Adding dispositions to create pedagogy-based Learning Analytics

How to cite:
Tempelaar, Dirk; Rienties, Bart and Nguyen, Quan (2017). Adding dispositions to create pedagogy-based Learning Analytics. Zeitschrift für Hochschulentwicklung, 12(1) pp. 15–35.

For guidance on citations see FAQs.

© [not recorded]
https://creativecommons.org/licenses/by-nc-nd/4.0/

Version: Version of Record

Link(s) to article on publisher’s website:
http://dx.doi.org/doi:10.3217/zfhe-12-01/02

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online’s data policy on reuse of materials please consult the policies page.

oro.open.ac.uk
Adding dispositions to create pedagogy-based Learning Analytics

Abstract

This empirical study aims to demonstrate how Dispositional Learning Analytics (DLA) can provide the missing link between Learning Analytics (LA) and pedagogy. Where LA based models typically do well in predicting course performance or student drop-out, they lack actionable data to easily connect model predictions with educational interventions. Using a showcase based on the learning processes of 1069 students in a blended introductory quantitative course, combining demographic and trace data from learning-management systems with self-reports of several contemporary social-cognitive theories, we analyse the use of worked-out examples by students. Students differ not only in the intensity of using worked-out examples but also how they position that use in the learning cycle. These differences can be described both in terms of differences measured by LA trace variables, as well as by differences in students’ learning dispositions. We conjecture that the second description has major advantages for designing educational interventions. Rather than focusing interventions on e.g. low learning activity, only a symptom of suboptimal learning, pedagogy-based interventions focus on potential causes of suboptimal learning, such as applying ineffective learning strategies.

Keywords
Dispositional Learning Analytics, actionable data

1 email: D.Tempelaar@MaastrichtUniversity.nl
1 Dispositional Learning Analytics

‘We can only see pedagogies through the data’, …, as data that ‘encapsulates pedagogic behaviour of the users’ [of the digital learning systems] (GRELLER & DRACHSLER, 2012, p. 53). But in order to have any pedagogical relevance, in order to use Learning Analytics (LA) ‘to evaluate different pedagogical strategies and their effects on learning and teaching through the analysis of learner data’ (GRELLER & DRACHSLER, 2012, p. 48), we will need beyond the mere collection of mere logs of student activity in digital learning systems. Beyond the issue of low predictive power of some of these logged activity data (TEMPELAAR, RIENTIES, & GIESBERS, 2015), the more important issue is that of lack of ‘actionable data’ (GASEVIC, DAWSON, & SIEMENS, 2015): it should be possible to link that data to pedagogical theory, in order to design pedagogy-based learning interventions when predictions signal the need to intervene.

In this contribution, we conjecture, and provide first evidence, that Dispositional LA (DLA, see BUCKINGHAM SHUM & DEAKIN CRICK, 2012; BUCKINGHAM SHUM & FERGUSON, 2012) has the potential to provide a pedagogy-based LA framework. Elsewhere (TEMPELAAR, RIENTIES, & NGUYEN, 2016) we have argued that the DLA infrastructure that combines learning data, generated in learning activities through the traces of the LMS, with learner data: student dispositions, values, and attitudes measured through self-report surveys, distinguish from other LA applications in generating data that is actionable. DLA applications not only provide prediction models that help identify students at risk, but do so using pedagogical descriptors, such as students high in deactivating negative learning emotions, or students using the suboptimal cognitive processing strategies of step-wise learning. Such descriptors are easily linked with instructional interventions based on pedagogical theories, and this way enable concrete actions, such as counselling activities directed at discovering where the negative learning emotions stem from, or practicing the use of deep learning processing strategies.

In this showcase study, companion paper of TEMPELAAR et al. (2016), we will focus on one specific trace variable: students calling for fully worked-out solutions.
What different pedagogical scenarios apply to this feedback option in the LMS? And what learning dispositions act as an antecedent of these scenarios? In answering these questions, we intend to demonstrate the great pedagogical advantage of extending LA into DLA.

