Tempelaar, D., Rienties, B., & Nguyen, Q. (2021). The Contribution of Dispositional Learning Analytics to Precision Education. *Educational Technology & Society*, *24* (1), 109-122.

The Contribution of Dispositional Learning Analytics to Precision Education

Dirk Tempelaar[1*], Bart Rienties[2] and Quan Nguyen[3]
[1]Maastricht University, School of Business and Economics, The Netherlands // [2]The Open University, Institute of Educational Technology, UK // [3]School of Information, University of Michigan, Ann Arbor, MI, USA //
d.tempelaar@maastrichtuniversity.nl // bart.rienties@open.ac.uk // quanngu@umich.edu
[*]Corresponding author

ABSTRACT: Precision education requires two equally important conditions: accurate predictions of academic performance based on early observations of the learning process and the availability of relevant educational intervention options. The field of learning analytics (LA) has made important contributions to the realisation of the first condition, especially in the context of blended learning and online learning. Prediction models that use data from institutional information systems and logs of learning management systems have gained a good reputation in predicting underperformance and dropout risk. However, less progress is made in resolving the second condition: applying LA generated feedback to design educational interventions. In our contribution, we make a plea for applying dispositional learning analytics (DLA) to make LA precise and actionable. DLA combines learning data, as in LA, with learners' disposition data measured through self-report surveys. The advantage of DLA is twofold: first, it improves the accuracy of prediction, specifically early in the module, when limited LMS trace data are available. Second, the main benefit of DLA is in the design of effective interventions: interventions that focus on addressing individual learning dispositions that are less developed but important for being successful in the module. We provide an empirical analysis of DLA in an introductory mathematics module, demonstrating the important role that a broad range of learning dispositions can play in realising precision education.

Keywords: Blended learning, Dispositional learning analytics, Educational intervention, Flipped learning, Precision education

1. Introduction

The "engineers" approach to education' seems to be a good way to define precision education casually. A more formal definition, derived from Yang (2019), states that precision education aims to identify at-risk students as early as possible and provide timely intervention through diagnosis, prediction, treatment, and prevention. These four components imply that the diagnosis of students' learning patterns, the prediction of students' learning outcomes, the treatment with learning strategy and activities, and prevention shape the four main research topics in precision education. The metaphor of engineering science does bring some more insights, whereby both system theory and control theory are important building blocks. System theory is engaged with the description of systems by mathematical models that enable the estimation of the state of the system and the prediction of the outcome of the system resulting from that state (Padulo & Arbib, 1974). Control theory is based on system theory and adds the dimension of steering the state of the system towards a predefined, desired constellation (Padulo & Arbib, 1974).

When comparing educational theory with system theory, precision education stands for the added dimension of being in control with regard to the educational attainments. To be "in control" in the system theory context, two so-called structural properties are required: observability and controllability (Padulo & Arbib, 1974). Controllability, also called reachability, refers to the interaction of system inputs and the system state. A system is controllable if, with proper input, every possible state can be reached. Observability relates to the interaction of state and output. A system is observable if the measurements of both inputs and outputs contain sufficient information to reconstruct any state (reconstructability is the alternative term for observability). Looking at the definition of precision education provided by Yang (2019), one can identify both aspects. Diagnosis and prediction refer to the observability aspect, whereas treatment and prevention refer to the controllability aspect.

Although precision education is a young discipline, different interpretations can be found as to what its focus is. Often, that difference is about a focus on both aspects of controllability and observability versus a focus on observability only. Next, most empirical research into precision education concentrates on the ability to predict students' academic performance and to identify students at risk for dropout. For example, Hart (2016, p. 209) suggested a focus on "creating data specifically for gaining a better understanding of the classification of learning disabilities at the individual level." Lian and Sangarun (2017), in the context of language education, defined the "precision project" in terms of providing accurate, detailed, timely, adaptive, and contextualised personalised data to facilitate intervention. Three empirical studies, Lu et al. (2018), Hsiao et al. (2019), and Huang et al. (2020), although acknowledging the role of intervention in precision education by phrasing that "*learning analytics is a conceptual framework and as a part of our Precision education used to analyze and predict students' performance and provide timely interventions based on student learning profiles*" (Lu et al., 2018, p. 221), focus on the role of learning analytics in providing accurate predictions of academic performance.

We make a plea for equal focus on prediction and intervention, observability and controllability in this contribution. We will materialise that plea by adopting a dispositional learning analytics (DLA) framework. It is our conjecture that adding dispositions of learners to learning analytics is a small step to take that may bring considerable benefits in terms of increasing observability, creating a better picture of the features of a learning issue, as well as increasing controllability, transforming predictions of poor performance or risk of dropout into meaningful interventions.

2. Learning analytics and dispositional learning analytics

Learning analytics (LA) refers to "the analysis and interpretation of educational data, such as the logs recorded in learning management systems, the interactive contents recorded in online discussion forums, or the learning process captured on video, to provide constructive feedback to learners, instructors or educational policymakers" (Hwang et al., 2018, p. 134). At the early stage of learning analytics, many scholars have focused on building predictive models based on data extracted from both institutional student information systems (SIS) and digital platforms that organise and facilitate learning, such as learning management systems and e-tutorials (LMS, taken them together). While these studies provide important markers on the potential of LA in education (Viberg et al., 2018), the findings were rather limited to the descriptive function of LA, which is mostly based on the "measurement, collection, analysis and reporting of data about learners and their contexts" (Siemens & Gašević, 2012, p. 1). Given the rigidity of SIS and LMS data, educators may encounter difficulties in designing pedagogically informed interventions. To overcome this shortcoming, Buckingham Shum and Crick (2012) proposed the Dispositional LA (DLA) infrastructure that combines learning data (i.e., generated in learning activities through traces of an LMS) with learner data (e.g., student dispositions, values, and attitudes measured through self-report surveys). Typical stakeholders of DLA applications are both students and teachers/tutors: these applications can be positioned at the meso- and micro-level (Ifenthaler, 2015).

It is our conjecture that, especially in precision education, these dispositions of learners are a crucial add-on to common LA functions. They are the dispositions that make the feedback both precise and actionable (Gašević et al., 2015; Tempelaar et al., 2017a). The term "actionable feedback," introduced into the LA community by Gašević et al. (2015), is crucial in pinpointing the aspect of precision education, not by default present in LA applications. Although the concept is well accepted in the LA community, not much empirical research has been published focusing on the extent to which LA feedback is actionable. Some exceptions are, for instance, Gašević et al. (2017), in which students' learning strategies are derived from log data, or Abdous et al. (2012), where learning behaviours in terms of frequency of chat messages and platform questions are analysed. Indeed, in a large scale application of LA of 1159 teachers in 231 courses containing 23K students over four years, Herodotou et al. (2020) found that although teachers appreciate rich predictive learning analytics data of whom might be at risk of dropping out, several groups of teachers struggled to put this advice into action.

Adding dispositions opens the perspective of enriching the LA learning feedback. Rather than intervening with the simple message to Maria "please catch up, you are lagging behind," the intervention can now contain some actionable aspects as "We see that your performance is not as good as most of your peers and, at the same time, we observe that you view the videos only after the class session has taken place, rather than before. It might be better to try viewing these videos before meeting in class." Still, this feedback is unidimensional in addressing one singular aspect of learning behaviour, and if no deviant behaviour can be detected in the learning strategies of the student lagging behind, we are still empty-handed with regard to actionable feedback. Therefore, the aim of DLA is to introduce a multidimensional perspective of learning dispositions into the practice of LA by identifying which constellations of personality characteristics go together with the risk of dropout or the prediction of underperformance. In Buckingham Shum and Deakin Crick (2012) and Buckingham Shum and Ferguson (2012) the source of learner data is found in the use of a dedicated survey instrument specifically developed to identify learning power (Deakin Crick et al., 2015): the mix of dispositions, experiences, social relations, values, and attitudes that influence the engagement with learning.

