CASE STUDY

Using SOWISO to realize interactive mathematical documents for learning, practising, and assessing mathematics

André Heck, Korteweg-de Vries Institute for Mathematics, University of Amsterdam, Amsterdam, The Netherlands. Email: a.j.p.heck@uva.nl

Abstract

Successes of computer aided learning in mathematics education have led to high expectations of teachers and learners. A teacher would like to be able to review what students actually do when learning online, how much progress they make, which difficulties arise during learning, and to which subjects attention must be paid in the next lessons. Learners expect that they can consult their instructional materials anywhere, anytime and device-independent, and that they receive immediate and intelligent feedback on their work. At the Faculty of Science of the University of Amsterdam we realize the envisioned interactivity and support in SOWISO. This is a cloud-based environment in which a course has the look-and-feel of an interactive module: texts, randomised examples and randomised exercises with automated feedback are integrated, and online formative and summative assessments are fully supported. We describe and evaluate a SOWISO-based course in basics mathematics for first-year psychobiology students, in which formative assessment and learning through worked examples are key elements of the instructional design.

Keywords: interactive mathematics documents, assessment driven learning, formative assessment, summative assessment, intelligent feedback.

1. Interactive mathematical documents

An interactive mathematical document can be understood as a collection of mathematical pages containing theoretical explanations, examples, exercises, and the like, delivered to the user through hypermedia. A fixed hierarchical structure is not required: users may take their own paths through the pages and they may study the same page, but get different examples and exercises. In the last decades, various systems offering interactive mathematical documents have been developed (Sangwin, 2013), ranging from homework and assessment systems like Maple T.A., STACK, WeBWorK, or Numbas, and notebooks and worksheets offered by systems like *Mathematica*, Maple, or Sage, to intelligent tutoring systems like ActiveMath. SOWISO differs from the aforementioned systems that it integrates instruction, practice and assessment in the form of an interactive online module.

The main characteristic of digital environments for assessment driven learning of mathematics is that each step in the learning path begins with a randomised question that a student is supposed (to try) to answer. If (s)he does not master it (completely), the environment often gives a hint or offers the option to show a worked example or to guide the student in a step-by-step approach. Hereafter the parameter-based implementation of the problem allows the loading of a new version of the problem for the student to demonstrate the newly acquired mastery. SOWISO extends these features with generous possibilities to give a student feedback during the process of completing a task, not only telling whether the given answer to a question was right or wrong, but also whether it meets requirements regarding the mathematical shape of a mathematical formula or the precision of a numeric answer, and what the mistake precisely was in case of an incorrect answer. Immediate feedback allows students to use the online mathematics document as a worksheet in which they progress line by line towards the final solution of a problem, while receiving feedback at each step. An example of a simplification task is shown in figure 1. The left-hand side of the equations are

always given as a prompt and not entered by the student. The student reasons by equivalence, that is, writes line by line equivalent expressions, starting with the given expression and ending with the simplified result in the requested form. Automated, intelligent feedback at each step guides the work.

Simplify the expression
$$\frac{4e^{7x}}{2e^{5x}}$$
 into the form $b \cdot e^{c \cdot x}$.

$$\frac{4e^{7x}}{2e^{5x}} = 2 \cdot \frac{e^{7x}}{e^{5x}} \quad \text{Not yet in the requested form.}$$
Did you simply copy the expression from the question or is a fraction still remaining?

$$\frac{4e^{7x}}{2e^{5x}} = 2 \cdot e^{7x} \cdot e^{-5x} \quad \text{Simplify further}$$

$$\frac{4e^{7x}}{2e^{5x}} = 2 \cdot e^{7x-5x} \quad \text{OK, but not yet in the requested form}$$

$$\frac{4e^{7x}}{2e^{5x}} = 2 \cdot e^{2x} \quad \text{Okay}$$

Figure 1. Use of the online document as a worksheet to solve a simplification task in small steps with intermediate feedback.

