the variable is

no more the strongest predictor: the indicator variable distinguishing math prior education at advanced level takes over that position. However, a substantial effect of successful participation in the summer course remains: the beta (standardized regression coefficient) exceeds 50% of the value of the beta of the predictor math at advanced level.
Shifting the focus to passing or failing the QM1 course as outcome variable, a similar picture emerges. The proper method is now that of logistic regression; Table 3 contains the outcomes of such a regression. The explained variation, expressed as the Nagelkerke R^2, equals 8.1%.

Table 3: Outcomes of effect analysis of summer course participation on QM1 passing rate with propensity score as covariate

	B (S.E.)	significance	Exp(B)
Propensity score	1.298 (0.456)	0.004	3.663
Advanced math dummy	0.967 (0.091)	0.000	2.629
Successful participation summer course dummy	1.097 (0.181)	0.000	2.996
Non-successful participation summer course dummy	-0.494 (0.149)	0.001	0.610

For the interpretation of the outcomes of the logistic regression, it is especially the last column of Table 3, which provides the changes in the odds of passing the QM1 course as the result of a unit change in the predictor variable, that deserves attention. Students' background characteristics that influence the participation in the summer course are the strongest determinant of the odds to pass QM1: see the coefficient of the propensity score. Next come the two indicator variables for math at advanced level in prior education and successful participation in the summer course, with the notable detail that predictive power of the summer course participation dummy exceeds that of the advanced math dummy.

Propensity score as stratification variable

The best protection against the impact of potential selection effects in a quasi-experimental research design with non-equivalent groups is offered by the matching approach, in our case through quintile stratification of all subjects on the basis of the propensity scores as stratification variable (Fraas et al., 2007; Guo & Fraser, 2010; Shadish et al., 2002; Yanovitzky et al., 2005). This literature suggests the creation of five strata, based on the quintiles of the distribution of the propensity scores. Each of these five strata this way contains subjects with propensity scores of the same magnitude, so that effect analysis within each stratum is minimally influenced by differences between subjects in their value on the propensity score, providing a correction for the selection effects that depend on background characteristics used in the estimation of the propensity scores. We applied this approach, and repeated the multiple regression analysis described in the last section for each of the five strata created by distinguishing the five quintiles of the propensity score. The outcomes of these regression analyses are collected in Table 4.

Table 4: Outcomes of effect analysis of summer course participation on QM1 total score with propensity score as stratification variable

Stratum 1: propensity score < 0.055	beta	*t*-value	significance
Propensity score	0.062	1.715	0.087
Advanced math dummy	0.365	10.065	0.000
Successful participation summer course dummy	0.061	1.674	0.095
Non-successful participation summer course dummy	-0.118	-3.251	0.001
Stratum 2: 0.055 < propensity score < 0.117			
Propensity score	0.022	0.563	0.573
Advanced math dummy	0.310	8.124	0.000
Successful participation summer course dummy	0.173	4.511	0.000
Non-successful participation summer course dummy	-0.050	-1.308	0.191
Stratum 3: 0.117 < propensity score < 0.166			
Propensity score	0.008	0.201	0.841
Advanced math dummy	0.222	5.697	0.000
Successful participation summer course dummy	0.127	3.257	0.001
Non-successful participation summer course dummy	-0.090	-2.315	0.021
Stratum 4: 0.166 < propensity score < 0.217			
Propensity score	0.041	1.063	0.288
Advanced math dummy	0.195	4.990	0.000
Successful participation summer course dummy	0.146	3.735	0.000
Non-successful participation summer course dummy	-0.117	-2.996	0.003
Stratum 5: 0.217 < propensity score			
Propensity score	0.007	0.191	0.849
Advanced math dummy	0.268	7.071	0.000
Successful participation summer course dummy	0.210	5.485	0.000
Non-successful participation summer course dummy	-0.068	-1.769	0.077

Stratification appears to achieve exactly what it is intended for: the influence of students' background characteristics, expressed as propensity score, is statistically insignificant in all five strata, where it had been the strongest predictor before stratification taking place. Since treatment and control group overlap in all five strata, we can do an even more thorough check of the adequacy of the propensity score based stratification by performing difference in means tests for all 42 predictors in all five strata (Guo & Fraser, 2010). Doing these test at a 5% significance level, we find 3, 2, 1, 0, 1 significant differences in strata 1-5 respectively, so 1.4 on average in each stratum. This compares quite well to the expected 2.1 significant difference one expects when testing at 5% level. As a reminder, in the complete data set we found 30 out of 42 differences to be significant, indicating that stratification indeed succeeds in taking away any selection effect induced by these 42 predictors.

