A Framework of Quality Indicators for Learning Analytics

Learning Analytics Community Exchange

A Framework of Quality Indicators for Learning Analytics

Learning Analytics Review 2

ISSN:2057-7494

By: Maren Scheffel

Published: 21st July 2015

Keywords: learning analytics, quality indicators, evaluation framework, group concept mapping

Abstract: The LACE project has established a first version of a framework of quality indicators for learning analytics, based on a group concept mapping study with experts. The group concept mapping approach is explained, and steps in the framework creation process described, as well as the framework itself. The framework was turned into an applicable tool and evaluated with a group of learning analytics experts. The results of the evaluation revealed several weak points in the first version of the framework, and the experts supplied several suggestions and recommendations on how to improve the framework further.

Contents

Executive Summary

The field of learning analytics has been growing steadily in recent years. The development and evaluation of more and more learning analytics dashboards bears witness to the belief that learning analytics can provide added value to learners as well as educators. There is, however, not as yet sufficient hard evidence for or against different types of learning analytics, and it is difficult to compare the results different tools and methods. The LACE project has therefore developed a proposal for a framework of quality indicators for learning analytics that supports the standardisation of the evaluation of learning analytics tools, and provides a means of capturing evidence of the impact of learning analytics on educational practices in a standardised manner.

A first version of the framework has been developed, by means of a group concept mapping study with experts from the field of learning analytics. The study collected statements from within the learning analytics community about what a quality indicator of learning analytics entails. This was followed by a process of sorting and rating of these statements, carried out by a group of learning analytics experts. The results were then used as input to several statistical techniques that led to the first version of the framework. This paper explains the group concept mapping approach and presents the steps in the framework creation process, as well as the framework itself.

In a second study we then conducted an evaluation with a small number of learning analytics experts by turning the framework into an applicable tool. The results of that study revealed several weak points in the first framework version and the experts supplied several suggestions and recommendations on how to improve the framework further. Again, the evaluation process is presented in this paper, together with the results of the study.

Introduction

Learning analytics (LA) is a field of research that builds on ideas from other fields such as process mining, business intelligence, data processing, information retrieval, technology-enhanced learning, educational data mining and data visualisation. Over the last few years, this research field has been growing steadily, gathering more and more attention in an educational world trying to understand and optimise learning. While closely related topics have already been addressed in the related domains mentioned, learning analytics now forms its own domain with associations, societies, courses, workshops, journals, etc., specifically dedicated to it. Many tools and dashboards have been developed to support learners as well as educators in various settings (e.g. primary schools, high schools, universities, workplace, etc.) and for various purposes (e.g. reducing drop-out rates, fostering awareness of different stakeholders, making learning and teaching processes more effective and efficient etc.). Reflecting the variety of these settings is the variety of the measures used to evaluate the success of learning analytics, and to compare the results of such evaluations. Additionally, from the LACE project's Evidence Hub we know that there is very little hard evidence for either the success or the failure of learning analytics tools. While the value of learning analytics is thus clearly recognised, little research has been done to provide a way to compare these tools and dashboards and their effect on learning with one another. Within the LACE project, we have therefore started work on a framework of quality indicators that can be used to measure and compare the impact of learning analytics and help standardise the evaluation of learning analytics tools.

The first version of the quality indicator framework has been developed using a group concept mapping approach. This process is described in the section "Framework Study 1". A detailed description of the study and its results can be found in Scheffel et al. (2014). As a next step we then evaluated this first framework version. The evaluation results are presented in the section "Framework Study 2". A detailed description can be found in Scheffel et al. (2015).

Framework Study 1

Methodology

Group concept mapping is a very structured approach that makes use of qualitative as well as quantitative measures. The methodology allows us to identify a group's common understanding of a specific issue, in our case quality indicators of learning analytics tools, i.e. it is a bottom-up approach in which the ideas submitted are generated by the community itself. The tool we used is available online[1] and consists of three steps: (1) brainstorming, (2) sorting and (3) rating. Apart from applying statistical techniques such as multidimensional scaling and hierarchical clustering, the tool also visualises the outcomes in a way that helps interpret the results.

The brainstorming phase took place during and shortly after the Learning Analytics and Knowledge conference 2014. The study was advertised through several channels such as personal contact, emails, Twitter and project websites, and was accessible to anyone via a link for a period of ten days. In total, 74 people participated resulting in a total of 103 statements about what a quality indicator for learning analytics is. For the sorting and rating phase we contacted 55 learning analytics experts

[1] http://www.conceptsystemglobal.com

ISSN:2057-7494

who had 14 days to carry out the tasks. 23 of them did the sorting task and 21 rated the collected ideas. For the sorting we asked the experts to cluster the statements according to their similarity in meaning in as many different clusters as they deemed appropriate. The rating was to be done according to a statement's *importance* and *feasibility* on a 1-to-7 scale each. Looking at the demographics of the experts showed that the average expert was a researcher at a university with an advanced expertise in learning analytics and more than 10 years of work experience.