We speak of intelligent feedback when it is based on an analysis of the student's answer or action, and when it goes beyond a simple reaction as correct/incorrect. Intelligent feedback is a detailed reaction based on (1) expert knowledge about the mathematical content; (2) a model in which common mistakes and actions are classified; and (3) knowledge about learning and instruction so that the most suitable feedback can be chosen. Part of the feedback design of a prototypical integration example is listed in table 1.

Table 1. Feedback whilst calculating the antiderivative of $\frac{1}{5}x^9$.

Answer	Feedback	Explanation
$\frac{1}{50}x^{10} + c$	Well done	Correct answer
$\frac{1}{5} \left(\frac{1}{10} x^{10} + c \right)$	Correct, but simplify further.	Unusual correct answer that can be simplified
$\frac{1}{50}x^{10}$	Don't forget the constant of integration	No constant of integration
$\frac{1}{50}x^{10} + \frac{1}{5}$	This is a particular solution, but other values than 1/5 can also be a constant of integration.	No general solution
$\frac{9}{5}x^{8}$	You computed the derivative instead of the antiderivative	Reading error?

Because one does not know in advance which formula a student enters, a computer algebra system is used as back engine to recognize mathematical expressions with the same semantic meaning and to respond with appropriate feedback. It is important to realize that equivalence of learner's input with the answer preferred by the lecturer is not good enough to label the answer as correct: when

asked to compute $\frac{1}{4} + \frac{1}{6}$, the answer $\frac{5}{12}$ is certainly correct and maybe $\frac{10}{24}$ is acceptable, but one can frown one's eyebrows with answers such as $\frac{1}{12} + \frac{1}{3}$, $\frac{1}{4} + \frac{1}{6}$, and $\frac{30}{72}$. Thus, it is important that automatic simplification of the back engine can be dealt with and that one can give appropriate feedback for different cases and at different stages of instruction.

Each randomised exercise can also be used as randomised example, so that a student can go through as many examples as needed. Figure 2 shows three instances of one and the same randomised example on powers.

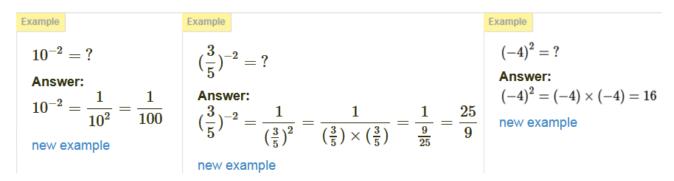


Figure 2. Repeated worked examples originating from one template exercise.

Learning from examples and example-problem pairs have been selected as the main mode of learning. Cognitive Load Theory (Kalyuga, 2011) has underpinned that use of worked examples is effective under certain conditions and is preferred by novices. A computer environment can provide the learner upon demand with as many good quality examples as needed and at the right level of detail to match the learner's needs.

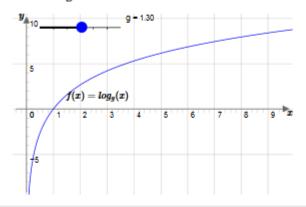
SOWISO offers many tools to implement the instructional design principles associated with effective examples based learning and tutoring. It is a multimedia authoring environment to create interactive, mathematical, HTML5 compliant texts that can be viewed in a standard web browser on any device. A conversion into generated course notes such as PDF, Word or ePub format is supported. JavaScript can be used to add interactivity to a web page: examples are JSXGraph elements for dynamic mathematics (Gerhäuser, et al., 2011) and EjsS simulations of computer models (Saenz, et al., 2015). The dynamic graph of the logarithmic function $log_g(x)$ shown in figure 3 allows students to explore how the graph and the properties of the logarithmic function depend on the value of the base. A video clip from Mathcentre (Matthews, Croft and Lawson, 2013) extends the written explanation. Extra information can be folded behind paragraphs and opened by clicking the 'More' button. Students can rate theory pages and ask questions about theory pages or worked-out solutions of exercises on the users' forum of the class.

