Student profiling in blended learning: relevance for learning analytics

ONBETWIST Werkpakket 5

Deliverable 5.4.7, Maart/April 2013

Dirk Tempelaar, Maastricht University School of Business and Economics

Bart Rienties, University of Surrey

Bas Giesbers, Maastricht University School of Business and Economics

Hans Cuypers, Eindhoven University of Technology, Department of Mathematics and Computing Science

André Heck, University of Amsterdam, Faculty of Science

Evert van de Vrie, Open University Netherlands, Faculteit Informatica

Henk van der Kooij, Utrecht University, Freudenthal Institute

(artikel ingediend)

Abstract

Research into blended learning seeks to understand the different ways learning can be supported by technology enhanced education, by contrasting tool use of students with different profiles. In applications of learning analytics, researchers seek to integrate individual learning dispositions with computer systems generated feedback in order to support learning regulation. This explorative, empirical study aims to demonstrate that combining both of these research traditions has much potential. In a large, undergraduate course of 986 students, two digital platforms for practicing and formative testing of introductory mathematics and statistics shaped the online component of the blend, together with problem-based learning as the face-to-face component. Within this design, both system tracking data from the digital platforms, and a rich set of learning dispositions based on social-cognitive learning theories, distinguishing adaptive and maladaptive learning behaviours originating from different implicit theories, were collected. We demonstrate that learning dispositions derived from these meaning systems are relevant for creating profiles of students with different tool use behaviour. From a learning analytics perspective, we demonstrate that both early tracking data, and learning dispositions, are important and relatively independent predictors of learning behaviour in the course, allowing for effective feedback to regulate individual learning.

Introduction

Learning analytics applications often do not succeed in achieving their full potential. In a recent plea to improve the state of learning analytics applications by introducing a Learning Analytics Acceptance Model, Ali, Asadi, Gašević, Jovanović, and Hatala (2013, p.121) describe many current applications even as 'superficial', containing no more than 'simple statistics such as, when and how many times students logged in, or low-level data about students' interactions with learning content'. Although this superficiality is easily understood by the complexity of collecting and integrating other learning related data beyond system track data, blue prints for learning analytics designs that integrate track data with other data sources are already around for some time. The design developed by Buckingham Sum and Deakin Crick (2012), proposing a learning analytics framework based on the integration of student learning dispositions and the use of generic learning management systems, suggests to be a crucial step in integrating more diverse data types.

A prime data source for most learning analytic applications is data generated by learner activities, such as learner participation in technology enhanced learning systems. This information is frequently supplemented by background data retrieved from learning management systems and other concern systems, as for example accounts of prior education. A combination with intentionally collected data, such as self-report data stemming from student responses to surveys, is however the exception rather than the rule. In their conceptual contribution, Buckingham Sum and Deakin Crick (2012) propose a learning analytics infrastructure that combines learning activity generated data with learning dispositions, values and attitudes measured through self-report surveys and feedback to students and teachers through visual analytics. Their proposal considers, for example, spider diagrams to provide learners insights into their relative positions of learning dispositions, values, attitudes and skills. Measurement of this 'learning power', as Buckingham Sum and Deakin Crick term the collection of this personal, intentionally collected data, is in their study through the self-report instrument ELLI, Effective Lifelong Learning Inventory.

Applications of learning analytics seem to be most fruitful in technology enhanced programs and large scale classes using technology enhanced learning. Undergraduate courses are the natural setting for this. Since the operationalization of learning dispositions is context dependent, we will adopt in this study Buckingham Sum and Deakin Crick's framework for integrating learning dispositions into technology generated learning feedback, but propose an alternative operationalization of these dispositions. We do so by connecting to another research strand: that of student profiling in blended or hybrid learning, or generally, technology enhanced learning (see e.g. Lust, Juarez Collazo, Elen, & Clarebout, 2012; Lust, Vandewaetere, Ceulemans, Elen, & Clarebout, 2011). In this research strand modern learning theories, such as social-cognitive learning theories, serve the role of providing the instruments that allow for student profiling. Theoretical frameworks of meaning systems surrounding implicit theories of intelligence, distinguishing maladaptive views that intelligence is fixed versus adaptive views that it is malleable (Dweck, 1999; Dweck & Master, 2008), have been demonstrated to play a crucial role in explaining learning behaviour in technology enhanced learning (Greene, Costa, Robertson, Pan & Deekens, 2010). Beyond implicit theories of intelligence itself, other components of these meaning system are effort beliefs, goal setting behaviour, intrinsic and extrinsic motivation, and learning strategies and learning regulation. In the current empirical, explorative contribution, we will integrate system tracking data from two digital learning platforms with intentionally collected learning dispositions (based on this meanings system framework) from a wide range of instruments in order to create student profiles. We will investigate which role these profiles can play in serving a feedback function, both for students and their teachers, in shaping learning behaviour in a technology enhanced learning environment. Specific focus in this research question is the contribution of both data sources of our learning analytics application: system data and intentionally collected student dispositions, and within the last data component, the contribution of individual instruments to the goals of learning analytics.

Meaning system of implicit theories

The theoretical framework applied in this study to operationalize learning dispositions, which we share with the Greene et al. (2010) study, is that of meaning systems based on implicit theories of intelligence. This framework has been developed by Carol Dweck, and most empirical research is firmly rooted in her monograph Dweck (1999), that describes the functions and origins of the meaning system and its components: implicit theories of intelligence, effort beliefs, goal setting behaviour, intrinsic and extrinsic motivation, and learning-regulation and learning strategies. Implicit theories of intelligence refer to beliefs people develop about the nature of their intelligence, and contrasts two opposite beliefs: that of the 'entity theorists', who view intelligence as being a fixed internal characteristic, and the 'incremental theorists', who believe that intelligence is malleable and can be cultivated by learning or practicing. Students endorsing an entity view are hypothesized to see effort as a negative characteristic, signalling lack of intelligence, whereas those with an incremental view develop a positive effort belief: exerting effort is the key to cultivating intelligence (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 1999). In turn, implicit theories and effort views impact students' goal setting behaviour. Students who view their intelligence to be fixed, find their goals in learning restricted to outperforming other students: normative performance goals. In contrast, students who regard their intelligence as something they can train, might learn out of interest (mastery goal), or the wish to achieve good grades (appearance performance goal). Next, "entity theorists" and "incremental theorists" are hypothesised to learn for different motives: more intrinsically motivated in the incremental theory case, and a tendency for stronger extrinsic motivation in the entity theory case. And lastly, implicit theories influence learning strategies and regulation according to the meaning system framework: incremental theorists tend to deep learning in a self-regulated manner, whereas entity theorists are inclined to surface learning approaches, who are in need for external regulation. Empirical research confirms the importance of endorsed meaning systems in shaping learning; focusing on learning in a technology enhanced context, for example, Greene et al. (2010) demonstrate the impact of implicit theories on self-regulated learning,