Sharing a similar systems' approach as described in Deakin Crick et al. (2015), we propose to operationalize dispositions with the help of instruments developed in the context of contemporary educational research, as to make the connection with educational theory and pedagogical interventions as strong as possible. Somewhat preluding on the outcomes of the empirical part of our contribution, the following illustrates the role of a broad spectrum of learning dispositions, which may make learning feedback actionable. Continuing with the example described above, when a student lags, a traditional LA application cannot go any further than providing a warning: "please catch up!"

If we extend the LA application as in Gašević et al. (2017) by deriving learning strategies, we can extend the feedback with reference to suboptimal strategies applied. However, this feedback is also at the symptom level and may trigger symptom management rather than looking for potential causes of these symptoms. In other words, the learning feedback is not that actionable as to suggest concrete forms of intervention, and the controllability condition is not yet fulfilled. In previous research of the authors (Nguyen et al., 2016; Rienties et al., 2019; Tempelaar et al., 2015; Tempelaar et al., 2017a, Tempelaar et al., 2017b; Tempelaar et al., 2018a, Tempelaar et al., 2018b; Tempelaar et al., 2019) and again in the empirical part of this contribution, we found that different constellations of learning dispositions might be associated with underperformance and lagging behind in the learning process. One such constellation may be based on personality characteristics of motivational types, such as disengagement. A very different constellation can be based on learning regulation strategies: the balance between self-regulation and external regulation of learning.

While typically education strives to achieve high levels of self-regulated learning, our empirical studies indicated that a minimum of external regulation, open up to the teachers' advice and the structure of the curriculum to a certain level, is indispensable. Both types of students, the disengaged and the too strong self-regulated, demonstrate the same symptoms: not being very active in doing the scheduled learning activities, lagging in what is measured (Tempelaar et al., 2015). However, the independent learner who makes his own way (indeed, more often his than her) is certainly not disengaged and requires feedback and intervention of a very different type than the disengaged learner. It is here that the dispositions component of DLA can be the crux to precision education, as we intend to showcase in the empirical part of our contribution. Rather than formulating specific research hypotheses, we will design our empirical study around this open research question: how can DLA contribute to prediction and intervention, to both dimensions of observability and controllability? Given the strong focus of most LA applications on observability issues, on the derivation of prediction models, we will emphasise the component of controllability in our study: in what respect can DLA better than LA facilitate the design of learning interventions directed at taking observed learning obstacles.

3. The educational context of the empirical study: blended learning and flipped classes

The learning context investigated in previous research by the authors as well as in the empirical part of this contribution is best described as large-scale introductory mathematics and statistics module using "blended" or "hybrid" learning in a business and economics university program in the Netherlands. Blended learning "combines online digital resources with traditional classroom activities and enables students to attain higher learning performance through well-defined interactive strategies involving online and traditional learning activities" (Lu et al., 2018, p. 220). The main learning component in our blend is face-to-face: Problem-Based Learning (PBL), in small groups (14 students), coached by a content expert tutor (Non & Tempelaar, 2016; Williams et al., 2016). Participation in these tutorial group sessions is required. Optional is the online component of the blend: the use of two e-tutorials or online learning and practising environments: the e-tutorial SOWISO to learn and practice mathematics (https://sowiso.nl/en/) and the e-tutorial MyStatLab to learn and practice statistics (https://www.pearsonmylabandmastering.com/northamerica/mystatlab/) (see Tempelaar et al., 2015; Tempelaar et al., 2017a, Tempelaar et al., 2017b; Tempelaar et al., 2018a, Tempelaar et al., 2018b; Tempelaar et al., 2019). This choice is based on the philosophy of student-centred education, placing the responsibility for making educational choices primarily on the student. Since most of the learning takes place in self-study outside class using the e-tutorials or other learning materials and class time is used to discuss solving advanced problems, the instructional format is best characterised as a flipped-class design (Hsiao et al., 2019; Lin & Hwang, 2018; Williams et al., 2016). The use of e-tutorials and achieving good scores in the practising modes of both e-tutorials is stimulated by making bonus points available for good performance in the quizzes. Quizzes are taken every two weeks and consist of items that are drawn from the same item pools applied in the practising mode. We chose this particular constellation as it stimulates students with limited prior knowledge to make intensive use of the e-tutorials. The bonus is maximised to 20% of what one can score in the exam.

The student-centred nature of the instructional design requires, first and foremost, adequate actionable feedback to students so that they can monitor their study progress and topic mastery. The provision of relevant feedback starts on the first day of the module when students take two diagnostic entry tests for mathematics and statistics. Feedback from these entry tests provides a first signal of the importance of using the e-tutorials. Next, the e-tutorials SOWISO and MyStatLab take over the monitoring function: at any time, students can see their performance in the practice sessions, their progress in preparing for the next quiz, and detailed feedback on their completed quizzes, all in the absolute and relative (to their peers) sense.

Our program is characterised by a large diversity in the student population: only about 20% of the students are educated in the Dutch high school system, most students are international, with a large share of European nationalities: no more than 5% of students are from outside Europe. High school systems in Europe differ strongly, most notably in the teaching of mathematics and statistics. Therefore, it is crucial that the first module is flexible and allows individual learning paths (Non & Tempelaar, 2016; Williams et al., 2016).

Learning dispositions measured at the start of the course were of affective, behavioural, and cognitive types (Rienties et al., 2019). The surveys had a prime role in supplying students with an individual data set required for doing a statistical project, resulting in a full response.

4. Methods for the current study

4.1. Participants

In the empirical part of this contribution, we investigated DLA's potentials applied to two cohorts of first-year students in our program: the cohorts of academic years 18/19 and 19/20. Profiling of students took place on the basis of a quartile split of the scores of all students who participated in the first quiz: 2,261 students (since the collection of survey data is part of a mandatory assignment, full data are available for all these students). Of these students, 40% were female, 60% male (with *Female* as indicator variable), and 35% followed the advanced mathematics track in high school education (with *MathMajor* as indicator variable).

4.2. E-tutorial trace data

Although students learned in two different e-tutorial systems, SOWISO for mathematics and MyStatLab for statistics, we focus the analysis on data gathered from the SOWISO platform since that platform allows us to trace every individual learning activity as a time-stamped record in the database of loggings. Students spend an average of 27 hours in SOWISO or 3.5 hours per week, out of the ten hours per week available to learn mathematics (that includes both self-study and class time). Learning can be divided into three consecutive learning phases. The first learning phase prepares the weekly tutorial session: students are expected to enter these sessions well prepared by self-studying the weeks' topic in advance so that the session itself can be used to solve advanced problems. The second learning phase starts after the tutorial session took place and runs until the bi-weekly quiz session. In phase two, students prepare the quizzes that bring them a bonus score. The third and last learning phase is the examinations' preparation: it starts after writing the quiz and continues until the writing of the exam begins. In each learning phase and for each of the seven weekly topics covered in the program, the students' mastery in the practice mode of the e-tutorial is measured: the proportion of problems successfully solved. Three learning phases and seven weekly topics imply a total of 20 mastery measures: *TGMasteryTopicWk1*, *QzMasteryTopicWk1*, *ExMasteryTopicWk1*, to *ExMasteryTopicWk7*, where *TG*, *Qz*, and *Ex* refer to tutorial group session (first learning phase), quiz session (second learning phase), and examination (third learning phase) (the last weekly topic is not included in any quiz).