Instead of $\log_g(x)$ the notation ${}^g{\log(x)}$ is also used. Furthermore: $\ln(x) = \log_e(x)$.

By moving the slider in the below figure you can get an idea of how the graph of the logarithmic function

$$f(x) = \log_q(x)$$

looks like for various values of the base g.

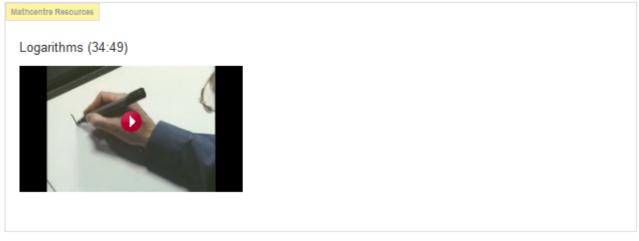


Properties

Some properties of a logarithmic function $f(x) = \log_g(x)$:

- f(1) = 0 (each graph of a logarithmic function passes through the point (1,0)).
- ullet If g>1, then f is an increasing function.
- ullet If 0 < g < 1, then f is a decreasing function.

more...



> continue ? ask question

Figure 3. Course page, illustrating some online features of SOWISO.

2. Case study: Basic mathematics for psychobiology

In this paper, experiences with the 'Basic Mathematics for Psychobiology' course in the first year of the psychobiology bachelor programme at the University of Amsterdam are described. It concerns the course settings in the study year 2015-2016, course results, and dispositional learning analytics (Buckingham Shum and Deakin Crick, 2012), in which learning data and learner data are combined as a means to provide instructors actionable feedback.

2.1. Course settings

One hundred and fifty-six first-year students (38 male, 118 female) participated in this course of ten weeks. The course design was assessment driven and examples based instruction with intelligent feedback to acquire mathematical cognitive skills needed by psychobiologists. It was the second study year that Basic Mathematics was taught without being embedded in some neuroscience course. Course evaluation and research done in the previous year by Hoek (2015) showed that students had difficulties with providing correct and useful self-explanations while studying worked examples, and could not always recognize the important differences between an exercise and a worked example that they used before. This had resulted in adaptations of the online module. The course content was reduced, and video clips and more detailed explanations in worked examples were added. An example of an annotated solution to an exercise is shown in figure 4.

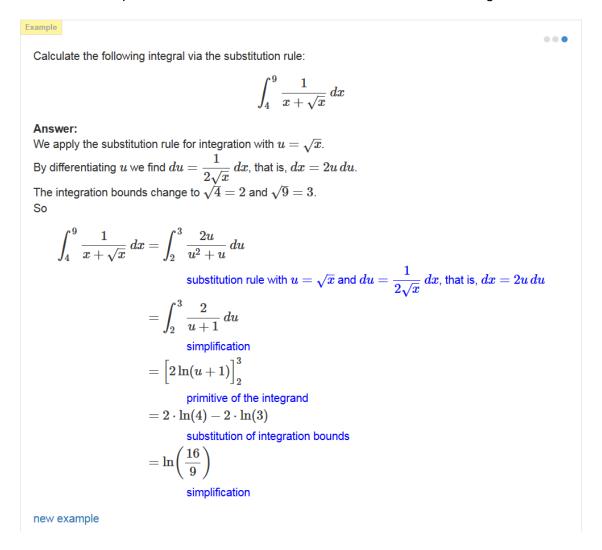


Figure 4. Annotated solution to an exercise, with verbose explanations of each step in blue.

The online module of this course in SOWISO is available in Dutch and English at the website www.nucleon.org/ (login name and password EAMS2016). The contents concern calculus and its applications in psychobiology: (1) Precalculus – calculating with numbers, calculating with letters, solving equations; (2) Chemical calculations; (3) Functions – all standard functions; (4) Differentiation and derivatives; (5) Differentials and integrals; (6) Complex numbers; (7) Models of growth – exponential growth, limited exponential growth, logistic growth; and (8) Ordinary differential equations – basic notions, separable 1st-order ODEs and linear 2nd-order ODEs with constant coefficients.