Dweck and co-authors have appended this theoretical framework with instruments for operationalizing constructs as incremental and entity theories of intelligence, positive and negative effort beliefs, goal choice, and learning, appearance performance and normative performance goals (Blackwell, Trzesniewski, & Dweck, 2007; Dweck, 1999; Grant & Dweck, 2003). These instruments have in common that they all distinguish two opposite positions: the adaptive versus the maladaptive ones. To better serve the student feedback function in our learning analytics project, we have extended the meaning system framework and its range of instruments with three theoretical learning models (i.e. academic motivations, subject attitudes, learning styles, and motivation and engagement). The corresponding instruments focus specifically on students' concrete learning behaviour, as opposed to the more subtle learning conceptions and orientations expressed by students' beliefs or implicit theories of intelligence. In doing so, we opted for frameworks that again allow for profiling into adaptive and maladaptive facets of learning. Self-determination theory (Ryan & Deci, 2000; see also Rienties, Giesbers, Tempelaar, Lygo-Baker, Segers, & Gijselaers, 2012), distinguishing autonomous and controlled aspects of motivation, learning styles theory (Vermunt & Vermetten, 2004; see also Hauptman & Cohen, 2011), distinguishing adaptive and maladaptive forms of learning strategies and learning regulation, and the motivation and engagement wheel framework of Martin (2007; see also Liang, Liu, Zhang, & Yang, 2010), composed of adaptive and maladaptive learning cognitions and behaviours, and subject attitudes serve this role.

Material and methods

1 Participants and educational context

This study involves the 2012/2013 cohort first-year students of a Business and Economics School in the Netherlands. This school's programme deviates from a conventional European university education in two important ways: it employs a student-centred learning approach of problem-based learning (PBL), and it has a strong international orientation— all degrees are offered fully in English and attract primarily non-Dutch students. Out of the 986 students on which this study is based, 78% had an international background (mostly European, with 37% originating from German speaking countries in Europe). 63% of the students were male and 37% were female. The participants' ages ranged from 17-29, with an average age of 19.7 years. The quantitative methods course that represents the context of this study contains an introduction into mathematics and statistics, and is delivered in the first eight weeks in the first term of the program, with a 50% study load.

2 The learning blend

The technology enhanced learning environment investigated in this paper combines face-to-face learning according to the problem-based learning principle, with online learning using two different MyLab (Pearson) tools at the same time: MyMathLab (MML), to practice and test mathematics, and MyStatLab (MSL), to practice and test statistics (see also McKenzie, Perini, Rohlf, Toukhsati, Conduit, & Sanson, 2013, Romero-Zalvidar, Pardo, Burgos, & Kloos, 2012, or Tempelaar, Cuypers, Van de Vrie, Heck, & Van der Kooij, 2013, for a discussion of the use of MyLabs in the context of learning analytics). The main principles of PBL are collaborative learning in small groups of students, steered by open-ended problems (Wilkerson & Gijselaers, 1996). In PBL, students take the stage and perform the leading part in small groups of fourteen students, where they discuss open, unstructured scientific and practical problems prepared by teachers.

The introduction of the online components MML and MSL was part of a national project into testing and test-steered learning, installed by the Dutch organisation SURF. The aim of the project is to experiment with diagnostic and formative testing in designing personalized, adaptive programs for learning introductory mathematics and statistics. In the local implementation of the program, a secondary aim of the experiment was to investigate how learning analytics could support the regulation of individual learning by providing feedback to both students and teachers.

The main reason for installing personalized learning paths, based on formative testing, is the strong heterogeneity in mastery and prior education in an introductory course with an international inflow. Every activity in MML or MSL starts with a test item, such as the one depicted in Figure 1. Students with mastery will solve the item and provide an answer, and swiftly move on to the next activity. Students without full mastery may ask for assistance (Help Me Solve This). Students with insufficient mastery may ask the tool to provide a step-by-step work example (View an Example). In both of the latter cases, the student will receive a new, equivalent but different, version of the activity as a next step in the learning process (i.e. items are parameterised) to solve it without using the scaffolds of the tool.

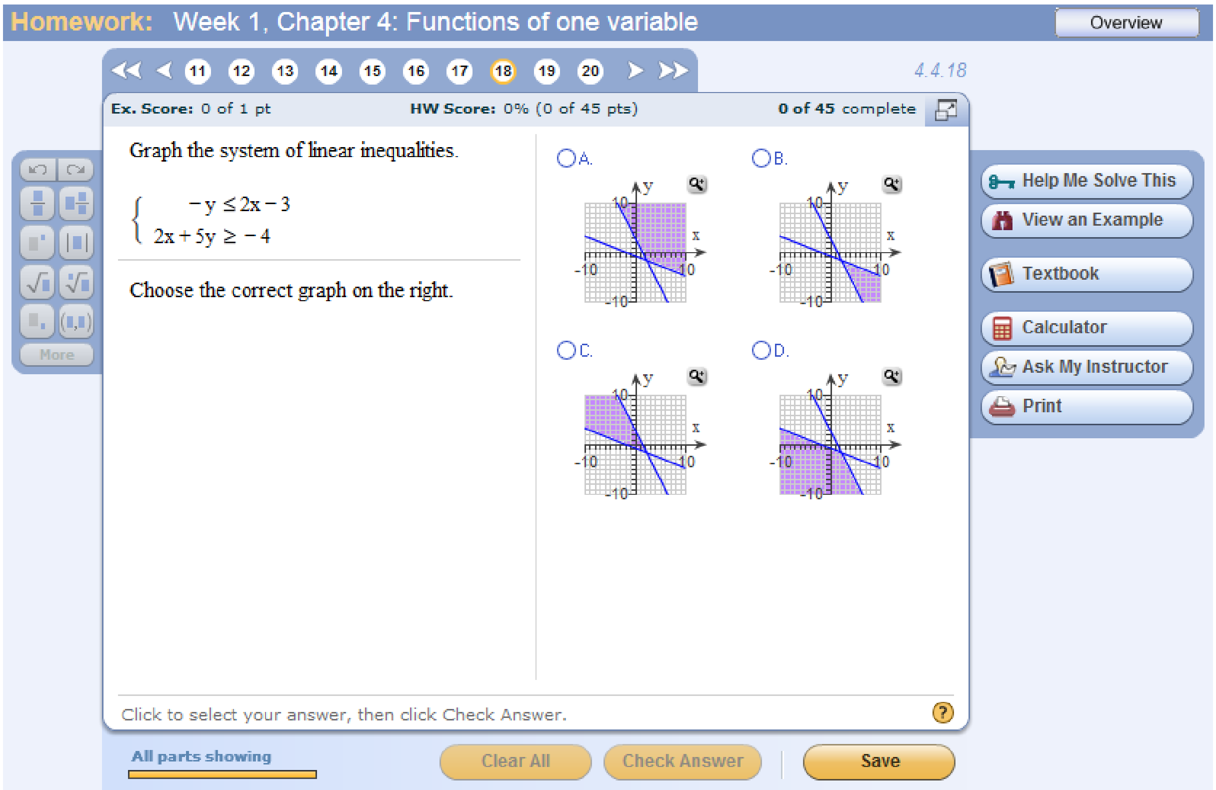


Figure 1. Screenshot of MML learning environment.