4.3. Performance data

Following the focus on mathematics learning, as explained above, two different types of performance indicators are available: exam score and quiz scores. There are three bi-weekly quizzes, covering the topics of the first two weeks (*MathQuiz1*), the second two weeks (*MathQuiz2*), and the third two weeks (*MathQuiz3*) (the last week topic not covered in any quiz). *MathExam* represents the score for the mathematics component in the final examination. Since this study aims to investigate the potential of early prediction of performance, most of our analysis is focused on the role of the first quiz that is administered in the third week of the module in predicting performance, and the relationships between learning dispositions and the score achieved in that first quiz:

MathQuiz1. The module starts with a diagnostic entry test, producing *EntryTest* as a measure of prior knowledge.

4.4. Disposition data

Five different instruments were applied to operationalize learning dispositions, all documented in full detail in previous studies (Tempelaar et al., 2015, Tempelaar et al., 2017a, Tempelaar et al., 2017b; Tempelaar et al., 2018a, Tempelaar et al., 2018b; Tempelaar et al., 2019). For space limitations, we limit the current description to the identification of survey scales adopted and refer to the above sources for a full elaboration.

Individual approaches to cognitive learning processing strategies and metacognitive learning regulation strategies were based on Vermunt's (1996) learning styles instrument. Processing strategies can be ordered from surface to deep learning approaches: *Memorising* and rehearsing and *Analysing* are two scales that represent different aspects of surface learning. *Relating* and structuring and *Critical* processing are two scales that represent different aspects of deep learning. *Concrete* processing is separate from this continuum and represents the tendency to learn strategically. Regulation strategies are decomposed into self and external regulation: Self-regulation of learning processes and results (*SelfRegProc*), Self-regulation of learning content (*SelfRegCont*), External regulation of learning processes (*ExtRegProc*), and External regulation of learning results (*ExtRegRes*), with Lack of regulation (*LackReg*) indicating a lack of regulation of any type.

Attitudes and beliefs toward learning quantitative topics were assessed with the SATS instrument (Tempelaar et al., 2007). It distinguishes Affect, cognitive competence (*CognComp*), *Value*, expected difficulty in learning, reversed (*NoDifficulty*), *Interest*, and planned *Effort*.

Learning emotions, both epistemic and activity type were measured on the basis of instruments developed by Pekrun (Pekrun & Linnenbrink-Garcia, 2012). Epistemic emotions are composed of positive emotions, *Curious* and *Excited*, negative emotions *Confused*, *Anxious*, *Frustrated*, and *Bored*, and the neutral emotion *Surprised* (Pekrun et al., 2017). Activity emotions were measured the module halfway, in the fourth week. Although these emotions are strongly associated with performance measures, we left them out of the analysis, given the focus of early prediction.

The instrument Motivation and Engagement Wheel (Martin, 2007) breaks down learning cognitions and learning behaviours into four categories of adaptive versus maladaptive types and cognitive versus behavioural types. *Self-belief*, value of school (*ValueSchool*), and learning focus (*LearnFocus*) shape the adaptive, cognitive factors, or cognitive boosters. Planning, task management (*TaskManagm*), and *Persistence* shape the behavioural boosters. Mufflers, the maladaptive, cognitive factors are *Anxiety*, failure avoidance (*FailureAvoid*), and uncertain control (*UncertainControl*), while self-sabotage (*SelfSabotage*) and *Disengagement* are the maladaptive, behavioural factors or guzzlers.

A recently developed 4x2 achievement goal framework by Elliott and coauthors (Elliot et al., 2015) was applied to include the self-perceived goal-setting behaviour of students. The instrument distinguishes two valence dimensions: approach and avoid, and four goal definition dimensions: task-based competence (striving to do the task correctly), self-based competence (do better than before), other-based competence (do better than others) and potential-based (to do the best one can) competence, resulting in eight scales: *TAP* or Task-Approach, *TAV* or Task-Avoid, *SAP* or Self-Approach, *SAV* or Self-Avoid, *OAP* or Other-Approach, *OAV* or Other-Avoid, *PAP* or Potential-Approach, and *PAV* or Potential-Avoid achievement goals.

4.5. Analyses

In this study, we choose to profile students in a very simple way: by quartile split of the first quiz score: *MathQuiz1*. In previous research (Tempelaar et al., 2015, Tempelaar et al., 2017a, Tempelaar et al., 2017b; Tempelaar et al., 2018a, Tempelaar et al., 2018b; Tempelaar et al., 2019), more advanced statistical techniques as cluster analysis or latent class analysis served the role of distinguishing students' different learning profiles. However, to distinguish student profiles with different dispositional characteristics, we do not need advanced statistical methods, which will be illustrated by our simple quartile split. In this study, the prime aim of predictive modelling is directed at predicting academic performance and distinguishing different groups or profiles that are best helped with different types of learning feedback based on different dispositional profiles (such as raising confidence for students being failure avoidant or helping to organise the study in students who lack planning skills). Since the first quiz score is the dominant early predictor of final course performance (see next section), a simple quartile split into four different profiles is adequate. Profile means are compared with ANOVA analyses. Differences in means are classified as statistically significant when p-values are below .01.

5. Experimental results and discussion

5.1. Prediction equations of final performance

The first step in the analysis is to consider the predictive power of alternative sets of predictors to explain the final exams' performance. Given that our investigations are directed at finding prediction models that allow for timely interventions, we restricted predictors to constructs measured before the start of the module or in the first three weeks of the module (allowing ample time for intervention in the later five weeks of the module). The alternative sets of predictor variables are:

- Demographics measured before the start of the module: gender, prior education, score in diagnostic entry test and dummy for year.
- Demographics plus the learning dispositions discussed above in the subsection of disposition data, all measured before the start or at the start of the module.
- Demographics plus learning dispositions plus trace variables collected in the first three weeks of the module: learning mastery in the e-tutorial for each of the topics covered in the first three weeks of the program, plus the score in the first quiz taking place in the third week of the module.
- Space prevents us to report the full regression models (mainly due to the 34 dispositional variables in the prediction model). Therefore, Table 1 restricts reporting the predictive power of the several models.

Table 1. Predictive power of three prediction models of final performance

Predictor set	R^2
demographics	0.184
demographics + dispositions	0.305
demographics + dispositions + traces	0.411

Table 1 emphasises one of the roles of dispositional variables (beyond their role in the actionable feedback): they empower the prediction equations raising predictive power of diagnostic entry test and prior education (gender is insignificant as predictor) with 12%. Trace variables add another 10.5%. Within the set of trace variables, the score in the first Quiz is the dominant predictor. In itself, MathQuiz1 explains 28.4% of variation in MathExam, leaving no more than a modest 12.7% of additional explained variation for all demographic, dispositional and trace variables. This outcome is fully in line with findings in previous studies (Tempelaar et al., 2015, Tempelaar et al., 2017a, Tempelaar et al., 2017b; Tempelaar et al., 2018a, Tempelaar et al., 2018b; Tempelaar et al., 2019): it is only in the early start of the module that demographic variables, dispositional variables or learning activity variables measured by traces in learning environments contribute substantially to the prediction of learning performance. Very early learning interventions, taking place before the first measures of cognitive types, such as quiz scores, are collected, are best based on the full set of predictors. However, as soon as these cognitive measures are included, they dominate all other types of predictors in the prediction equations.