Treatment of each mathematical subject ended with a randomised digital assessment for which the pass mark was 7.5 (out of 10). Besides having a summative role, the main purpose of these tests was assessment for learning: students could objectively verify their learning progress in mastering a mathematics subject. These tests were short, but there were many. In total, students had to take forty-nine tests. Only in case they passed all tests, they received a course mark, regardless of positive results on the final examination. The final exam took place in a digital examination room (figure 5). The exam was taken in SOWISO. The questions were randomised having two versions of the same difficulty. Students solved the exercises by pencil and paper, entered their final answers in the SOWISO test, and submitted both their papers and online test. The digitalisation of the exam not only reduced the correction process, but also allowed students to compare their answers with the worked-out solutions and made the review process before the examination easier and more effective.



Figure 5. Picture of the digital examination room with psychobiology students at work.

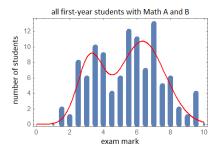
2.2. Course results

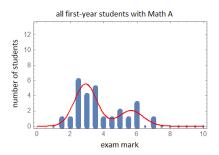
Data of student activity were collected by logging students' actions, their performance in the digital learning environment in formative assessments, and from evaluation reports of the students. Personal data of students such as mathematics background, initial level of mathematical competence, mathematics anxiety, test anxiety, and motivation and engagement, were collected in SOWISO using validated questionnaires (Hopko, et al., 2003; Spielberger, 1980; Martin, 2007).

By the mathematics background of the students we mean the mathematics examination programme that students took at upper pre-university level, namely, Mathematics A or B. Mathematics A is meant for students who prepare for academic studies in social or economic sciences, or other related studies. Its core subjects are statistics and probability and a little calculus. Emphasis is on applying mathematics. Mathematics B is the mathematics needed for exact sciences and technical studies,

and its core component is calculus. In addition to the difference in core mathematical subjects the purpose and type of problems of the two curricula differ. The original intentions for Mathematics A, based upon realistic mathematics education, were (HEWET, 1980, p.19): "Mathematics A is intended for students who will have little further education in mathematics in their academic studies, but who must be able to use mathematics as an instrument to a certain extent. In particular, we have in mind those who have to prepare themselves for the fact that subjects outside the traditional sciences are more frequently being approached with the use of mathematics." However, soon after its introduction in 1989, the initial objectives were not met anymore in the central Mathematics A final examinations and nowadays open ended problems generally tend to ask of the students only to do a calculation, draw a graph, read a graph, substitute a few values in a formula, and so on, and do not really assess modelling and problem solving competencies.

A significant percentage (30%) of the first-year psychobiology students had taken at secondary school the Mathematics A examination programme, which prepared them less for exact sciences than the Mathematics B examination programme taken by the rest of the first-year students. This became obvious when the exam results of the basic mathematics course were inspected. Only 19% of the students with Mathematics A background passed the exam, whereas 69% the students with Mathematics B background were successful; see the cross tabulation in table 2. In figure 6, three histograms of exam marks are plotted: the left, middle, and right diagram show the distribution of marks for all first-year students with known mathematics background (107), for first-year students with Mathematics A background (26) and for first-year students with Mathematics B background (81), respectively. On the left-hand side and in the middle there appear to be binormal distributions of marks, whereas it becomes a normal distribution for Mathematics B students only. These diagrams suggest that there was a split between the students based on their mathematics background, which can almost be described by combining the distributions for the two groups. Indeed, when the Mann-Whitney statistic was calculated to determine whether there was any statistically significant difference in exam marks regarding the mathematics background (U = 1752, z = 5.10, p < 0.001), a statistically significant difference between students with Mathematics A and Mathematics B was found, with large effect size (0.49). A cross tabulation (table 2) found that Mathematics B performed better than the Mathematics A students on the final exam. From the diagnostic entry test in the first week of the course we already knew that students with Mathematics B background entered the course with a statistically significant higher competence level than their peers with Mathematics A background. The mathematical knowledge and skills gap between these two groups of students did not get smaller during the course.