3 Learning analytic materials of system generated type

In line with the Buckingham Sum and Deakin Crick (2012) proposal, differentiating, between data generated by computer systems and data intentionally collected, we focus first on the systematically generated data components in section 3.3.1-3.3.5. The MML and MSL learning platforms provide both tracking data related to students' activities in homework and quizzing modes, and performance data. Other performance data is from the exam system. Registration systems provide demographic data as nationality, prior education, age and gender. Course specific learning management systems collect performance on entry tests.

3.1 Registration system

The broad range of data available in the registration system is condensed into three indicator (dummy) variables. The first is an indicator for Female students (female students generally outperforming male students), an indicator for International students, and an indicator of advanced mathematics in prior education (MathMajor). All high school systems distinguish (at least) two levels of mathematics education: A-level, preparing for sciences (advanced track, like Calculus BC), or preparing for social sciences (basic track, like Calculus AB). Other variables are less relevant, mainly because of too little variation (such as age) or too much variation (44 different nationalities, but only seven nationalities counting 10 or more students).

3.2 Entry testing

Entry tests for mathematics and statistics, stemming from the national SURF project, are applied to measure variation in prior mastery in the first day of the course. The mathematics entry test score (MathEntry) allows a break down into the topics Algebraic Skills (AlgebraEntry), Logarithms & Powers (LogPowerEntry), and Equations (EquationsEntry).

3.3 Tool tracking

Both MML and MSL tools provide important logging data on a weekly basis: mastery of the weekly topics, expressed as the proportion of all weekly homework activities successfully completed (MML-Mastery and MSL-Mastery), and weekly connect time needed to achieve those mastery levels (MML-Time and MSL-Time). Mastery and time data are strongly collinear (for mathematics, r = 0.66, for statistics, r = 0.72, measured over the whole course), and for that reason, a third tracking index is determined as an efficiency measure: MML-Efficiency and MSL-Efficiency are defined as the ratio of mastery and time, or the number of items a student can do per hour of connect time.

3.4 Tool performance

To stimulate students to fully utilize the e-tools, two-weekly quizzes were installed that allow students to earn a bonus score on top of their final exam score. These quizzes were administered in the MML and MSL tools, and consist of subsets of the homework items students have practiced in the immediate two weeks before. However, all items are parameterised, so quiz items are best described of equivalent versions of the homework items. Mathematics and statistics quiz performance generate bonus scores toward the exam: MathBonus and StatsBonus.

3.5 Exam system

The performance in the final exam allows a breakdown into two partial scores: score in the mathematics part (MathExam) and in the statistics part (StatsExam).

4 Intentionally collected learning analytic materials: the dispositions

Learning dispositions have been measured by a wide range of seven self-report surveys referring to several constructs from the meaning systems framework and related social-cognitive learning theories.

4.1 Implicit theories of intelligence

Measures of both entity (intelligence as fixed) and incremental (intelligence is malleable) implicit theories of intelligence were adopted from Dweck's Theories of Intelligence Scale – Self Form for Adults (1999). This scale consists of eight items: four Entity Theory statements and four Incremental Theory statements.

4.2 Effort beliefs

Measures of Effort beliefs were drawn from two sources: Dweck (1999) and Blackwell (2002). Dweck provides several sample statements which are designed to portray Effort as a Negative concept—i.e. exerting effort conveys the view that one has low ability, and Effort as a Positive concept—i.e. exerting effort is regarded as something which activates and increases one’s ability. In addition, Blackwell´s full set of Effort beliefs (2002) were used, comprising five positive and five negative items (see also Blackwell et al., 2007).

4.3 Achievement Goals.

Goals have been operationalized by the Grant and Dweck (2003) instrument. Different from the dominant approach of differentiating approach and avoidance valences within a multiple goal perspective, this instrument describes several goal facets without using the avoidance valence. As with other multiple goal perspectives, the instrument defines goals to be of mastery/learning type, versus performance type. In the learning goal definition, it distinguishes the two alternative forms: Challenge-Mastery and Learning. And without using the avoidance valence, it distinguishes four types of performance goals—two of appearance nature: Outcome and Ability Goals, and two of normative nature: Normative Outcome and Normative Ability Goals.

4.4 Motivation and Engagement

The Motivation and Engagement Scale –University/College (MES-UC: Martin, 2007) measures university or college students’ motivation and engagement. The MES-UC consists of four scales and eleven subscales subsumed under the four scales. The first scale is the adaptive cognition scale, which reflects students’ positive attitudes and orientations to academic learning, and is composed of the subscales Self-Belief, Valuing School, and Learning Focus. The second scale, adaptive behaviour, reflects students’ positive behaviours and engagement in academic learning, and contains the subscales Persistence, Planning, and Study Management. Third, students’ attitudes and orientations that inhibit academic learning are collected in the impeding or maladaptive cognition scale, including the subscales Anxiety, Failure Avoidance, and Uncertain Control. Finally, the maladaptive behaviour scale reflects on students’ problematic learning behaviours, and includes the subscales Self-Handicapping and Disengagement.

4.5 Academic motivations.

The Academic Motivation Scale (AMS; Vallerand et al., 1992) is based upon Ryan and Deci’s (2000) model of autonomous and controlled motivation. The AMS consists of 28 items, to which students respond according to the question stem “Why are you going to college?” There are seven subscales on the AMS, of which three belong to the intrinsic motivation scale: Intrinsic Motivation to Know, Intrinsic Motivation to Accomplish, and Intrinsic Motivation to Experience Stimulation. Together with Identified Motivation, these constitute autonomous motivation. Controlled motivation is composed of subscales Introjected Motivation and External Regulation. All subscales together constitute a motivational continuum reflecting the degree of a student's self-determined behaviour. The final scale, A-Motivation, indicates the extreme of the continuum: the absence of regulation, either externally directed or internally.

4.6 Learning styles

The Inventory of Learning Styles (ILS) instrument, developed by Vermunt (Vermunt & Vermetten, 2004), has been used to assess preferred learning approaches. In this study, we apply the cognitive processing strategies and metacognitive regulation strategies of ILS, which are composed of five different scales. The two processing strategy scales relating and structuring and critical processing together compose the Deep Learning strategy, whereas memorizing and rehearsing, together with analysing, constitute the Stepwise Learning strategy (also called surface learning in several theories of learning). The third processing strategy is Concrete Learning (also called strategic learning). Similarly, two regulation scales self-regulation of learning processes and self-regulation of learning content together compose the strategy Self-Regulation, hypothesized to be prevalently used by deep learning students. The two regulation scales external regulation of learning processes and external regulation of learning results constitute the External Regulation strategy, supposed to be characteristic for stepwise learners. The third regulation strategy signals any absence of self or external regulation: Lack of Regulation.