Other regression outcomes are quite similar to the outcomes achieved on the complete data set, with the first stratum producing slightly deviant outcomes. In that first stratum, the quintile of students with the lowest score for students' background characteristics that contribute to participation in the summer course, the positive effect of successful participation in the summer course is outshined by the negative effect of failing the summer course. This different position of the first stratum is an artefact of the way the strata are created: due to the very low propensity scores of students in this first stratum, that stratum counts by far the fewest number of participants of the summer course, and amongst those participants, the large majority drops out of the summer course (amongst the 660 students in this stratum, there are only 23 participants in the summer course, of which 17 drop out). The other four strata, each containing many more summer course participants and especially many more successful participants, all demonstrate the same patterns as found in the full data set: the largest effect is that of the indicator variable of prior math education at advanced level, with the treatment effect of successful participation in the summer course in the second position, having an effect size of at least 50% of the effect size of advanced math. Average treatment effect (ATE; Guo & Fraser, 2010), calculated by averaging the stratum-specific differences of the mean QM1 scores, equals t = 4.75, which is clearly significant at 5% or 1% level. Focusing on the four strata with a substantial amount of summer course participants, so excluding the first stratum form the calculation of the ATE due to the very small number of participants, the statistical significance even achieves a value of t = 11.27.

Redoing the logistic regression analysis to determine the treatment effect of participation in the summer course on the passing rate of the QM1 course after stratifying the data set into five strata based on the quintiles of the distribution of the propensity scores, we achieve equivalent outcomes. Within each of the strata, the propensity score has no statistically significant effect anymore on passing rate. And except for the first quintile, where success in the summer course appears to be insignificant for the QM1 passing rate, the other four strata demonstrate significant effects of both math prior education at advanced level and successful participation in the summer course, with the odds-ratio of the last everywhere exceeding the one of the first. The average treatment effect, ATE, for the success rate equals t = 2.59 when calculated over all five strata, and up to t = 9.18 when calculated over the four non-sparse populated strata, so statistically significant at 5% level.

Summarizing the outcomes of the statistical analyses: a direct comparison of academic success of students successfully participating the summer course, and that of non-participants, demonstrates a large treatment effect. The treatment effect exceeds the effect of math schooling at advanced level, with math schooling at basic level as reference. However, part of this effect may be caused by selection bias due to the voluntary participation in the summer course. In order to decompose the total treatment effect into such a selection effect, and a corrected treatment effect, propensity scores of summer course participation were estimated based on a very wide range of learning related students' background data. After stratification based on these propensity scores the presumption of the existence of selection bias was confirmed, and correction for this selection effect did indeed diminish the treatment effect, but still a very substantial treatment effect remained, of the size of about half of the effect size of being schooled at advanced level.

CONCLUSION AND DISCUSSIONS

Many first year university programs contain elements of remedial education: only after revisiting topics that have been taught in the last grades of high school, the program continues with the coverage of completely new topics. Our study suggests that such an approach to bridging is insufficient: important deficiencies exist in the mastery of more basic mathematical competencies, and these stay outside the scope of such a refreshment approach. Based on our experiences with entry tests over a sequence of years, the UM has opted for a very broad and basic coverage of topics in our math summer course, and individual learning routes controlled by repeated administrations of adaptive, diagnostic tests.

Effect analysis suggests that this kind of bridging education is very effective: the non-corrected effect of successful participation in the summer course exceeds the effect of math prior schooling at advanced level, with basic schooling as reference. The relevant research design of this study is however that of a quasi-experimental setup with non-equivalent groups, requiring a correction of the calculated treatment effect for potential selection effects. Correction on the basis of the propensity score method indicates that indeed part of the non-corrected treatment effect should be attributed to the circumstance that participants in the summer course possess more favourable background characteristics for achieving academic success in their study, than students who choose not to participate in the summer course. At the same time, after correction for the non-equivalent composition of both groups, a substantial treatment effect remains, in the order of size of about half the effect size of being educated at advanced math level in high school.