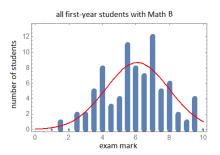


Figure 6. Distribution of exam marks: a binormal distribution for all first-year students with known mathematics background (left) and for first-year students with Mathematics A background (middle) and a normal distribution for first-year students with Mathematics B background (right).

Table 2. Cross tabulation of pass-fail exam results and mathematics background of the students.

	pass	fail	total
Mathematics A	5 (19%)	21 (81%)	26 (24%)
Mathematics B	56 (69%)	25 (31%)	81 (76%)

2.3. Mathematics anxiety and test anxiety

The level of mathematics anxiety amonst first-year students was measured via the Dutch translation of the Abbreviated Math Anxiety Scale (Hopko, et al., 2003). The mean AMAS score (M = 2.9, SE = 0.43) of 155 students was low on a scale from 1 (no anxiety at all) to 10 (panic) and in agreement with students' self-estimates. We calculated the Mann-Whitney statistic to determine whether there were any statistically significant differences in mathematics anxiety regarding mathematics background and gender. No significant differences were found at the significance level of 0.05 (U = 2491, z = 1.12, p = 0.263 for gender and U = 1806, z = 1.76, p = 0.079 for mathematics background). This is in agreement with result collected in previous study years in life sciences and in agreement with the data and results of Erurtan and Jansen (2015), who found in their studies on the relationship between children's emotional experiences of mathematics and their mathematics performances that only differences in test anxiety (boys had lower test anxiety than girls) but not in mathematics anxiety appeared amongst Dutch pupils in grades 3-8 (ages 7-15 years).

The level of test anxiety amonst first-year students was measured via the Dutch translation of the Test Anxiety Invenory (Spielberger, 1980). The mean TAI score (M = 3.7, SE = 0.96) of 153 students was low on a scale from 1 (no anxiety at all) to 10 (panic) and in agreement with students' self-estimates. We calculated the Mann-Whitney statistic to determine whether there were any statistically significant differences in test anxiety regarding mathematics background and gender. No significant difference was found at the significance level of 0.05 (U = 1915, z =0.98, p = 0.328) for mathematics background. But a statistically significant gender difference in TAI scores (U = 2887, z = 2.97, p = 0.003) was found, with moderate effect size (0.24). Male students had lower test anxiety than female students. This is in agreement with results collected in previous study years in life sciences and in agreement with the results of Erurtan and Jansen (2015).

The AMAS scores were significantly but moderately related to the TAI scores (Spearman's rho $r_s = 0.56$, p (two-tailed) < 0.001, with bias corrected and accelerated bootstrap 95% confidence interval [0.40, 0.69]). This is not surprising because AMAS has items (in the mathematics evaluation anxiety component of the scale) that are comparable with items in TAI. Concerning the relationship between emotional experiences of mathematics and mathematical performance, only mathematics anxiety was statistically negatively related with exam scores at the significance level of 0.05, but it was a weak correlation (Spearman's rho $r_s = -0.16$, p (one-tailed) = 0.04, with bias corrected and accelerated bootstrap 95% confidence interval [-0.33, 0.02]).

2.4. Dispositional learning analytics

The 'Motivation and Engagement Wheel' framework (Martin, 2007) includes both thoughts and behaviours that play a role in learning and consequently in course performance. Both are subdivided into adaptive and maladaptive forms. Adaptive thoughts (the 'boosters') consist of self-belief (SB), value of school (V), and learning focus (LF), whereas adaptive behaviours consist of planning (PLN), task management (TM), and perseverance (P). Maladaptive thoughts (the 'mufflers') include (test) anxiety (A), failure avoidance (FA), and uncertain control (UC), and lastly, maladaptive behaviours (the 'guzzlers') include self-sabotage (SS) and disengagement (D). These behaviours and thoughts are assessed via the 'Motivation and Engagement Scale – University / College' (MES-UC) self-report instrument. One hundred and fifty-four (154) students filled out the MES-UC questionnaire at the