4.7 Subject attitudes

Students' subject attitudes based on the expectancy * value theory (Wigfield & Eccles, 2000) are measured with the Survey of Attitudes Toward Statistics (SATS) developed by Schau and co-authors (Schau, Stevens, Dauphinee, & Del Vecchio, 1995). The instrument distinguishes six adaptive aspects of students' subject attitudes, amongst which two expectancy factors that deal with students' beliefs about their own ability and perceived task difficulty: CognitiveCompetence and NotDifficult, and three subjective task-values Affect, Interest and Value. The sixth aspect, Effort, is assumed to be the outcome of the process of weighting expectancy against value.

5 Procedure

In the first weeks of the term, the students were asked to complete the self-report questionnaires that shape the intentionally collected data component, as described above. The instruments apply a 7-point response Likert scale, ranging from 1 (strongly disagree) to 7 (strongly agree). These self-reports were introduced to supply students with relevant data in their individual statistical projects, but next to that, provide students feedback on their current learning approach(es) and what could be done to improve these, and to serve research purposes. All students consented to their data, in an anonymous format, being used for educational and research purposes. Since dispositions data is a necessary ingredient for the required student project, complete response of all students for all instruments is implied.

Students had access to the systems generated data component on a continuous basis, and given the student-centred focus of the program, utilization of this feedback data to regulate individual learning had been the prime responsibility of students. Teachers received MML and MSL track data on a weekly basis, to assist students in their monitoring and learning regulation. Learning dispositions, the intentionally collected component of the data, became available to students after completion of respective surveys. This feedback referred to absolute scoring, as well as relative scoring with peer students as benchmark, with graphical tools similar to the spider diagrams proposed by Buckingham Sum and Deakin Crick (2012). In line with recommendations of McKenzie et al. (2013), we provided feedback to students based on individual sources of information, thus describing partial student profiles (rather than holistic scores across the seven learning disposition instruments). The integration of these partial profiles into an overall profile would have required modelling outcomes as discussed in this paper, that allow to specify the relative weights of the several sources of information.

6 Statistical analyses

The investigation of relationships between learning behaviour at the one side, and aspects of student profiles at the other side, is the major aim of this study. We will do so in a bivariate context first, using simple correlations. Variables being part of the learning dispositions appear to be collinear. In a multivariate context, this collinearity gives rise to suppression effects: meaningful relationships between disposition variables and learning behaviour may be dominated by other and stronger relationships, and remain hidden due to these suppression effects. To avoid so, we opt for using the bivariate context.

Next to investigating individual dispositions, we aim to cluster students on measured learning behaviour: track data of e-tool use in the MML and MSL learning platforms. This approach resembles the modelling approach in McKenzie et al. (2013) of categorizing all students into four groups on the basis of ML use patterns, be it that we apply K-means cluster analysis to create learner groups. As a last step in the statistical analysis, we investigate how well we can predict cluster membership of individual students, on the basis of both components of learning analytics data, early track data and learning dispositions, or a combination of the two types of data. Multinomial logistic regression (the extension of binary logistic regression to more than two categories) is used to predict cluster membership. In the comparison of the quality of different multinomial logistic models, all based on different sets of information, we apply the Nagelkerke pseudo R2 and the percentage of correct predictions in each of the clusters, as fit criteria. All analyses are performed in SPSS 20.

Results

4.1 Descriptive statistics of tool use as measure learning behaviour

Out of the 20 hours of weekly study load available for the course, students spent an average 7.5 hours per week practicing in the MML and MSL platforms (there are 8 weekly class hours, with the remaining time being self-study for the tutorial sessions outside the e-platform; self-reported total study time for the course is 20.1 hours/week). Table 1 provides a breakdown of tool use data into individual weeks, and includes, beyond connect time data, also mastery data, expressed as a percentage of all assigned weekly homework activities.

Table 1: Tool use descriptive statistics

	Week1	Week2	Week3	Week4	Week5	Week6	Week7	Overall
MML-Mastery	84%	84%	80%	75%	72%	74%	68%	77%
MSL-Mastery	84%	84%	81%	80%	72%	70%	66%	77%
MML-Time (hours)	5.22	2.80	5.96	6.55	5.35	2.54	3.52	31.94
MSL-Time (hours)	2.18	1.96	2.72	3.57	5.01	2.98	2.41	20.84

Mastery is at a relative high level: in both tools, overall mastery is beyond 75%. For a large majority of students, mastery is even nearly complete. This is also visible from Figure 2, which depicts the outcomes of a cluster analysis on seven weekly MML mastery data combined with seven weekly MSL mastery data. K-means cluster analysis results in an optimal solution consisting of four clusters. Cluster 4 represents the cluster of students achieving approximately full mastery in both tools in all weeks. It is by far the largest cluster, containing 652 of the 986 students (66%), thus explaining why average mastery as laid out in Table 1 is at such high level. Cluster 1, containing 133 students, is composed of students who opted out from using the e-tool component in the learning blend. In between, the Cluster 2 (98) and Cluster 3 (103) contain students who opted in during the first weeks, but strongly decreased the intensity of usage over the later weeks of the course.

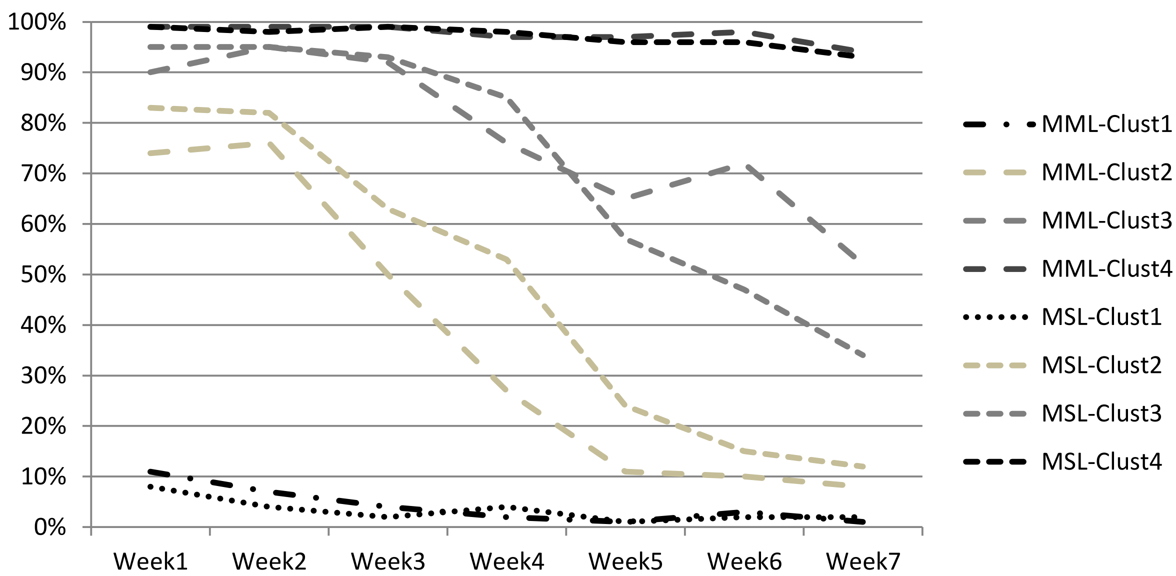


Figure 2. Average mastery levels in MML and MSL of four clusters of students.