Results

The first analysis the tool provides is a point map that shows the outcome of the multidimensional scaling. Figure 1 shows the point map of the 103 statements collected in the brainstorming phase. The statements that are close together on the map are also close to one another in meaning according to the experts' sorting.

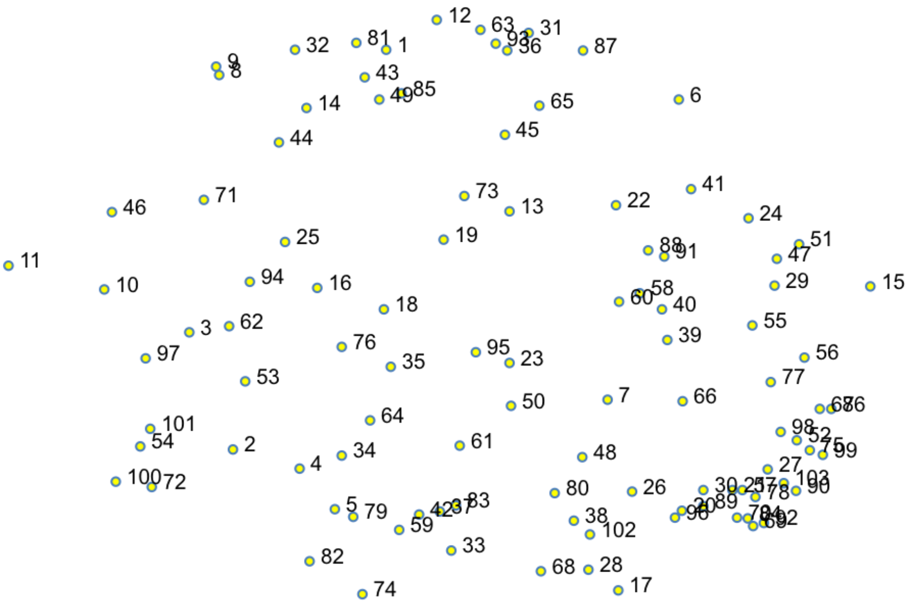


Figure 1 - Point map of the 103 quality indicators.

Once the point map has been created, the tool carries out a hierarchical cluster analysis. It does so by offering several cluster solutions to the point map in the form of a replay map that ranges between 2 and 15 clusters. Figure 2 shows the replay map with 15. We carefully looked at the different solutions generated by the tool, going back and forth between them. We looked at the statements in each cluster, and checked whether the merging or unmerging of clusters from one step to the next made sense, i.e. whether the statements in the clusters at any given time were as similar in theme as possible and the clusters were sufficiently diverse to describe the full set of statements. The solution with 8 clusters seemed most sensible to us.

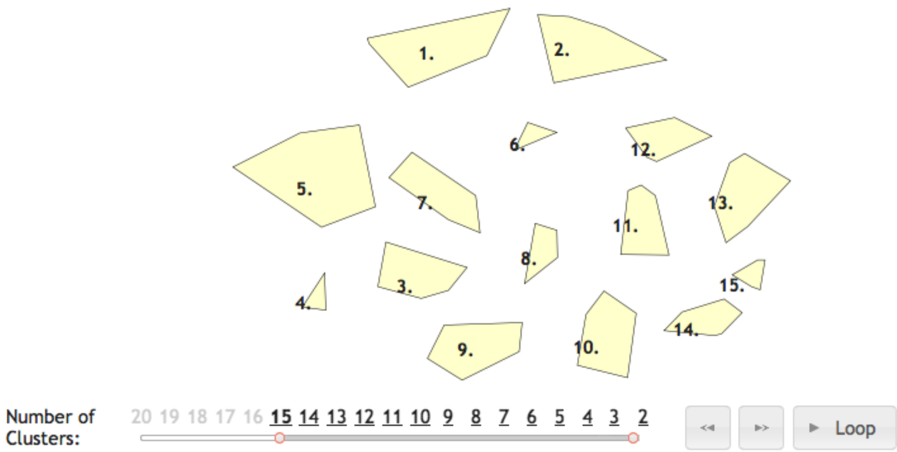


Figure 2 - Replay map showing 15 clusters.