start of the course. Cronbach's alpha for the instrument was 0.76, close to the literature value of 0.78. No statistical significant differences between students with Mathematics A and B background were found on all scales of the MES-UC questionnaire, except for test anxiety: students with Mathematics A background scored higher on this scale. Gender differences were found: female students scored statistically significantly higher on boosters and guzzlers (in particular on planning, study management and test anxiety), but the effect size was small.

The effect of adaptive and maladaptive behaviours and thoughts on the activity level of students in the online module and on the course performance measured by the mark on the final exam was investigated. Activity level was measured by the number of theory pages that a student has viewed and the number of exercises that a student had tried to solve (by entering at least one response). Figure 7 represents the correlations between the MES-UC scales and the exam mark. It shows that self-belief, persistence and planning had the largest positive impact on the exam mark. As expected, maladaptive thoughts and behaviours (especially self-sabotage) had a negative impact.

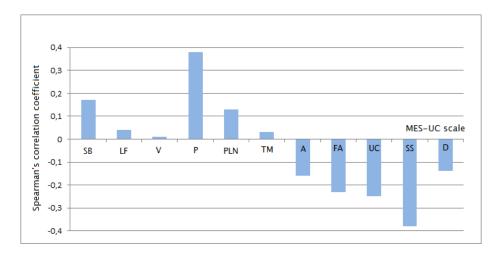


Figure 7. Correlation between MES-UC scales and the exam marks.

Figure 8 represents the correlations between the MES-UC scales and the activity level of a student. As expected, adaptive behaviours had a positive impact on the activity of students in the module. But adaptive thoughts apparently had less impact; self-belief may even have a negative impact. Anxiety and uncertain control had a positive impact on the activity level of students, whereas selfsabotage had amongst the maladaptive behaviours the strongest negative impact on the student's activity level. These findings are in good agreement with results found in other studies, for example in mathematics and statistics courses for first-year economy students (Tempelaar, Rienties, and Nguyen, 2016). When the Jonckheere-Terpstra test (N = 121) was calculated to determine whether there was a statistically significant difference in exam scores and level of practise (with three categories of 'no drop out', 'drop out as soon as new mathematics was taught', and 'no or irregularar participation'), a statistically significant difference (J(2) = 334, z = 5.47, p < 0.001) was found, with large effect size (0.50), between students who regularly practised and did not drop out, and those students who dropped out when it became difficult because of new mathematical subjects or who practised irregularly or not at all during the course. Students with Mathematics A background dropped out significantly more than students with Mathematics B Background, maybe because they were insufficiently prepared for the course. In summary, one could conclude that looking only at worked-out solutions does not lead to better marks, but that sufficient practising and stamina really helps to be successful.

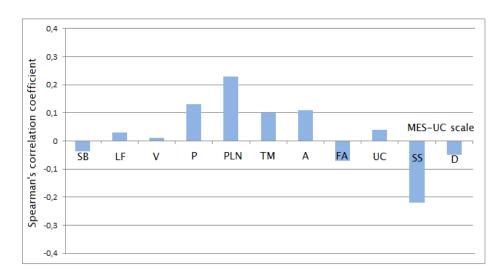


Figure 8. Correlation between MES-UC scales and the student activity level.

3. Conclusions

This case study and a similar case study about a mathematics course for computer science students (Heck and Brouwer, 2015) illustrate that the SOWISO environment offers an e-learning environment for mathematics in which students

- practise more than in a classical instructional setting;
- get automated feedback on their work without direct involvement of a teacher or tutor;
- use interactive material for self-study and practise mathematics;
- ask questions on the forum and communicate with peers;

and in which teachers

- create interactive online course notes (re-using instructional materials);
- monitor the students' level of activity and learning progress;
- carry out digital assessments (formative and summative).