The relevance of creating student profiles based on tool use can be seen by comparing passing rates of the four clusters: passing rates for the course are 38% for students in cluster 1, a similar 39% for cluster 2 students, next jump to 61% for cluster 3, to finish with 83% in cluster 4. Although opting out of using the e-tools might have been for reasons of superior prior mathematics and statistics education, making practicing and testing in the e-tools an unnecessary exam preparation, this seems to be valid for no more than a small minority of clusters 1 & 2 students. In the majority of cases, opting out from using e-tools is a strong early signal for failing the course.

2 Antecedents and consequences of tool use

As clarified in the method section, the several aspects of student profiles are collinear. To be able to see the whole range of relevant relationships, we will focus on bivariate relationships first, expressed as simple correlations. To allow for easy comparison of the effects of different antecedents, we opt for presenting correlations graphically. With the current sample size, all correlations larger than .06 in absolute size are statistically significant at 5% level, and all larger than .09 in absolute size significant at 1% level.

Demographic and prior mastery variables are included in the left part of Figure 3. The three panels corresponding to the indicator variables demonstrate that Female students achieve higher mastery levels, both by spending more time, and working more efficient, than male students; that International students spend a lot more time, but due to limited efficiency, achieve mastery levels only slightly higher, than Dutch students; and that MathMajor students excel in efficiency for doing math. As one would expect, high prior mastery, expressed as score in the mathematics entry test or its three topic components, predicts high mastery in the tool, and high efficiency, against lower than average time investment.

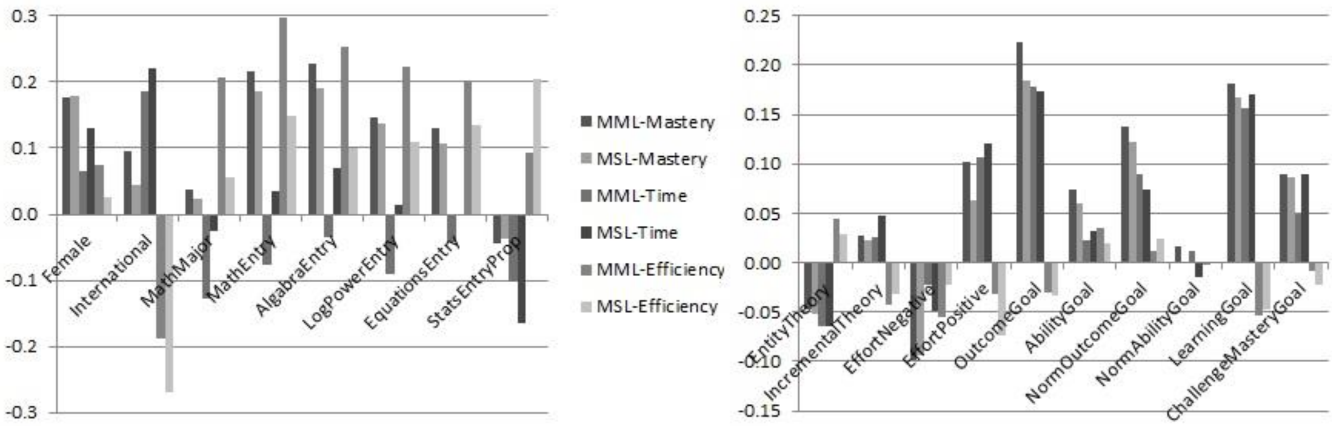


Figure 3. Correlations of tool use variables and demographic and prior mastery data (left), and implicit theories, effort beliefs and goals (right).

Both implicit theories, the Entity and Incremental Theory views, have non-significant relationships with the tool use variables. Effort beliefs have a slightly stronger impact: Effort Positive beliefs stimulate practicing in the e-tools, Effort Negative beliefs deter. But the strongest impact is by achievement goals: both the Outcome Goal, and the Learning Goal, stimulate the use of e-tools, whilst the Learning Goal does have a favourable impact on efficiency at the same time: see the right part of Figure 3.

Academic motivations have only limited impact on tool use: see the left part of Figure 4. Intrinsic Motivation to Know, and Identified Regulation, are beneficial for achieving higher mastery levels. Lack of any motivation, as in A-Motivation, has no impact on tool use time, but a negative on tool use efficiency, thus causing lower mastery levels.

The adaptive versus maladaptive break down of variables from the motivation and engagement wheel break is clearly visible from the right part of Figure 4. Adaptive behaviours and thoughts positively impact both time and mastery in the two tools. Maladaptive behaviours and thoughts have a detrimental effect, with however two exceptions. Students high in Uncertain Control tend to spend more time in the tools, but work inefficiently, the balance of the two being that their mastery is somewhat lower. Anxiety has about the same pattern, be it that anxiety has a very strong impact on tool use time, thus completely balancing lower levels of efficiency.