The outcomes of the effect analysis suggest that the chosen format for bridging education, to know that of an online summer course with a very broad coverage of basic math topics, and learning controlled by individual, adaptive testing, is a very efficient one to bridge math skills deficiencies. The average study load of being successful in the summer course is much less than the difference in study load between high school math education at advanced, versus basic level. Notwithstanding, the treatment effect of successful summer course participation is about 50% of the effect size of advanced prior math education. The question if such an outcome is unique for the chosen format of bridging education, or that other formats, like offering additional bridging classes parallel to regular education as part of the first year university program –a format used by many Dutch and European bridging initiatives–, is as effective, suggests to be an important question for future research.

ACKNOWLEDGEMENTS

The authors would like to thank the EU Lifelong Learning programme funding the S.T.E.P. project and the Dutch SURF funding the NKBW project being part of the Nationaal ActiePlan e-Learning, which enabled this research project. This publication reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

REFERENCES

Attewell, P., Lavin, D., Domina, T., Levey, T. (2006). New Evidence on College Remediation. *Journal of Higher Education*, *77*(5), 886-924.

Bahr, P. R. (2008). Does Mathematics Remediation Work? : A Comparative Analysis of Academic Attainment among Community College Students. *Research in Higher Education*, *49*(5), 420-450.

Bailey, T. (2009). Challenge and Opportunity: Rethinking the Role and Function of Developmental Education in Community College. *New Directions for Community Colleges*. No. 145, Spring 2008, 11-30.

Bettinger, E. P., Long, B. T. (2009). Addressing the Needs of Underprepared Students in Higher Education : Does College Remediation Work? *The Journal of Human Resources*, *44*(3), 736-771.

Brants, L., Struyven, K. (2009). Literature Review on Online remedial Education: A European perspective. *Industry and Higher Education*, *23*(4), 269-276.

Calcagno, J. C., Long, B. T. (2008). *The Impact of Postsecondary Remediation Using a Regression Discontinuity Approach: Addressing Endogenous Sorting and Noncompliance*. NBER working paper series, 14194.

Doignon, J. P., Falmagne J. C. (1999). *Knowledge Spaces*. Berlin: Springer.

Dweck, C. S. (1999). *Self-theories: their role in motivation, personality, and development*. Philadelphia, PA: Psychology Press.

Eccles, J.S., Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Reviews Psychology, 53*, 109-132.

Elshout-Mohr, M., van Daalen-Kapteijns, M. M., Meijer, J. (2001). *Constructie van het instrument "Rapportage Autonoom Studeren"*. Amsterdam: SCO-Kohnstamm Instituut en Instituut voor de Leraren Opleiding (ILO).

Entwistle, N. J., Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behavior and influences of learning environments. *International Journal of Educational Research*, *41*(6), 407-428.

Falmange, J., Cosyn, E., Doignon, J., Thiéry, N. (2004). *The Assessment of Knowledge, in Theory and in Practice*. Retrieved December 1, 2010, from http://www.aleks.com/about_aleks/Science_Behind_ALEKS.pdf.

Fraas, J. W. (2007). *A Comparison of Propensity Score Analysis to Analysis of Covariance: A Case Illustration*. Invited paper American Educational Research Association annual conference Chicago.

Grant, H., Dweck, C. S. (2003). Clarifying Achievement Goals and Their Impact. *Journal of Personality and Social Psychology, 85*(3), 541–553.

Guya, F., Mageau, G.A., & Vallerand, R.J. (2003). On the hierarchical structure of self-determined motivation: A test of top-down, bottom-up, reciprocal, and horizontal effects. *Personality and Social Psychology Bulletin*, *29*(8), 992-1004.

Jamelske, E. (2009). Measuring the impact of a university first-year experience program on student GPA and retention. *Higher Education, 57*(3), 373-391.

Kozeracki, C. A. (2005). Editor's notes. *New Directions for Community Colleges*, No. 129, Spring 2005, 1-4.

New Directions for Community Colleges (2008). No. 144, Winter 2008: *Are Community Colleges Underprepared for Underprepared Students*?