## 2 Use of fully worked out solutions

The manner students seek feedback in their self-regulated learning activities constitutes one aspect of pedagogic behaviour (GRELLER & DRACHSLER, 2012). Worked-out examples represent one of the several feedback formats in computer-enhanced environments (DUFFY & AZEVEDO, 2015), formats that amongst others differ in the amount of guidance or assistance provided to students. Pedagogics has identified four main instructional approaches for assisting learners in problem-solving (MCLAREN, VAN GOG, GANOE, KARABINOS, & YARON, 2016), with varying degrees of learner support. The problem-solving approach is positioned in the low guidance end of the continuum, offering little or no feedback to learners. Tutored problem solving provides learners with feedback and hints to solve the problem or construct the schema when learning is stuck. This approach interferes with the learning process only when help is needed; hence, it ensures learners actively attempt to solve the problems. Erroneous examples present learners with flawed examples and instruct them to find, explain, and fix the errors. And at the high end of learner support, MCLAREN et al. (2016) position the use of worked-out examples.

The use of worked-out solutions in multi-media based learning environments stimulates gaining deep understanding (RENKL, 2014). When compared to the use of erroneous examples, tutored problem solving, and problem-solving in computer-based environments, the use of worked examples is the more efficient pedagogical scenario in that it reaches similar learning outcomes in less time and with less learning efforts (MCLAREN et al., 2016). Studies as the above cited are typically laboratory studies, with students assigned to one of the several experimental conditions, each representing one unique pedagogical feedback scenario. In authentic
settings, students mix and mingle diverse pedagogical feedback scenarios, and do so in different orders. So in authentic settings, it is not a matter of what single pedagogical scenario students select out of a list of options, but what combination they prefer, and in what order. As an example of individual differences in pedagogic behaviour unique for authentic settings: some students will avoid using worked-out examples, other students use worked-out examples to start up new learning cycles, whereas a third category uses worked-out examples in the very end of their learning.

Beyond detecting individual differences in preferences for pedagogical scenarios, a next step is to explain these out of differences in learning dispositions. Studies in gender differences in learning mathematics suggest e.g. that female students would profit much stronger than male students from having worked-out examples available at the very start of learning new mathematical concepts (BOLTJENS, 2004). If so, one would expect that to become visible in the trace data of an authentic session where students can choose their own preferred combination of pedagogical scenarios. LA-based models that encompass traces of all relevant pedagogical scenarios, as in the approach suggested by KOEDINGER, MCLAUGHLIN, ZHUXIN JIA, & BIER (2016), may lead to not only knowledge of preferred pedagogical scenarios and their relationship to learning dispositions, but also to their efficiency.

In a digital learning environment as applied in our empirical study, any attempt students do to solve an exercise, can have three different outcomes: the student successfully solves the exercises, provides an incorrect answer, or does not provide any answer, but calls for a worked-out solution. In each of these cases, a student can call for a supportive Hint. These functionalities are examples of Knowledge of the Correct Response (KCR) and Knowledge of Result/response (KR) types of learning feedback; see Narciss (2008). As indicated before, individual differences exist both in the intensity of using worked-out examples, and their timing: in the start, or at the end of each learning cycle. In our study, students undertake on average 1.35 attempts per exercise, using one hint per eight exercises, and asking on average 0.37 worked-out solutions per exercise. As an approximation for what stage of the learning cycle students use the feedback mode of fully worked-out
solution, we constructed a SolutionOrder variable indicating the position of the call of the solution in the series of attempts of any exercise. The variable ranges from zero to one, with lower values indicating that the call takes place in the initial learning phase, and higher values indicating that the call is positioned at the end of the learning process, such as the last attempt preparing for the quiz.