To find labels for the clusters we looked at the ones automatically suggested by the tool as well as the statements in the clusters and their overarching theme. Figure 3 shows the eight clusters and the labels we settled on: *1. Data: open access*, *2. Data: privacy*, *3. Acceptance & uptake*, *4. Learning outcome*, *5. Teacher awareness*, *6. Learning performance*, *7. Learning support*, and *8. Student awareness.*

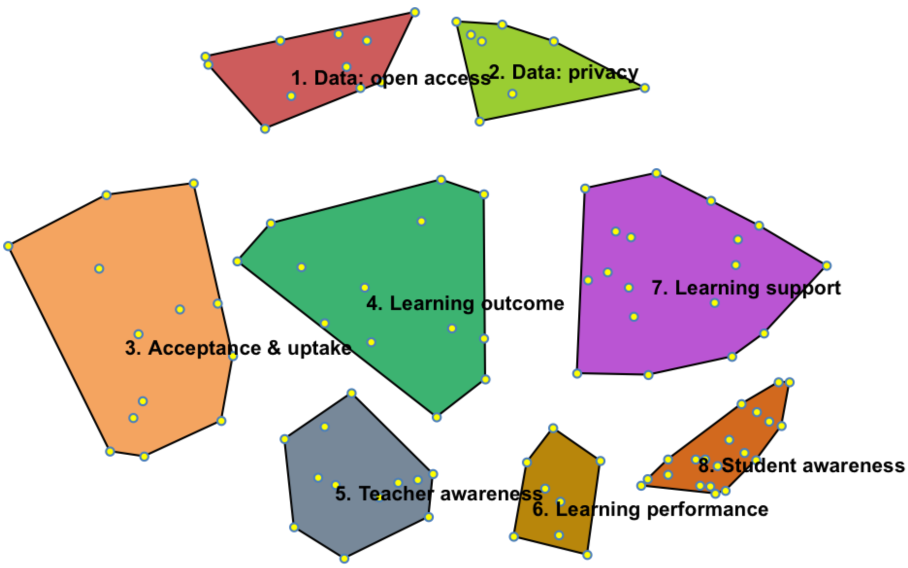


Figure 3 - Cluster map with labels.

The two clusters in the North (1, 2) both deal with data, access, methods, algorithms, transparency and privacy, i.e., with technical issues, while the clusters in the South (5, 6, 8) deal with awareness, reflection, performance and behavioural change of students and teachers, i.e., with human issues. The "technical North" (*Data: open access* and *Data: privacy*) and the "human South" (*Teacher awareness*, *Learning performance* and *Student awareness*) are bridged by a wide layer of learning-related clusters (*Learning outcome* and *Learning support*). Apart from the North-South view, one can also look at the map with an East-West perspective: The three Eastern clusters (*Data: privacy*, *Learning support* and *Student awareness*) are more concerned with issues during the learning process while the Western clusters (*Data: open access*, *Learning outcomes*, *Teacher awareness* and

Learning performance) are slightly more concerned with issues of learning output and results. This division is of course not to be seen strictly, but these groupings clearly show a thematic tendency.

The tool then overlays the experts' ratings for *importance* and *feasibility* of all the statements included in each cluster. Figure 4 shows the rating map for *importance*. The more layers a cluster has, the more important the experts deemed it. There are several things we can deduce from this visualisation. The learning-related middle layer (clusters 4 and 7) is deemed highly important by the learning analytics experts, as are all Eastern clusters (2, 7, 8). Generally one can thus say that the focus of *importance* is on the learning process-related clusters.