5.2. Student profiles by quartile split

Four student profiles are determined as the outcome of a quartile split of *MathQuiz1* score. That score expressed as a proportion, the three quartiles are .458, .667 and .804. The four quarters have unequal size, due to the stepwise nature of the *MathQuiz1* scores. Q1 counts 473 students scoring below 0.458; Q2 counts 648 students scoring between .458 and .667; Q3 counts 573 students scoring between .667 and .804, and Q4 counts 567 students scoring .804 or beyond. Average scores of all four quarters for the three quiz scores and the final exam score, all expressed as proportions are depicted in Figure 1.

All profile differences are strongly statistically significant and substantial in size (ANOVA F-values > 200, p-values < 10^{-10}). Since the profiling is based on *MathQuiz1* scores, differences are largest for that variable, but also substantial in the other scores. Given that the passing norm is 55%, profile differences in passing rates are even higher. With most students in Q1, the first quarter, failing, and nearly all students in the third and fourth quarters, Q3 and Q4, passing.

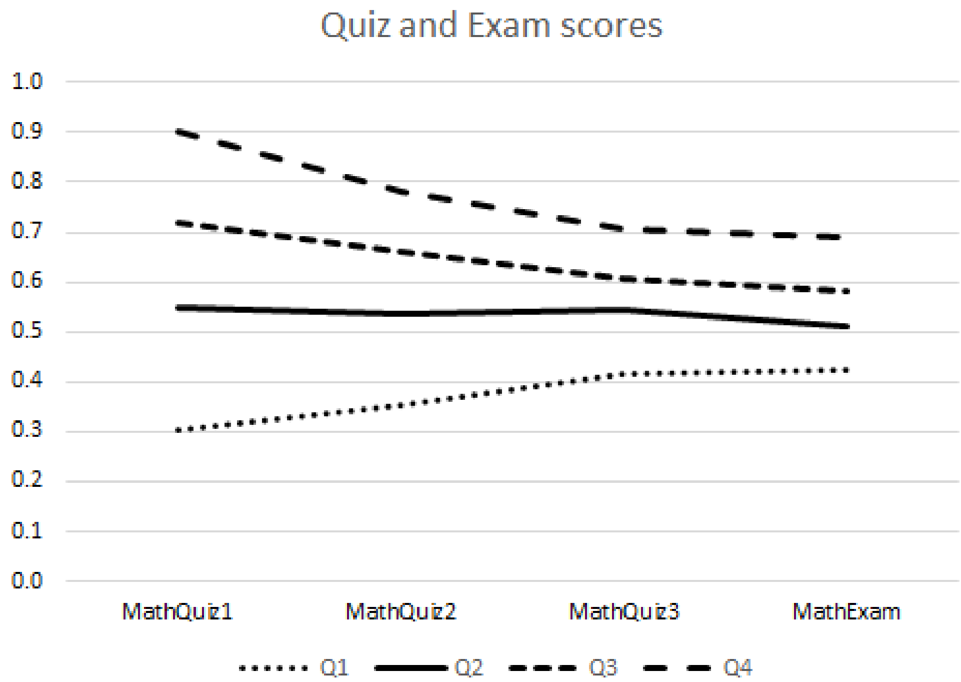

Figure 1. Profiling students by quartile split: Quiz and exam performance

5.3. Student profiles by quartile split: Traces of learning activity

In a traditional LA application, the first thing to check is if the profiling is validated by differences in levels of learning activity. Our application focuses on the mastery level defined as the proportion of successfully solved practice problems achieved in the e-tutorial in the first three weeks of the module. Every week has its own topic to be covered, so mastery levels refer to three different topics: see the three panels of Figure 2. In every topic, we distinguish three subsequent learning phases: mastery achieved in the preparation of the tutorial session (*TGMastery*), master achieved in the preparation of the quiz session (*QzMastery*), and mastery achieved in the preparation of the examination (*ExMastery*). Mastery levels are cumulative: mastery in subsequent phases add to the mastery achieved in previous ones.

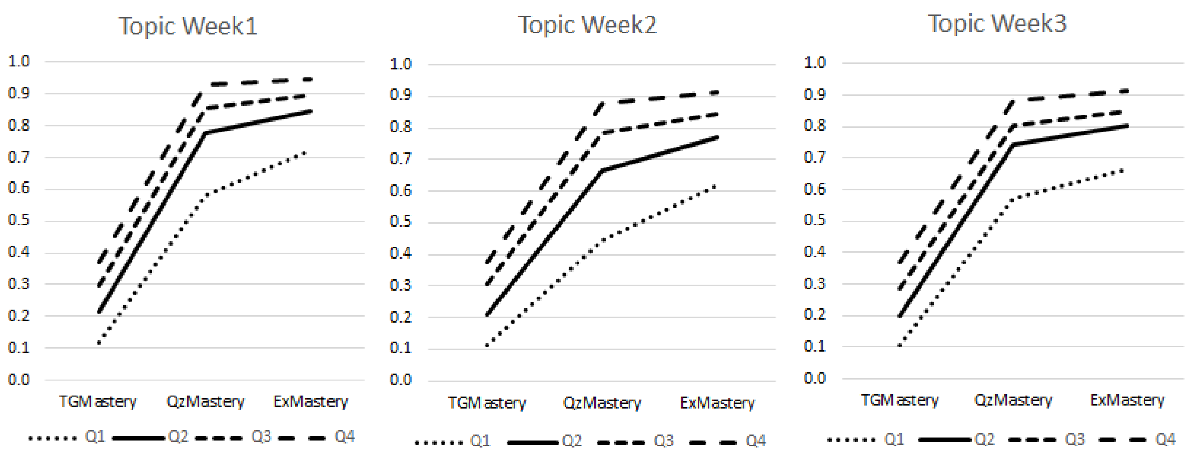

Figure 2. Profiling students by quartile split: Mastery by learning activity in the first three weeks

We observe considerable differences between profiles, all strongly statistically significant (ANOVA F-values > 50, p-values < 10^{-10}). Figure 2 suggests that most of the learning takes place in the second phase: after the tutorial session took place, in preparation for the quiz.

5.4. Student profiles by quartile split: Demographics

Prior education plays a major role in the explanation of the composition of the four different profiles. The most considerable differences are in the *MathMajor* variable: 58% of students in Q4 are educated in the science preparing track in high school, against no more than 15% in Q1 (ANOVA $F = 81.5$, p-value $< 10^{-10}$). These differences in prior education show up in the *EntryTest*: scores students achieve in the diagnostic entry test taken before the module starts (ANOVA $F = 133.7$, p-value $< 10^{-10}$). There are no gender differences between the four quarters: see Figure 3 (ANOVA $F = 1.7$, p-value = .162).