2.1 Context of the empirical study

This empirical study is based on a large-scale course introductory mathematics and statistics as service topics, using an educational system best described as a ‘blended’ or ‘hybrid’. The main component is face-to-face: problem-based learning (PBL), in small groups (14 students), coached by a content expert tutor (SCHMIDT, VAN DER MOLEN, TE WINKEL, & WIJNEN, 2009). Participation in these tutorial groups is required. Optional is the online component of the blend: the use of the two e-tutorials SOWISO for mathematics, MyStatLab for statistics (TEMPELAAR et al., 2015). This choice is based on the philosophy of student-centred education placing the responsibility for making educational choices primarily on the student. However, although optional, the use of e-tutorials and achieving good scores in the practicing modes of the digital environments 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 practicing mode. We chose this particular constellation as it stimulates students with limited prior knowledge to make intensive use of the digital platforms. The bonus is maximized to 20% of what one can score in the exam.

The subject of this study is the 2015/2016 cohort of first-year students, who in some way participated in learning activities in the SOWISO digital tool: 1080 students. We restrict this study to learning activities in the SOWISO tool, because of the richness of trace data generated by the tool, in comparison to the MyStatLab tool. A large diversity in the student population is present: only 23.8% were educated in the Dutch high school system, 45.7% of the students were educated according to the German Abitur system. In the investigated course, students work an
average 9.7 hours in SOWISO, 12% of the available time of 80 hours for learning in both topics.

2.2 Instruments and procedure

Our study combines two different data sources: trace data of the SOWISO learning environment, and self-report survey data measuring learning dispositions. Trace data is both of product and process type (AZEVEDO et al., 2013). SOWISO reporting options of trace data are very broad, requiring making selections from the data. First, all dynamic trace data were aggregated over time, to arrive at static, full course period accounts of trace data. Second, from the large array of trace variables, a selection was made by focusing on process variables most strongly connected to alternative pedagogical behaviours of students. These include the alternative feedback modes preferred by students.

The five types of track data for both topics appear to be collinear: in general, active students spend more time in the e-tutorials, making more attempts, achieving higher mastery, and in doing so using more hints and examples. In total, six trace variables were selected:

- Mastery in the tool, the proportion of exercises successfully solved as product indicator;
- Time in the tool: total connect time;
- #Attempts: total number of attempts of individual exercises;
- #Solutions: total number of worked-out solutions called;
- SolutionOrder: phase in the learning process where worked-out solution is called for;
- #Hints: total number of Hints called for.

In this study, we will make another selection with regard to the self-report surveys measuring student learning dispositions. More than a dozen were administered, ranging from epistemological conceptions about the role of intelligence in learning, to academic buoyance in the learning itself. We will focus here on a selection of six
instruments measuring aspects of self-regulated learning (SRL), feedback seeking, achievement goal setting and learning emotions, since these dispositions have been investigated in recent LA studies (see AZEVEDO et al., 2012; DUFFY & AZEVEDO, 2015, and references therein).

3 Empirical studies into role of dispositions

This first, preliminary section of the empirical research intends to verify that our case satisfies the requirement of a traditional LA application: that trace data are informative for the relevant performance indicators of the course, implying that trace-based prediction models have the potential to signal students at risk. Fig. 1 contains bivariate correlations of the three performance indicators MathExam, the score in the final exam on the Math questions, MathQuiz, the total score in the three Math quizzes, and CourseScore, the final total score for the course, built from quiz scores and final exam scores, and containing both Math and Statistics as topics. A fourth variable added to Fig. 1 is the indicator variable Female, as to check the existence of differences in revealed preferences in using alternative pedagogical scenarios. The six different trace variables are described in the previous section. Starting with the issue of gender differences in revealed feedback scenario use: these seem to be absent. Correlations of #Solutions, SolutionOrder, and #Hints with Gender are non-significant. Correlations of Mastery, Time and #Attempts are at the border of significance, but do not signal difference in scenarios, but only that female students use the same scenarios in a slightly more intensive manner.