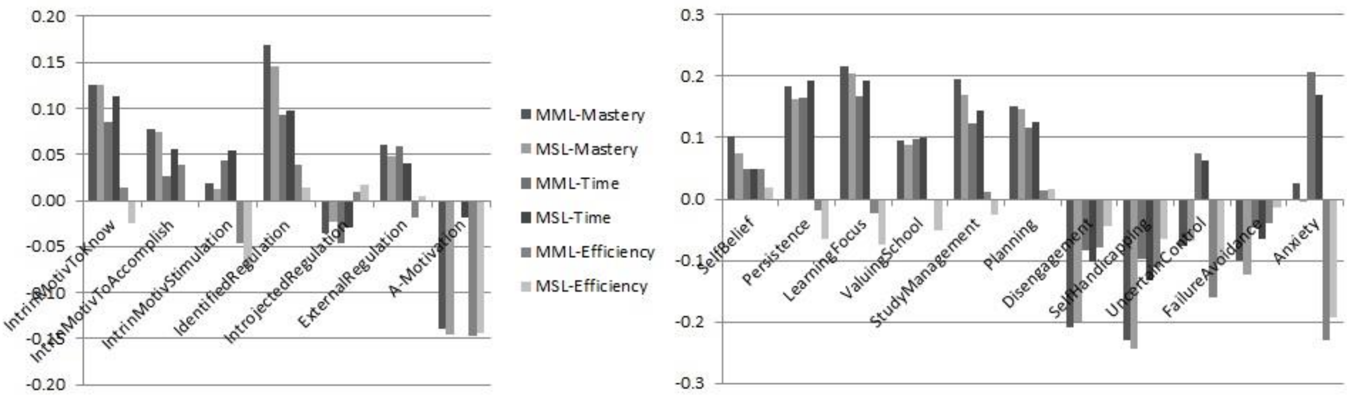
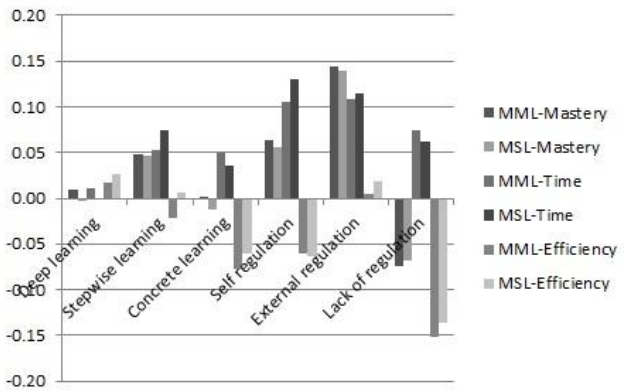


Figure 4. Correlations of tool use variables and academic motivations (left), and motivation and engagement (right).

Relationships between learning styles and tool use are more modest, especially for learning strategies: Stepwise Learners tend to be slightly more active, but correlations are small. Learning regulation is however a better predictor of tool use, with different patterns. Students who prefer Self-Regulation spend more time in the tools, but do so relatively inefficiently: their mastery does not profit one-to-one. The opposite case is for students with a preference for External Regulation: they gain most in mastery. Lack of regulation shows up first and for all in inefficient tool use: students high in Lack of Regulation spend more time in the tool, but achieve lower mastery, as a consequence of very low efficiency: see the left part of Figure 5.



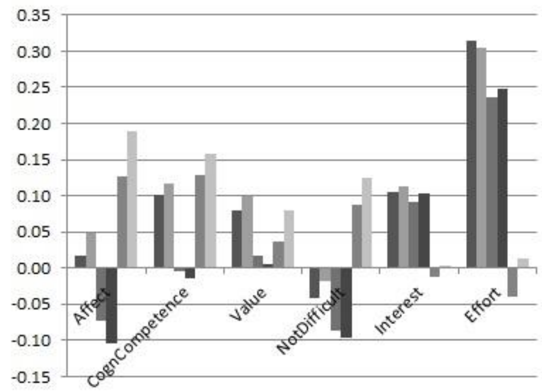


Figure 5. Correlations of tool use variables and learning styles (left) and subject attitudes (right).

Highest correlations for the determinants of tool use are to be found amongst the students attitudes towards learning mathematics and statistics. High self-efficacy, visible from the Cognitive Competence variable, results in high efficiency whilst investing same amounts of time, and thus higher mastery. The role of Affect is somewhat similar: less time, but highly efficient tool use. A very powerful but rather different pattern is found in planned Effort in learning mathematics and statistics: it positively impacts both mastery and time, but not efficiency: see the right part of Figure 5.

Most impressive correlations are not amongst tool use and its antecedents, but amongst tool use and performance as its consequence: Figure 6 provides these for six different performance measures. Math tool mastery and math time are the main predictors, with a very strong role in predicting the quiz performance, expressed as bonus score achieved in quizzes (MathBonus, StatsBonus). Tool efficiency scores positively impact course performance, and do so stronger for the final exam, than the quizzes.

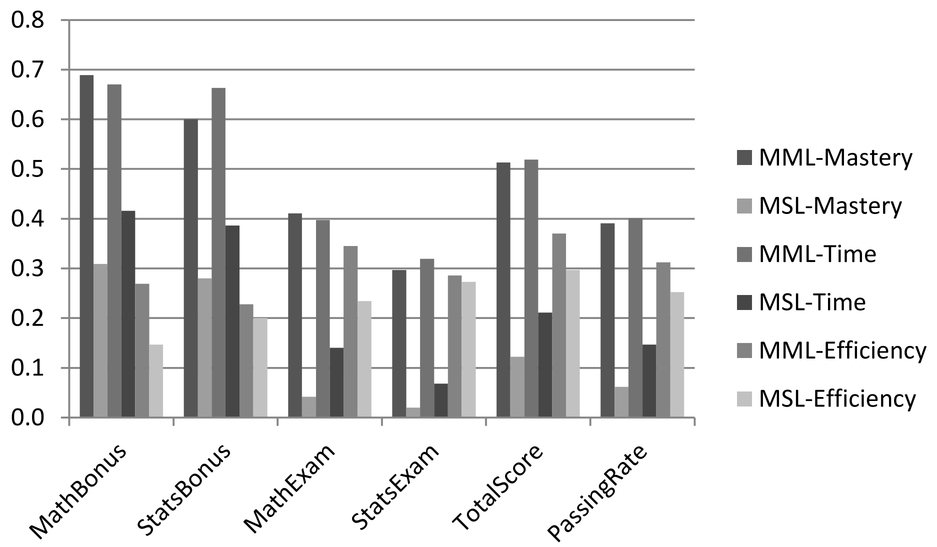


Figure 6. Correlations of tool use variables and course performances.

3 Predicting cluster membership by student profiles, early track data, or both

In the first part of this section reporting on the descriptive statistics, outcomes of a cluster analysis were provided that distinguishes four different patterns of tool use. The strong ties between tool use and course performance, as visible from Figure 6, suggest that the ability to predict cluster-membership is an important aim of any Learning Analytics application. In our context of having access to two different types of data, the question of the relationship between clustering based on tool use and student profiles can be phrased as: what dispositions and what system data best distinguish students in these four different clusters. Multinomial logistic regression is used to predict cluster membership of individual students; it generates linear prediction equations that best distinguish students from the base group (chosen as cluster 4) from each of the other groups. To report in detail on all regression equations for all relevant predictor sets is beyond the scope of this contribution; instead, we will focus on comparing the relative fit of the several multinomial logistic regression models predicting cluster membership when distinguishing five different sets of predictors that differ with regard to the time these predictor sets are available:

- Student learning dispositions (as discussed in 3.4.1 to 3.4.7), plus system generated data available at the very start of the course (3.3.1);
- Week1 track data (3.3.3), that is system generated data collected in the first week of the course, and available at the end of week 1;
- The combination of both of these data sets, that is the sum of all system generated and intentionally collected data available at the end of week 1 (including entry test data, 3.3.2);
- Week1 and week2 track data, that is system generated data collected in the first and second week of the course, and available at the end of week 2;
- The combination of all data available at the end of week 2, both system generated, and intentionally collected data.