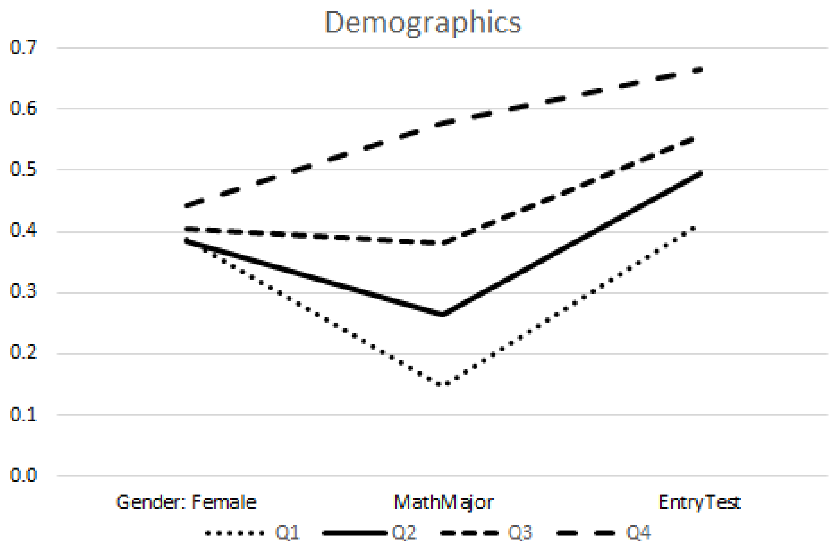

Figure 3. Profiling students by quartile split: Gender, prior education, and entry test score

5.5. Student profiles by quartile split: Learning processing and learning regulation strategies

Profile differences of learning processing and regulation strategies are at a much lower level, as Figure 4 illustrates. Statistical significant differences exist for *Relating* and structuring (ANOVA $F = 4.0$, p-value = .008), one of the deep learning components, for *ExtRegRes*, the external regulation of learning results (ANOVA $F = 3.9$, p-value = .009), and for Lack of regulation (ANOVA $F = 10.6$, p-value $< 10^{-7}$). Concerning learning processing: the higher the profile, the higher the level of deep learning, the lower the level of concrete learning. Concerning learning regulation: the higher the profile, the higher the level of external regulation of learning results, the lower the level of a lack of regulation.

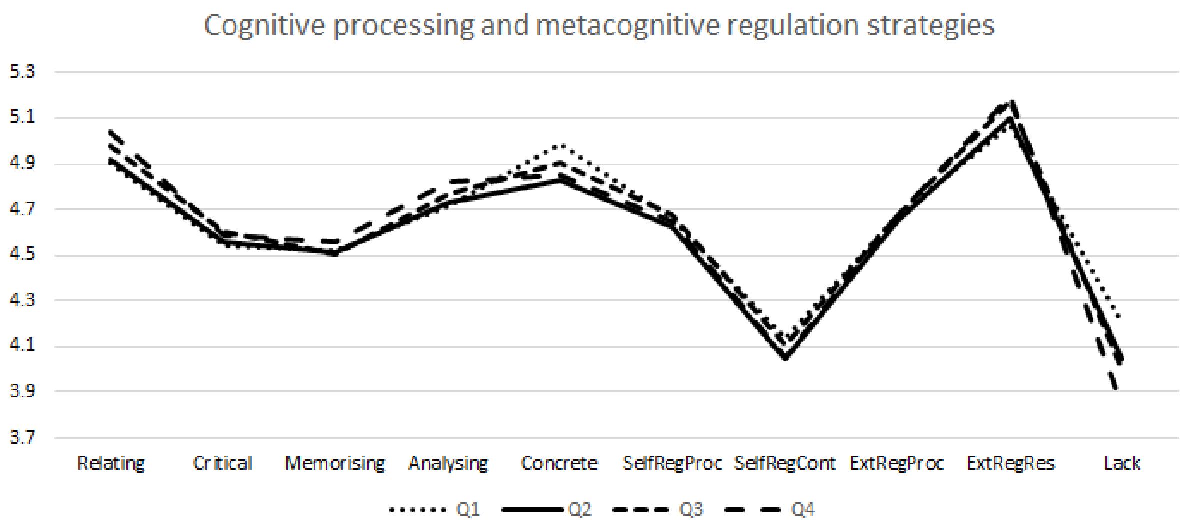

Figure 4. Profiling students by quartile split: learning processing and regulation

5.6. Student profiles by quartile split: Learning attitudes and beliefs

More considerable profile differences pop up when analysing attitudes and beliefs towards learning mathematics. All facets, except *Effort*, demonstrate statistically significant differences (ANOVA $F > 8.0$, p-values $< 10^{-4}$). However, the size of the differences is especially large for the *Affect* and *CognComp* variables, with levels in these attitudes increasing with higher quartiles. See Figure 5.

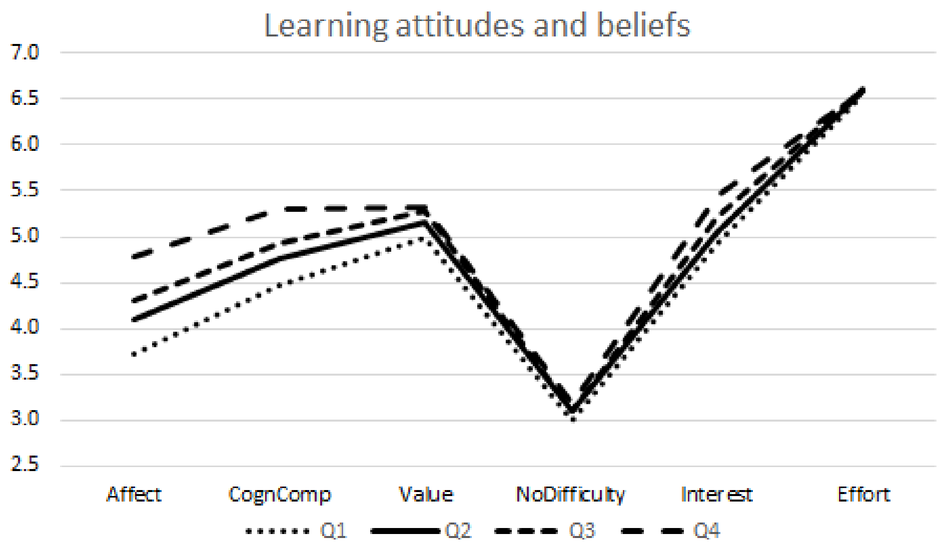

Figure 5. Profiling students by quartile split: Learning attitudes and beliefs

5.7. Student profiles by quartile split: Epistemic learning emotions

The valence dimension of epistemic emotion splits the graph of profile means into two mirrored patterns. Positive emotions *Curious* and *Excited* are positively related to the ordering of quarters; *Surprise*, hypothesised as a neutral emotion, is the only epistemic emotion without differences in profile means; all other differences are statistically significant (ANOVA $F > 10.4$, p-values $< 10^{-7}$). Negative epistemic emotions *Confused*, *Anxious*, *Frustrated*, and *Bored* demonstrate levels that are inversely related to the quarters' order with substantial profile differences: see Figure 6.

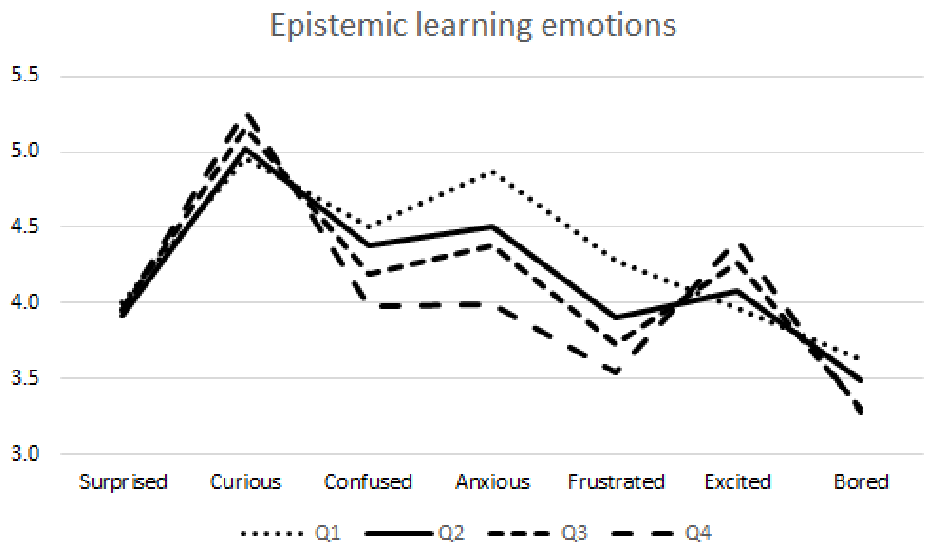

Figure 6. Profiling students by quartile split: Epistemic learning emotions

5.8. Student profiles by quartile split: adaptive motivation and engagement

Cognitive and behavioural motivation and engagement constructs of adaptive type demonstrate different patterns. The cognitive scale *Self-belief* demonstrates a small but significant profile difference (ANOVA $F = 5.8$, p-value = .0006); *LearnFocus* (ANOVA $F = 3.5$, p-value=.015) and *Value School* do not (ANOVA $F = .50$, p-values = .684). The behavioural constructs *Planning* and *Persistence* demonstrate profile differences of more substantial size, in the direction that higher ordered quarters reach higher mean levels (ANOVA $F = 6.0$ and 15.1, p-values = .0005 and 10^{-9}). No mean differences exist for *TaskManagm* (ANOVA $F = .35$, p-value = .791): see Figure 7.