Onderwijsraad. (2006). Reinforcing Knowledge in Education; study. Den Haag: Onderwijsraad. English summary. Retrieved December 1, 2010, from http://www.onderwijsraad.nl/upload/english/publications/reinforcing_knowledge_in_education.pdf.

Onderwijsraad. (2007). Reinforcing Knowledge II; advice. Den Haag: Onderwijsraad. English summary. Retrieved December 1, 2010, from www.onderwijsraad.nl/upload/english/publications/samenvatting_versteviging_van_kennis_ii_engels.pdf.

Onderwijsraad. (2008). A Successful Start in Higher Education; advice. Den Haag: Onderwijsraad. English summary. Retrieved December 1, 2010, from http://www.onderwijsraad.nl/upload/english/publications/a_succesfull_start_in_higher_education.pdf

Ratelle, C. F., Guay, F., Vallerand, R. J., Larose, S., Senécal, C. (2007). Autonomous, controlled, and amotivated types of academic motivation: A person-oriented analysis. *Journal of Educational Psychology*, *99*(4), 734-746.

Rienties, B., Kaper, W., Struyven, K., Tempelaar, D. T., Gastel, L. van, Vrancken, S., Jasińskad, M., Virgailaitė-Mečkauskaitėe, E. (in press). Describing the current transitional educational practices in Europe. *Interactive Learning Environments.*

Ryan, R.M. & Deci, E.L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well being. *American Psychologist*, *55*(1), 68-78.

Schneider, M., Carnoy, B., Kilpatrick, J., Schmidt, W. H., & Shavelson, R. J. (2007). Estimating Causal Effects: *Using Experimental and Observational Designs*. Washington, DC: AERA.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Boston, NY: Houghton Mifflin Company.

Tempelaar, D. (2006). The role of metacognition in business education. *Industry & Higher Education*, 20(5), 291-298.

Tempelaar, D. T., Gijselaers, W. H., Schim van der Loeff, S., & Nijhuis, J. F. H. (2007). A structural equation model analyzing the relationship of student personality factors and achievement motivations, in a range of academic subjects. *Contemporary Educational Psychology*, 32 (1), 105-131.

Tempelaar, D. T. & Rienties, B. (2008), Remediating Summer Classes and Diagnostic Entry Assessment in Mathematics to Ease the Transition from High School to University. In *Proceedings Student Mobility and ICT: Can E-LEARNING overcome barriers of Life-Long learning* (pp. 9-17). Maastricht: FEBA ERD Press.

Vallerand, R.J., Pelletier, L.G., Blais, M.R., Brière, N.M., Sénécal, C., & Vallières, E.F. (1993). On the assessment of intrinsic, extrinsic, and amotivation in education: Evidence on the concurrent and construct validity of the Academic Motivation Scale. *Educational and Psychological Measurement, 53*(1), 159-172.

Vermunt, J. D. (1996). *Leerstijlen en sturen van leerprocessen in het Hoger Onderwijs*. Amsterdam/Lisse: Swets & Zeitlinger.

Yanovitzky, I., Zanutto, E., Hornik, R. (2005). Estimating causal effects of public health education campaigns using propensity score methodology. *Evaluation and Program Planning*, *28*(2), 209-220.

Key Terms & Definitions

Adaptive e-tutorial: web-based tutorial that intelligently adapts to student interaction and knowledge level.

ALEKS: Assessment and LEarning in Knowledge Spaces, adaptive e-tutorial for mathematics and other subjects that allow for hierarchical classification in knowledge spaces.

Bridging education: remedial education with the aim to ease the transition between subsequent educational systems.

Effectiveness: systematic investigation of the impact of educational innovations.

Postsecondary remediation: bridging education in the transition from high school to college or university.

Propensity score: probability that a student opts for participating in voluntary remedial education, or to opt out, given this students' background characteristics: the covariates.

Quasi-experimental design: research design in an educational effectiveness study in which effectiveness is based on the comparison of achievements of two groups of students, the experimental group and the control group, but where groups are not created through random assignment.

Selection effect or selection bias: distortion of statistical analysis resulting from experimental and control groups being non-equivalent. In educational research, selection effects result quite often from self-selection: instead of comparing groups designed by random assignment, group composition is the result of students opting to participate, or not, different types of education.