Table 2 contains the outcomes of the several regression models. In terms of overall quality, the predictive power of dispositions and Week1 track data for predicting cluster membership are of the same order of magnitude. Combining dispositions and Week1 track data as explanatory variables, generates however a far better prediction model (whereas the addition of Week2 track data does not add much).

Table 2: Predicting cluster membership using dispositions, track data, or both

	Pseudo R^2	% predicted correct			
Explanatory variables	Nagelkerke	Cluster1	Cluster2	Cluster3	Cluster4
Dispositions	.37	35%	15%	8%	99%
Week1 track data	.43	85%	7%	2%	98%
Dispositions +Week1 track data	.68	100%	40%	25%	97%
Week1,2 track data	.49	83%	25%	6%	98%
Dispositions +Week1,2 track data	.68	100%	48%	16%	97%

Focussing on the predictor set consisting of dispositions and week1 track data, the regression equation that distinguish clusters 4 and 2, and the one that distinguishes clusters 4 and 3, are of most interest. Cluster 4 students distinguish themselves from cluster 2 students on the basis of, in order of beta weights: week1 MML-Mastery, week1 MSL-Mastery, academic motivations A-motivation and Introjected Regulation, the two attitudes Affect and Effort, Disengagement, the Logarithms & Powers component score in EntryTest, and the Normative Outcome Goal. In distinguishing cluster 4 from cluster 3 students, again week1 MML-Mastery and week1 MSL-Mastery scores are the most powerful predictors, supplemented with a different set of dispositions related predictors: attitude Value, Self-belief, Effort negative, and Failure Avoidance.

Summary and discussion

Learning theory, specifically social-cognitive theories of student learning (Dweck, 1999), suggest a wide range of learning related individual difference variables that may allow the generation of student profiles that predict learning behaviour, such as the intensity to use the technology enhanced components when learning in a blended learning environment (Lust et al., 2011). In our empirical study of 986 undergraduate business students, we have investigated the role of implicit theories of intelligence, effort beliefs, achievement goals, motivation and engagement, learning styles, academic motivation and subject attitudes as antecedents of tool use, hypothesizing that adaptive aspects of student profiles would stimulate tool use, whereas maladaptive aspects of profiles would hinder tool use. These broad hypotheses are generally corroborated in the empirical analysis using bivariate correlations and multinominal logistic regression. Adaptive goal setting behaviour, focussing on learning itself, or focusing on achieving outcomes in learning, contribute to tool use. As does the positive effort belief. Autonomous academic motivations, here intrinsic motivation to know and identified regulation, both contribute. In contrast, the maladaptive state of lack of motivation (A-Motivation) and the maladaptive negative effort belief, based on the view that it signals lack of intelligence, in combination with an entity theory view of intelligence, hinder tool use. Extending student profiles to include motivations and engagement and subject attitudes, repeats this pattern: all adaptive thoughts and behaviours, and the attitudes interest and planned effort, positively impact tool use, maladaptive behaviours have a negative impact on tool use.

However, there are two types of deviations from this general pattern in the profiles, which both find their explanation in the specific context. First, students with high prior mastery have no need to intensively use the e-tools in order to achieve a satisfactory course performance. This is visible from the correlational pattern of the prior education dummy variable (MathMajor), but also in the two attitude variables Affect and NotDifficult: less connect time, much higher efficiency, similar mastery.

The second deviating pattern refers to a sub classification of the maladaptive aspects of learning dispositions. It is most clearly demonstrated on the basis of the maladaptive thoughts and behaviours of the motivation and engagement wheel (Figure 4). The severely maladaptive aspects, the maladaptive behaviours, have a strong negative impact on all tool use variables, and especially the mastery variable. But the mildly maladaptive aspects, the maladaptive thoughts, demonstrate a mixed pattern, most clearly visible in the Anxiety variable: more time combined with lower efficiency, resulting in similar mastery. This extends to the learning styles (Figure 5): in learning style theories, deep learning and the ability to self-regulate one’s own learning, are regarded as the most adaptive profiles. However, correlations of these variables demonstrate a mixed pattern, or are absent: the most adaptive group of learners find sufficient support from the face-to-face component in the learning blend, and do not need (above average) use of the e-tool. Stepwise or surface leaning, and being dependent on others to regulate one's learning, external regulation, are examples of less adaptive learning styles. Given that the type of face-to-face learning relevant to this study, problem-based learning, is quite demanding with regard to the capabilities of students to self-direct their study, one can easily imagine that the support the e-tools offer in the learning process is most valuable for students who lack some of these skills: the less adaptive students, in need for external help to regulate their learning. As was previously reported by (Tempelaar, Niculescu, Rienties, Giesbers, Gijselaers, 2012), truly maladaptive students do not seem to profit from this digital support, as is clear from the correlational pattern of the Lack of Regulation variable, as well as Disengagement en Self-Handicapping.

The existence of such clear relationships between tool use and aspects of learning dispositions implies that these dispositions can serve the role as predictors for tool use in the course, and thus as predictors for the classification of students in the different clusters describing patterns of tool use. Under the assumption that these learning dispositions are available at the start of the course, we have investigated how well we can predict students' cluster-membership over time, with two different set of predictors: the history of available tracking data, as in many learning analytics applications, and learning dispositions, as in blended learning research. Week 1 tracking data does a satisfactory job in predicting clusters, as signalled by the explained variation. When looking at the percentages correct predicted, it is clear that it finds most of the explained variation in separating cluster 1 students, those who fail to become active e-learners, from the other three clusters. Explained variation by learning dispositions data is slightly less, but not very different (see also Table 2). It however does not require week 1 data, so can be available at the start of the course. Next, it finds its predictive power more equally distributed over the first three clusters (cluster 4 being by far the largest one, has high correct predictions in any method). The complementary ability to predict cluster membership suggests that combining the two predictor sets would boost the predictive power of the model. Which indeed is the case: dispositions and week1 track data together explain two thirds of the variation, provide about full prediction of the two extreme clusters, 1 and 4, and partial prediction of the two intermediate clusters, 2 and 3 (again, Table 2). In fact, predictive power related to these two most difficult to distinguish clusters raises stronger than in an linear, additive manner. Extending to the second week tracking data does not improve predictive power a lot: it is the combination of very early tracking data, and dispositions data, that is most helpful in identifying different types of learning behaviour.

Our study provides evidence to the conjecture in Ali et al. (2013) that currently many learning analytics applications are too superficial, based on track data only, and providing no more than simple statistics of such log data or interaction data. While track data based learning analytics tools were quite able to identify the inactive (Cluster 1) students, the two groups of students who are potentially at risk (Clusters 2-3) were not well detected by models based on early track data. We have demonstrated that supplementing track data with learning disposition data of the type that has been applied in recent studies into blended learning (Green et al., 2010; Lust et al., 2011, 2012), increases the relevance of learning analytics by improving the identification of Cluster 2-3 students.