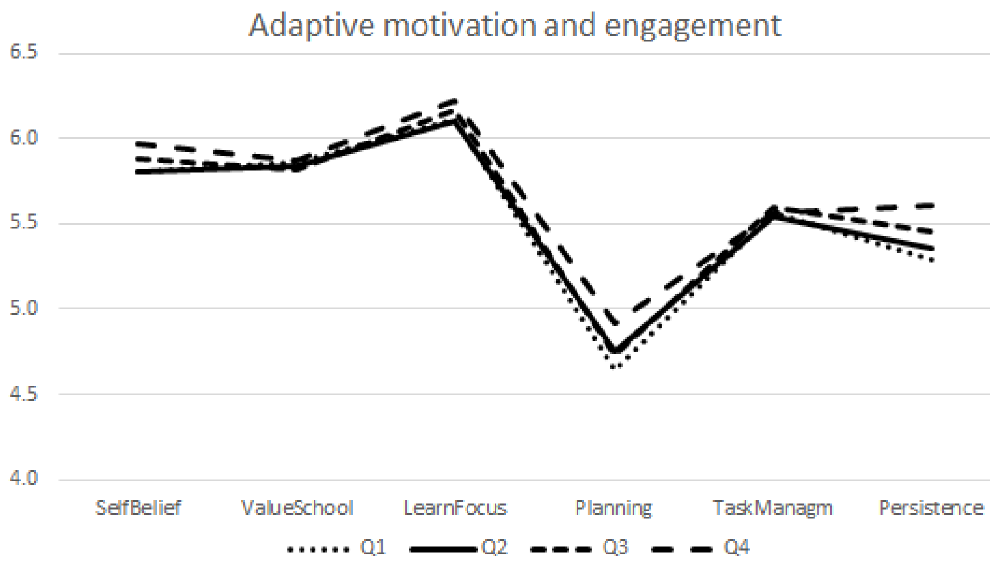

Figure 7. Profiling students by quartile split: Adaptive motivation and engagement

5.9. Student profiles by quartile split: maladaptive motivation and engagement

The same breakdown of cognitive and behavioural mean profile levels is visible in the next figure, Figure 8, providing patterns for maladaptive motivation and engagement constructs. The differences in profile means of the maladaptive cognitive constructs, *Anxiety* and *UncertainControl* (ANOVA $F = 10.6$ and 12.3, p-values < 10^{-6} and 10^{-7}), as well as those of the behavioural constructs *SelfSabotage* and *Disengagement* (ANOVA $F = 15.7$ and 4.2, p-values < 10^{-9} and .005), are all statistically significant but modest in size. Comparing Figure 7, we see that the pattern is inverted: the first quarter scores highest, the fourth quarter scores lowest.

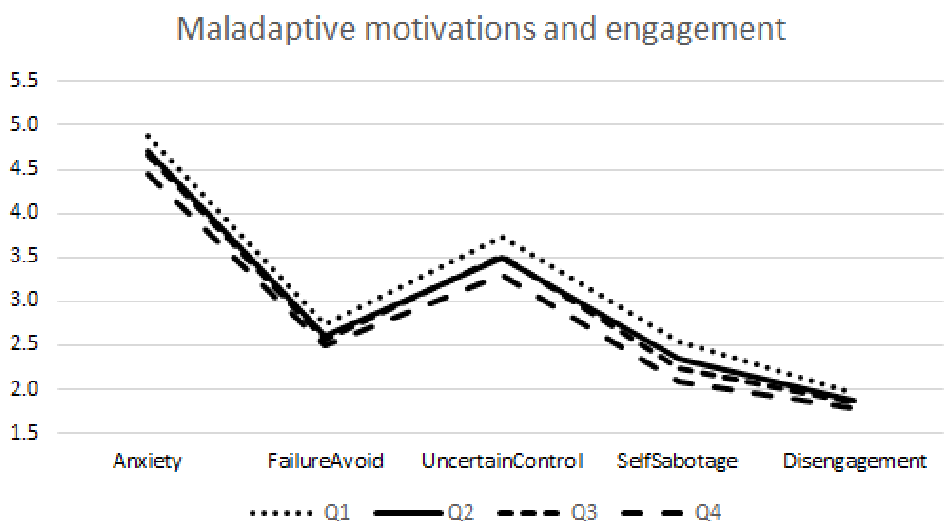

Figure 8. Profiling students by quartile split: Maladaptive motivation and engagement

5.10. Student profiles by quartile split: goal orientations

The fifth learning dispositions instrument administered is that of goal orientations. It provides a remarkable pattern of profile means: see Figure 9. First, profile differences between the four definitions of achievement goals, task-based, self-based, other-based, and potential-based, tend to exceed the differences observed in the valence dimension: approach versus avoidance. Next: differences in the self-related goals are absent (ANOVA F = .90 and .03, p-values = .443 and .992). The other three aspects of the definition, task-based goals (ANOVA F = 6.1 and 4.7, p-values < .001 and .003), other-based goals (ANOVA F = 10.4 and 5.6, p-values < 10^{-6} and .001), and potential-based goals (ANOVA F = 5.4 and 4.7, p-values = .001 and .003), demonstrate significant differences that are substantial in the case of other-based goals. That is: high-achieving students distinguish most from low-achieving students in their competitive learning motivation.

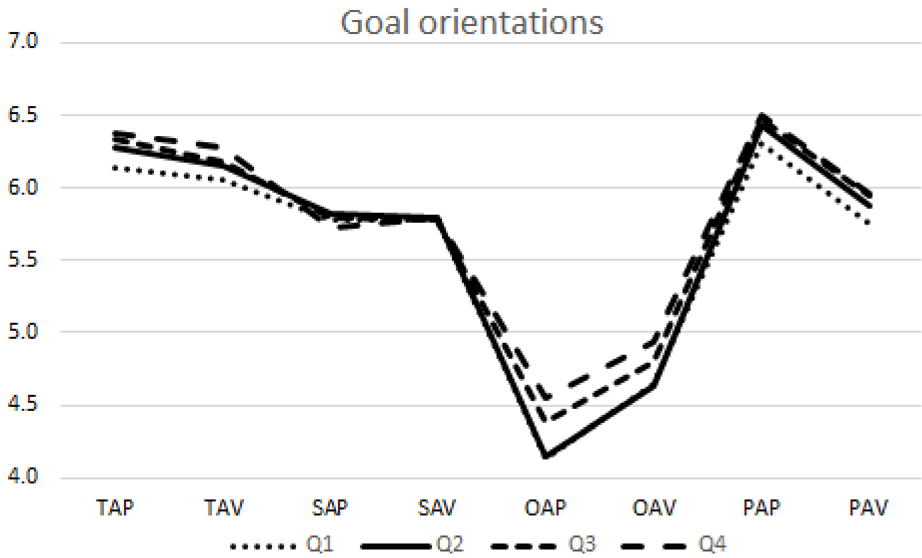

Figure 9. Profiling students by quartile split: Goal orientations

6. Conclusions

The "LA component" of our analysis results in outcomes that are representative for many empirical LA studies that focus on the prediction of academic performance: Lu et al. (2018), Hsiao et al. (2019), and Huang et al. (2020). Yes, there are several "early predictors" of module performance, of which in our study, the first quiz score dominates predictive power. Still, other variables like demographics and learning dispositions do contribute to explained variation. And yes again, trace measures of students' learning activity behaviours are associated with the first quiz score, suggesting an intervention directed at students with low levels in the e-tutorial in the first weeks of the module. Since we collected trace data at the individual level, such learning activity based feedback to students need not be at the profile level but can be tailored towards individual levels, as precision education suggests.