More research is needed to find out if only procedural knowledge and skills have been mastered, or if problem solving skills have been developed, how long the acquired competencies retain, and whether also transfer of mathematical competencies has been reached. Anyway, this course has proved to many of the students that they are able to master mathematical knowledge and skills needed by psychobiologists. At least, when the students have a proper mathematics background: in this study we found that mathematical knowledge and skills gap between students with Mathematics A and B background was already significant at the start of the mathematics course and did not improve during the course. We also found that mathematics anxiety and test anxiety were low amongst the Dutch students, but female students reported a slightly higher level of test anxiety. The correlation between mathematics anxiety and mathematics performance was significant but weak in this study. No statistically significant relationship was found between test anxiety and exam performance. Dispositional learning analytics provided actionable feedback: in this study we learned for example that we should stimulate students (1) to actively use the interactive mathematical documents and do the exercises instead of only reading the text; and (2) not to rely on just one example, but to use more examples and read the theoretical explanations about them.

4. References

Buckingham Shum, S. and Deakin Crick, R., 2012. Learning Dispositions and Transferable Competencies: Pedagogy, Modelling and Learning Analytics. In: *Proceedings of the 2nd International Conference on Learning Analytics & Knowledge*. New York: ACM Press. pp.92-101. Available at: http://oro.open.ac.uk/32823/1/SBS-RDC-LAK12-ORO.pdf [Accessed 9 October 2016].

Erturan, S. and Jansen, B., 2015. An investigation of boys' and girls' emotional experience of math, their math performance, and the relation between these variables. *European Journal of Psychology of Education*, 30(4), pp.421-435.

Gerhäuser, M., Valentin, B., Wassermann, A. and Wilfahrt, P., 2011. JSXGraph – dynamic mathematics running on (nearly) every device. *Electronic Journal of Mathematics and Technology*, 5(1).

Heck, A. and Brouwer, N., 2015. Digital assessment-driven examples-based mathematics for computer science students. In N. Amado & S. Carreira (eds.), *Proceedings of the 12th International Conference on Technology in Mathematics Teaching* (ICTMT12). University of Algarve, Portugal. pp.403-411. Available at http://hdl.handle.net/10400.1/6081 [Accessed 9 October 2016].

HEWET (1980). HEWET-Report (in Dutch) from the 'werkgroep van advies voor de herverkaveling eindexamenprogramma's Wiskunde I en Wiskunde II vwo [Working group for advice on the reallocation of the examination programmes Mathematics I and II VWO] (1980). The Hague: Staatsuitgeverij.

Hoek, R., 2015. *Setting the Right Example*. Master thesis, University of Amsterdam. Available at: http://www.scriptiesonline.uba.uva.nl/document/634733 [Accessed 9 October 2016].

Hopko, D.R., Mahadevan, R., Bare, R.L. and Hunt, M.K., 2003. The abbreviated math anxiety scale (AMAS). *Assessment*, *10*(2), pp.178-182.

Kalyuga, S., 2011. *Managing Cognitive Load in Adaptive Multimedia Learning*. Charlotte, NC: Information Age Publishing.

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), pp.413-440.

Matthews, J., Croft, T. and Lawson, D., 2013. Mathcentre 2013. MSOR Connections, 13(1), pp.34-40.

Saenz, R., Esquembre, F., Garcia, F.J., de la Torre, L., and Dormido, S., 2015. An architecture to use Easy Java-JavaScript Simulations in New Devices. *IFAC-PapersOnLine*, 48(29), pp.129-133.

Sangwin, C., 2013. Computer Aided Assessment of Mathematics. Oxford: Oxford University Press.

Spielberger, C., 1980. The Test Anxiety Inventory. Palo Alto, CA: Consulting Psychologist Press.

Tempelaar, D.T., Rienties, B., and Nguyen, Q., 2016. Achieving actionable learning analytics through dispositions. *IEEE Transactions on Learning Technologies* [Accepted].