Correlations of course performance are quite strong and confirm the prospect of LA: trace variables are crucial building blocks of predictive models of course performance.
3.1 Achievement goals as pedagogic antecedents

Applications of achievement goal theory in LA studies typically employ the two*two framework of goals, distinguishing two goal definitions, mastery goals against performance goals, and two goal valences, approach goals against avoiding goals (see e.g. DUFFY & AZEVEDO, 2015, and references therein). In this study, we apply an extended version of this framework, distinguishing beyond the approach and avoid valence dimensions, four different goal definitions: Task, Self, Other, and Potential goal types (ELLIOT, MURAYAMA, KOBEISY, & LICHTENFELD, 2015). Task, Self, and Potential goals use as a basic standard to define competence the task itself, oneself in the past, and one’s own future potential, respectively. Other goals are normative of character, using a standard based on the comparison with others.

Fig. 1: Task, Self, Other, & Performance achievement goals as antecedents
Both cognitive or product traces (Mastery in tool) and activity or process traces (Time in the tool, #Attempts, #Solutions, #Hints) are positively related to all goal setting types, see Fig. 2. Strongest relationships are for the classical Task goals and for future directed self-related goals: Potential. Within these definitions, stronger impacts exist for the Approach than Avoid valence of goals.

![Graph showing relationships between achievement goals and traces](image)

**Fig. 2:** Task, Self, Other, & Potential achievement goals as antecedents

The negative relationships with SolutionOrder indicate that students scoring high on the Task and Other achievement goals, of both Approach and Avoid valences make use of the worked-out solutions more early in the learning process, relative to students scoring high on the two Self-related goals, both of past and future (Potential) type.

An alternative operationalization of achievement goals that avoids the use of the avoidance valence for goal setting is based on the learning and appearance achievement goal framework proposed by GRANT and DWECK (2003). Correlations depicted in Fig. 3 demonstrate that the Outcome goal provides the strongest stimulus to be active in the digital learning environment, followed by the second
appearance goal of non-normative type: Ability. Lower scores are visible for both the normative versions of the appearance goals and the two types of learning goals.

Fig. 3: Learning and Appearance achievement goals as antecedents

3.2 Cognitive processing strategies as pedagogic antecedents

Self-regulated learning dispositions decompose into preferred processing strategies of students, and metacognitive regulation strategies (VERMUNT, 1996). Processing strategies allow for an ordinal classification from two deep learning orientations, Critical processing and Relating, through Concrete processing, to two surface or step-wise learning orientations: Analysing and Memorising. Fig. 4 demonstrates the relationships of these student dispositions, and tool trace data.
Higher order processing strategies appear to be unrelated to process type of trace data, in contrast to lower order processing strategies. Especially students scoring high on Memorising as preferred strategy, distinguish from other users in high levels of activity, and subsequently high levels of mastery. Students inclined to Memorise distinguish from other strategy preferences in the way they use solutions: early in the learning cycle for those students with deeper strategies, equally spread out in the memorising strategy.

### 3.3 Metacognitive regulation as pedagogic antecedents

The second component of SRL is the metacognitive component, specifying students’ preferences in the regulation of the learning (VERMUNT, 1996). The two main types are a preference for self-regulation versus a preference for regulation by others, or external regulation. Both distinguish two aspects: the regulation of the learning process and the learning content. A third main type is that of lack of regu-
lation. Fig. 5 exhibits the relationships between these five dispositions, and the tool trace data.

![Fig. 5: Metacognitive learning regulation strategies as antecedents](image)

Self-regulation is only very weakly related to tool trace data, in contrast to external regulation. Students, who need external help in regulating their learning, profit from the support by the digital tool. They are more active than other students and reach higher mastery levels. At the same time, they use the worked-out solutions primarily at the start of the learning cycle. This position is mirrored in students who lack regulation; low on mastery and activity levels, high on the SolutionOrder score.

### 3.4 Help-seeking behaviour as pedagogic antecedents

A further facet of SRL is the help-seeking behaviour of students: of Instrumental type, of Executive type, or Avoiding help-seeking type (PAJARES, CHEONG, & OBERMAN, 2004). Here, both the students who avoid help-seeking at all, and the
students seek help with the main goal that someone else solves the problem for them, labelled as executive help-seeking, represent the mal-adaptive types of help-seeking. Instrumental help-seekers search for help as part of their own learning process. How differences in preferred help-seeking behaviour impact tool use, is visible from Fig. 6.