However, does such an intervention approach go beyond the symptom level? Are low activity levels really the cause of underperformance, or no more than a symptom of still hidden causes? In the "D" component of our DLA approach, we demonstrated that underperformance concerning our early performance measure is associated with a whole range of affective, behavioural, and cognitive learning dispositions. Learning attitudes, cognitive learning processing strategies, metacognitive learning regulation strategies, epistemic learning emotions, goal orientations, motivation and engagement, all have their role in explaining variation in the first quiz score. Since our module is the first module students take in the first term of the first year of university, and the dispositions are measured at the start of the module, these dispositions represent learning approaches and beliefs students acquired during six years of high school education and upon transferring to university, bring into our university classes. These suggest being one of the true causes of underperformance: the tendency to follow learning processing strategies of surface type rather than a deep type, the tendency with regard to learning regulation to study quite independent from the curriculum of the school, negative epistemic emotions that developed in high school mathematics classes, insufficient planning skills or lack of persistence, goal orientations that lack a competitive nature.

After linking underperformance with profiles of learning dispositions, the next stage is that of intervention. Intervention studies describe a rich arsenal of interventions linked to dispositional profiles, of which we will highlight some examples related to learning processing and regulation. For instance, Vermunt and Vermetten (2004) identified five "phenomena of dissonance in student learning patterns," with incompatibility of learning strategies and lack of integration between learning strategies being two examples. Especially in the situation of students transferring from one type of education to a new type, as in our context, Vermunt and Vermetten (2004) frequently observed such dissonant patterns. Their main solution, see Vermunt (2003), is in the adaptation of the teaching context by applying innovative teaching methods and practices. This was done in our context by providing a range of alternative formats and approaches for learners to comprehend mathematics in statistics, catering to different learning dispositions.

Other intervention studies see, e.g., Donche et al. (2012), enrich the range of intervention options by including different student feedback types. These researchers distinguish four different learning profiles, low and high self-efficacy in combination with low and high levels of learning regulation, and demonstrate that students of these profiles are characterised by different preferences for learning feedback and are thus best facilitated in their learning in different ways. For example, for a student like Maria, who was mostly watching videos on math problems after a class, understanding whether this comes from low self-efficacy/intense anxiety or procrastination/deficient self-regulation might fundamentally influence the type of support intervention. If it is the former, one could generate automatic feedback like "students like you Maria, who have expressed some anxiety in doing math have benefited tremendously by watching math video 34 before going to class. Therefore, it might be a good idea to consider watching this video 34 before class, but you may also watch it afterward." For the latter type of procrastination student one could provide automatic feedback like "student like you Maria who occasionally struggle to keep a clear study agenda, we recommend that you watch video 34 at least 3 hours before your next class meeting. This will help you to get well prepared for the class meeting."

Therefore, in precision education, our focus is not at the group level but the individual level. In the empirical part of our contribution, we focussed on the properties of four different profiles. The quartile split employed does no more that illustrate that differences in performance and differences in activity levels are associated with differences in learning dispositions. This happens at the individual level even stronger than at the group level: where in our analyses the mean levels of the four quarters may be indistinguishable for some of the dispositions, they may result in large variation at the individual level. It is that variation at the individual level that should serve as input for educational interventions.

Most learning dispositions of behavioural and cognitive type refer to study skills: think of planning, task management, learning processing, and regulation strategies. Every university has programs where these study skills are trained, often organised in extracurricular counselling classes. Therefore, making the step from diagnosis to intervention can be a small one. There are, however, two important differences between current counselling activities and DLA embedded interventions. The first is the sense of urgency: since the counselling is decoupled from the curriculum, it is difficult to demonstrate its relevance. The second difference relates to the generic nature of current counselling. The outcome of DLA is not only a disposition profile that gives students insights into their relative strengths and weakness but attached to that insight in the role these dispositions play in achieving good performance. That second insight will differ from module to module: dispositions relevant for learning mathematics will not be the same as dispositions important in, say, the study of languages. In this combination of these two types of information, DLA can bring its unique contribution to precision education.

The availability of such a broad range of disposition measurements as available in our study will be the exception rather than the rule. From that perspective, this study serves more as a showcase of what can be done with rich disposition data for precision education, where the way of getting such rich data may not be readily generalisable. An important facet of the richness of the data is having a full response of all students, where typically response rates of self-report surveys tend to be low and, typically, the missed cases represent students low in motivation and high in dropout risk, exactly those students it is crucial to have data about. A less crucial facet of the richness of data is the multitude of different disposition surveys. Disposition data tends to be collinear; that is, students with less favourable attitudes will tend to follow less adaptive learning strategies, or depend strongly on external types of motivation. The availability of specific interventions will govern in such a situation the choice of what type of survey instruments to apply: the ultimate goal of precision education is to prevent dropout rather than predict dropout.

References

Abdous, M., He, W., & Yen, C.-J. (2012). Using data mining for predicting relationships between online question theme and final grade. *Educational Technology & Society*, *15*(3), 77–88.

Buckingham Shum, S. & Deakin Crick, 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*, 92-101. New York, NY: ACM. doi:10.1145/2330601.2330629

Buckingham Shum, S. & Ferguson, R. (2012). Social learning analytics. *Educational Technology & Society*, *15*(3), 3-26.

Deakin Crick, R., Huang, S., Ahmed Shafi, A., & Goldspink, C. (2015). Developing resilient agency in learning: the internal structure of learning power. *British Journal of Educational Studies*, *63*(2), 121-160. doi:10.1080/00071005.2015.1006574

Donche, V., Coertjens, L., Vanthournout, G., & Van Petegem, P. (2012). Providing constructive feedback on learning patterns: An Individual learner's perspective. *Reflecting Education*, *8*(1), 114–131.