Profiles relating social-cognitive learning theories, incorporating constructs as implicit theories of intelligence, effort beliefs, achievement goals, motivation and engagement, learning styles, subject attitudes and academic motivation, not only have predictive power in themselves, but appear to be strongly complementary to the other data component of track data, this way creating an ideal mix.

Practical Implications, limitations and further research

The use of the two MyLab learning tools contributed to both student performance and student experience; outcomes that are in line with McKenzie et al. (2013). Beyond providing a fully blended course, we experimented with what Mckenzie et al. (2013, p. 125) described as 'essential component underpinning both the implementation and evaluation of blended learning environments: learning analytics'. Feedback from practicing, formative assessments and quizzes in MML and MSL were reported to students and teachers, as well as reports from the several learning dispositions instruments. However, in the current experimental setup, this reporting functionality was limited. The first limitation was in the multitude of sources of feedback data: reporting in the current stage focussed on the separate aspects of learning related information, leaving it to students and teachers to create insights of using multiple feedback reports together. Second, the feedback process was mostly unobserved: beyond the observation of students consulting their quiz performance in the two ML systems, all other parts of student feedback were beyond measurement. As a consequence, the current study does not allow any investigation into to effectiveness of the learning analytics based feedback, similar to the investigation of the effectiveness of using the e-tools. In future research, we intend reorganize the feedback reporting of the learning analytics information as to make its use better observable, and by incorporating multiple feedback sources, this way allowing for such effectiveness studies.

In this study, we have explored the role of several instruments from a rich set of learning dispositions data, much richer than what is typically attainable in regular classes. Discovering the relative strength of this broad range of learning powers has been an important goal of this exploration. One main outcome is the strong collinearity that exists in these dispositions data, implying that a small subset will serve about the same role as the complete data set. In our current example, designing a learning analytics application for blended learning applying disposition data, the use of the motivation and engagement wheel, together with the learning styles instrument, would provide sufficiently rich information to distinguish adaptive and slightly and strongly maladaptive student profiles, in order to help students and teachers in learning regulation. It is only when counselling comes into the picture, that is helps knowing that types of maladaptive behaviour may originate from specific implicit theories or effort beliefs.

Acknowledgements

The current project has been financed by SURF-foundation as part of the Testing and Test Driven Learning program.

References

Ali, L., Asadi, M., Gašević, D., Jovanović, J. & Hatala, M. (2013). Factors influencing beliefs for adoption of a learning analytics tool: An empirical study. *Computers & Education*, 62, 130-148.

Blackwell, L. S. (2002). *Psychological mediators of student achievement during the transition to junior high school: The role of implicit theories*. Unpublished doctoral dissertation, Columbia University, New York.

Blackwell, L., Trzesniewski, K., & Dweck, C. S. (2007). Implicit theories of intelligence predict achievement across an adolescent transition: A longitudinal study and an intervention. *Child Development*, 78, 246-263.

Buckingham, S. S. & Deakin, C. R. (2012). Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics. In: S. Buckingham Shum, D. Gasevic, & R. Ferguson (Eds.), *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge*, pp. 92-101. New York: ACM Press.

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

Dweck, C. S., & Master, A. (2008). Self-theories motivate self-regulated learning. In D. H. Schunk, & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning: Theory, research, and applications* (pp. 31–51). New York, NY: Lawrence Erlbaum Associates.

Greene, J. A., Costa, L.-J., Robertson, J., Pan, Y., & Deekens, V. M. (2010). Exploring relations among college students' prior knowledge, implicit theories of intelligence, and self-regulated learning in a hypermedia environment. *Computers & Education*, 55, 1027-1043.

Hauptman, H. & Cohen, A. (2011). The synergetic effect of learning styles on the interaction between virtual environments and the enhancement of spatial thinking. *Computers & Education*, 57, 2106-2117.

Liang, Z., Liu, H., Zhang, C., & Yang, S. (2010). Simulation of Learners' Behaviors Based on the Modified Cellular Automata Model. *Intelligent Information Management*, 2, 563-568.

Lust, G., Juarez Collazo, N., Elen, J., Clarebout, G. (2012). Content Management Systems: Enriched learning opportunities for all? *Computers in Human Behavior*, *28*, 795-808.

Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., Clarebout, G. (2011). Tool-use in a blended undergraduate course: In Search of user profiles. *Computers and Education*, 57, 2135-2144.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology*, 77, 413-440.

McKenzie, W. A., Perini, E., Rohlf, V., Toukhsati, S., Conduit, R., & Sanson, G. (2013). A blended learning lecture delivery model for large and diverse undergraduate cohorts. *Computers & Education*, 64, 116-126.

Rienties, B., Giesbers, B., Tempelaar, D. T., Lygo-Baker, S., Segers, M., & Gijselaers, W. H. (2012). The role of scaffolding and motivation in CSCL. *Computers & Education*, 59, 893-906.

Romero-Zalvidar, V.-A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education*, 58, 1058-1067.

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: classic definitions and new directions. *Contemporary Educational Psychology*, 25, 54–67.

Schau, C.; Stevens, J., Dauphinee, T. L., & Del Vecchio, A. (1995). The development and validation of the Survey of Attitudes Toward Statistics. Educational and Psychological Measurement, 55, 868-875.

Tempelaar, D. T., Cuypers, H., Van de Vrie, E., Heck, A., Van der Kooij, H. (2013). Formative Assessment and Learning Analytics. In: D. Suthers & K. Verbert (Eds.), *Proceedings of the 3rd International Conference on Learning Analytics and Knowledge*, pp. 205-209. New York: ACM Press.

Tempelaar, D. T., Niculescu, A., Rienties, B., Giesbers, B., & Gijselaers, W. H. (2012). How achievement emotions impact students' decisions for online learning, and what precedes those emotions. *Internet and Higher Education*, 15, 161-169.

Vallerand, R. J., Pelletier, L. G., Blais, M. R., Brière, N. M., Senécal, C., & Vallières, E. F. (1992). The academic motivation scale: A measure of intrinsic, extrinsic, and amotivation in education. *Educational and Psychological Measurement*, 52, 1003–1017.

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-Driven Research to Support Learning and Knowledge Analytics. *Educational Technology & Society*, 15, 133–148.

Vermunt, J. D. & Vermetten, Y. (2004). Patterns in Student Learning: Relationships Between Learning Strategies, Conceptions of Learning, and Learning Orientations. *Educational Psychology Review*, 16, 359-385.

Wigfield, A., & Eccles, J. S. (2000). Expectancy -- Value Theory of Achievement Motivation. *Contemporary Educational Psychology*, 25, 68-81.

Wilkerson, L., & Gijselaers, W. H. (1996). *Bringing problem-based learning to higher education: Theory and practice*. San Francisco: Jossey-Bass.