Fig. 6: Help-seeking behaviour as antecedents

The tendency to avoid help-seeking blocks students in using the tool, and building mastery in the tool. And when students with such tendency use the tool, they are inclined to use it in a suboptimal way: to let the tool find the solutions. Seekers of executive help demonstrate a similar pattern, be much less outspoken, whereas seekers of instrumental help demonstrate the opposite pattern.

3.5 Epistemic learning emotions as pedagogic antecedents

Learning emotions of epistemic type distinguish in emotions with positive, negative and neutral valence (PEKRUN & MEIER, 2011). Positively valenced emo-
tions, Enjoyment and Curiosity, are positively related to all tool trace data, with the exception of SolutionOrder. Negatively valenced emotions, Anxiety, Frustration, Confusion, and Boredom, exhibit the opposite pattern, whereas the neutrally valenced emotion Surprise is unrelated to tool trace data: see Fig. 7.

![Graph of Epistemic Learning Emotions as Antecedents](image)

Fig. 7: Epistemic learning emotions as antecedents

4 Discussion and conclusion

Female students do not distinguish themselves from male students in terms of the pedagogical scenarios they apply. In that respect does the expectation derived from the research of BOLTJENS (2004) not come true: there exists a gender difference in overall activity in the tools, with female students being more active participants, but that difference is a generic one. There is no evidence of any differences in revealed preferences with regard to the use of different pedagogical scenarios.
Where our findings do not contribute to the gender-related aspect of the use of worked-out examples, they do contribute to the theorizing on the role of worked-out examples in learning. Different from previous research, all taken place under laboratory conditions with tightly controlled opportunities to use pedagogical scenarios, our authentic setting does allow differentiating between two alternative ways of using worked-out examples. Some students use the worked-out examples early in the learning cycle, most probably as the first encounter with a new mathematical topic. This variant of the worked-out examples pedagogical scenario proves to be an efficient one, corroborating the findings achieved in laboratory research (RENKL, 2014; MCLAREN et al., 2016). However, there exists another variant of the same pedagogical scenario that does not share these positive characteristics. That is the variant where students use the worked-out examples only late in the learning cycle, maybe shortly before upcoming quizzes. This specification of the pedagogical scenario is not effective, given the negative correlations of the SolutionOrder variable with all performance indicators, visible in Fig. 1.

What type of students tends to use these worked-out examples in a suboptimal way? Sections 3.1 to 3.5 provide a description in terms of learning dispositions. For students who set achievement goals using own performance, either in the past or in the future, as a standard, correlations are basically zero. But students setting a Mastery goal, as well as students setting Other goals taking peers as their standards, tend to postpone the use of worked-out examples. The alternative goal setting framework of GRANT and DWECK (2003) confirms this finding: it is the Outcome goal, both in non-normative and normative versions, that correlates most strongly to delayed use of worked-out examples.

Marked differences are also visible in the trace data between students disposed to use superficial cognitive processing strategies, Analysing and especially Memorising, and all other processing strategies. Stepwise learners, VERMUNTs’ (1996) term for learners who focus on analysing and memorising, are much stronger facilitated by the digital learning environment than the Deep learners: they are a lot more active in the e-tutorial, and by spending more time reach higher mastery levels, without falling into the inefficient learning behaviour of late calling of work-
out examples. In contrast, deep learners do not need intensive practicing, can rely on the face-to-face component as the main mode of learning, and when calling for examples from the digital tool, do it mostly in an optimal way: at the start of the learning cycle.

Metacognitive regulation of learning, both with regard learning content and learning processes makes a similar difference as between deep and step-wise learners visible. Self-regulated learners, like deep learners, do not depend on the digital component of the learning blend. And where they do use the e-tutorials, they tend to use examples in the optimal way: at the start of their learning. External-regulated learners depend on their learning environment: their teachers, but also the digital tools. They are making use of these tools more frequently than students at average. Still, they are applying worked-out examples in the same way: mostly to start the learning cycle. It is the learner that lacks any regulation, either of self or external type, who mirrors the position of the externally regulated learner: low in activity, and when using worked-out examples, at the end of the learning cycle.