Elliot, A. J., Murayama, K., Kobeisy, A., & Lichtendfeld, S. (2015). Potential-based achievement goals. *British Journal of Educational Psychology*, *85*, 192-206. doi:10.1111/bjep.12051

Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning, *TechTrends*, *59*(1), 64-71. doi:10.1007/s11528-014-0822-x

Gašević, D., Jovanović, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Earning Analytics*, *4*(1), 113–128. doi:10.18608/jla.2017.42.10

Hart, S. A. (2016). Precision education initiative: Moving toward personalized education, *Mind, Brain, and Education*. doi:10.1111/mbe.12109

Herodotou, C., Rienties, B., Hlosta, M., Boroowa, A., Mangafa, C., & Zdrahal, Z. (2020). The Scalable implementation of predictive learning analytics at a distance learning university: Insights from a longitudinal case study. *The Internet and Higher Education, 45*, 100725. doi:10.1016/j.iheduc.2020.100725

Hsiao, C. C., Huang, J. C. H., Huang, A. Y. Q., Lu, O. H. T., Yin, C. J. & Yang, S. J. H. (2019). Exploring the effects of online learning behaviors on short-term and long-term learning outcomes in flipped classrooms. *Interactive Learning Environments*, *27*(8), 1160-1177. doi:10.1080/10494820.2018.1522651

Huang, A. Y. Q., Lu, O. H. T., Huang, J. C. H., Yin C. J. & Yang, S. J. H. (2020). Predicting students' academic performance by using educational big data and learning analytics: evaluation of classification methods and learning logs. *Interactive Learning Environments*, *28*(2), 206-230. doi:10.1080/10494820.2019.1636086

Hwang, G.-J., Spikol, D., & Li, K.-C. (2018). Guest editorial: Trends and research issues of learning analytics and educational big data. *Educational Technology & Society*, *21*(2), 134–136.

Ifenthaler, D. (2015). Learning analytics. In J. M. Spector (Ed.), *Encyclopedia of Educational Technology* (Vol. 2, pp. 447–451). Thousand Oaks, CA: Sage.

Lian, A.-P. & Sangarun, P. (2017). Precision language education: A Glimpse into a possible future. GEMA Online® *Journal of Language Studies*, *17*(4). doi:10.17576/gema-2017-1704-01

Lin, C.-J., & Hwang, G.-J. (2018). A Learning analytics Approach to investigating factors affecting EFL students' oral performance in a flipped classroom. *Educational Technology & Society*, *21*(2), 205–219.

Lu, O. H. T., Huang, A. Y. Q., Lin, A. J. Q., Ogata, H., & Yang, S. J. H. (2018). Applying learning analytics for the early prediction of students' academic performance in blended learning. *Educational Technology & Society*, *21*(2), 220–232.

Martin, A. J. (2007). Examining a multidimensional model of student motivation and engagement using a construct validation approach. *British Journal of Educational Psychology*, *77*(2), 413-440. doi:10.1348/000709906X118036

Nguyen, Q., Tempelaar, D.T., Rienties, B., & Giesbers, B. (2016). What learning analytics based prediction models tell us about feedback preferences of students. *Quarterly Review of Distance Education*, *17*(3), 13-33.

Non, A. & Tempelaar, D. (2016). Time preferences, study effort, and academic performance. *Economics of Education Review*, *54*, 36–61. doi:10.1016/j.econedurev.2016.06.003

Padulo, L. & Arbib, M. A. (1974). *System theory*. Philadelphia: W. B. Saunders Company.

Pekrun, R., & Linnenbrink-Garcia, L. (2012). Academic emotions and student engagement. In S. L. Christenson, A. L. Reschly, & C. Wylie (Eds.), *Handbook of Research on Student Engagement* (pp. 259–282). doi:10.1007/978-1-4614-2018-7_12

Pekrun, R., Vogl, E., Muis, K. R., & Sinatra, G. M. (2017). Measuring emotions during epistemic activities: The Epistemically-Related Emotion Scales. *Cognition and Emotion*, *31*(6), 1268-1276. doi:10.1080/02699931.2016.1204989

Rienties, B., Cross, S., & Zdrahal, Z (2017). Implementing a learning analytics intervention and evaluation framework: What works? In B. Kei Daniel (Ed.), *Big data and learning analytics: Current theory and practice in higher education* (pp. 147–166). doi:10.1007/978-3-319-06520-5_10

Rienties, B., Tempelaar, D., Nguyen, Q., & Littlejohn, A. (2019). Unpacking the intertemporal impact of self-regulation in a blended mathematics environment. *Computers in Human Behavior*, *100*, 345-357. doi:10.1016/j.chb.2019.07.007

Siemens, G., & Gašević, D. (2012). Guest editorial - Learning and knowledge analytics. *Educational Technology & Society*, *15*(3), 1–2.

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 achievement motivations and personality factors in a range of academic subject-matter areas. *Contemporary Educational Psychology*, *32*(1), 105-131. doi:10.1016/j.cedpsych.2006.10.004

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning analytics in a data-rich context. *Computers in Human Behavior, 47*, 157-167. doi:10.1016/j.chb.2014.05.038

Tempelaar, D. T., Rienties, B., Mittelmeier, J., & Nguyen, Q. (2018a). Student profiling in a dispositional learning analytics application using formative assessment. *Computers in Human Behavior*, *78*, 408-420. doi:10.1016/j.chb.2017.08.

Tempelaar, D. T., Rienties, B., & Nguyen, Q. (2017a). Towards actionable learning analytics using dispositions. *IEEE Transactions on Education*, *10*(1), 6-16. doi:10.1109/TLT.2017.2662679

Tempelaar, D.T., Rienties, B., & Nguyen, Q. (2017b). Adding dispositions to create pedagogy-based learning analytics. *Zeitschrift für Hochschulentwicklung, ZFHE*, *12*(1), 15-35.

Tempelaar, D., Rienties, B. & Nguyen, Q. (2018b). Investigating learning strategies in a dispositional learning analytics context: the case of worked examples. In *Proceedings of the International Conference on Learning Analytics and Knowledge (LAK'18)* (pp. 201-205). doi:10.1145/3170358.3170385

Tempelaar, D., Nguyen, Q., & Rienties, B. (2019). Learning feedback based on dispositional learning analytics. In M. Virvou, E. Alepis, G. Tsihrintzis, & L. Jain (Eds.), *Machine Learning Paradigms* (Vol. 158, pp. 69-89). doi:10.1007/978-3-030-13743-4_5

Vermunt, J. D. (1996). Metacognitive, cognitive and affective aspects of learning styles and strategies: A Phenomenographic analysis. *Higher Education*, *31*, 25–50. doi:10.1007/BF00129106

Vermunt, J. D. (2003). The Power of learning environments and the quality of student learning. In E. de Corte, L. Verschaffel, N. Entwistle, & J. Van Merriënboer (Eds.), *Advances in learning and instruction series. Powerful learning environments: Unravelling basic components and dimensions* (pp. 109–124). Bingley, UK: Pergamon/Elsevier Science Ltd.

Vermunt, J. D., & Vermetten, Y. J. (2004). Patterns in student learning: relationships between learning strategies, conceptions of learning, and learning orientations. *Educational Psychology Review*, *16*, 359–384. doi:10.1007/s10648-004-0005-y

Viberg, O., Hatakka, M., Bälter, O., & Mavroudi, A. (2018). The Current landscape of learning analytics in higher education. *Computers in Human Behavior, 89*, 98-110. doi:10.1016/j.chb.2018.07.027

Williams, A., Sun, Z., Xie, K., Garcia, E., Ashby, I., Exter, M., Largent, D., Lu, P., Szafron, D., Ahmed, S., Onuczko, T., Smith, J., Tempelaar, D. T., Bitting, K. S., Olcott Marshall, A., Christensen, E., Tseng, H., & Walsh, J. (2017). Flipping STEM. In L. Santos Green, J. Banas, & R. Perkins (Eds.), *The Flipped College Classroom, Conceptualized and Re-Conceptualized, Part II* (pp. 149-186). Switzerland: Springer International Publishing. doi:10.1007/978-3-319-41855-1_8.

Yang, S. J. H. (2019). Precision education: New challenges for AI in education [conference keynote]. In *Proceedings of the 27th International Conference on Computers in Education (ICCE)* (pp. XXVII-XXVIII). Kenting, Taiwan: Asia-Pacific Society for Computers in Education (APSCE).