Help-seeking behaviour is one of the metacognitive dispositions of strong relevance to learning in e-tutorial systems (PAJARES et al., 2004). Instrumental help seeking as the single adaptive version of help seeking resembles deep learning and self-regulated learning in its correlational pattern: lack of relationship with activity levels, weak tendency to use worked-out examples early in the learning cycle. Again, this is mirrored in both executive help seeking and avoidance of help seeking: the mal-adaptive versions. Both correlate negatively with activity in the digital tool, and positively with the delayed use of worked-out examples.

Learning emotions too distinguish adaptive and mal-adaptive types (PEKRUN & MEIER, 2011). Enjoyment and Curiosity are strong examples of the first type. Students scoring high on these two positive and activating emotions do indeed show higher activity levels and optimal use of worked-out examples. In contrast, negative, de-activating emotions, here represented by Frustration, Confusion, and Boredom, block students from using the e-tutorials, and act as stimuli to postpone the use of worked-out examples till later parts of the learning cycle. Surprise and
Anxiety take a special position in the continuum of epistemic emotions: they lack an unequivocal position on the valence and activation dimensions. Anxiety is a negative emotion that can be activating or deactivating, depending on the context. Relationships between Anxiety and activity measures are weak but tend to be negative. The impact of Anxiety on the position of worked-out examples in the learning cycle is, however, clear: Anxiety pushes the use of examples backward in time. Surprise has no straightforward valence, nor activation dimension. In our context, this translates in the complete absence of any impact of the epistemic emotion Surprise.

‘Traditional’ LA applications focus on discovering relationships between trace data and course performance, in order to provide learning feedback related to activities for which these trace data are available. In the context of our own empirical study, such LA type of analysis would learn that higher levels of learning activity and higher mastery levels resulting from these activities contribute to better performance in the course, but that every activity does not translate one-to-one to better performance. In our context: calling for worked-out examples improves performance in general, but the timing of that call makes a strong difference. Looking at worked-out examples at the start of a new learning cycle appears to be a much more effective learning strategy, than calling that same worked-out example at the end of the learning cycle. In order to prevent students from using suboptimal learning strategies, one would come up with campaigns informing students about what learning strategies have proven to be effective, and what ones are suboptimal, but trying to make students aware of these differences, the data itself does not lend to real interventions.

The disposition dimension of DLA adds this aspect of gathering data that is not only predictive but also actionable (GASEVIC et al., 2015). Knowing that students who in their learning of new academic topics depend strongly on the most step-wise processing strategy, the strategy of memorizing, tend to postpone the use of worked-out examples, opens a way to intervention: not by just persuading students to use the worked-out examples earlier, but by focussing on more effective learning strategies. Same for epistemic emotions: pressing one to become more active in
practicing will have low success rates when that low activity level is caused by learning boredom. In such a case, any measure to stimulate learning activity is likely to have adverse effects. From that perspective, the introduction of DLA does even more than provide actionable data: it allows the interventions to be directed at the true causes of the underperforming, rather than its symptoms.

5 ACKNOWLEDGEMENT

The project reported here has been supported and co-financed by SURF-foundation as part of the Learning Analytics Stimulus and the Testing and Test-Driven Learning programs.

6 References


Authors

Dr. Dirk TEMPELAAR || Maastricht University, School of Business and Economics || Tongersestraat 53, 6211 LM Maastricht
https://www.maastrichtuniversity.nl/nl/d.tempelaar
D.Tempelaar@MaastrichtUniversity.nl

Dr. Bart RIENTIES || Open University UK, Institute of Educational Technology || Tongersestraat 53, 6211 LM Maastricht
http://www.open.ac.uk/people/bcr58
Bart.Rienties@open.ac.uk
Quan NGUYEN || Maastricht University, School of Business and Economics || Tongersestraat 53, 6211 LM Maastricht

Q.Lenguyen@Student.MaastrichtUniversity.nl