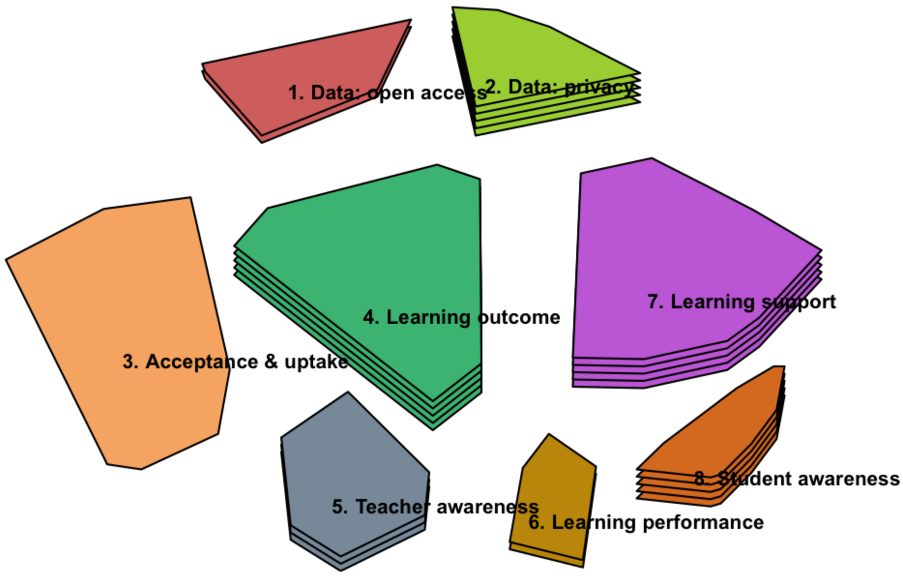


Figure 4 - Rating map on importance.

Figure 5 shows the rating map for the *feasibility* rating. Again, the more layers a cluster has, the more feasible the experts rated it. Analysing the feasibility ratings shows a clear North-South divide. The technically oriented clusters in the North (1,2) are deemed most feasible by the experts, followed by the learning-related layer in the middle (4,7) and concluded by the human-related clusters in the South (5,6,8).

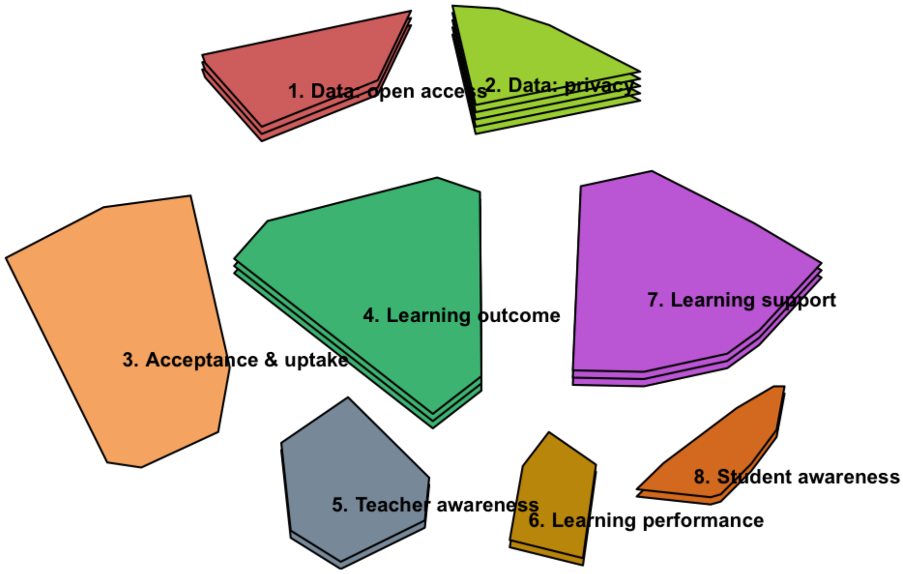


Figure 5 - Rating map on feasibility.

When deciding upon the criteria of the framework it is important to find a good trade-off between the *importance* ratings and the *feasibility* ratings. Due to the high *importance* of the clusters about *Data: privacy*, *Learning outcome*, *Learning support* and *Student awareness* it seems to be sensible to use them as a basis for the criteria of the framework. From the results of the GCM study we can thus identify four topic areas that can be turned into criteria for the framework: the first deals with anything related to data, algorithms, transparency and privacy (criterion *Data Aspect*, quality indicators *Transparency*, *Data Standards*, *Data Ownership*, and *Privacy*). The second topic area concerns support for students and teachers during the learning process, i.e., while using learning analytics tools (criterion *Learning Support*, quality indicators *Perceived Usefulness*, *Recommendation*, *Activity Classification*, and *Detection of Students at Risk*). The third topic area deals with the results at the end of the learning process, i.e. any issues of output, consequence, performance, outcome etc. In this case, however, it is not primarily to be seen in relation to individual student performance, e.g., their grades, but refers to the learning analytics tools' results and outcomes (criterion *Learning Measures and Outputs*, quality indicators *Comparability*, *Effectiveness*, *Efficiency*, and *Helpfulness*). The fourth topic area contains the quality indicators about students and educators during the learning processes (criterion *Objectives*, quality indicators *Awareness*, *Reflection*, *Motivation*, and *Behavioural Change*). As we consider indicators of organisational issues to be an important aspect when considering the evaluation of learning analytics tools, we decided to add a fifth criterion to the framework (criterion *Organisational Aspects*, quality indicators *Availability*, *Implementation*, *Training of Educational Stakeholders*, and *Organisational Change*). Figure 6 shows the framework we created from the results of the group concept mapping study.

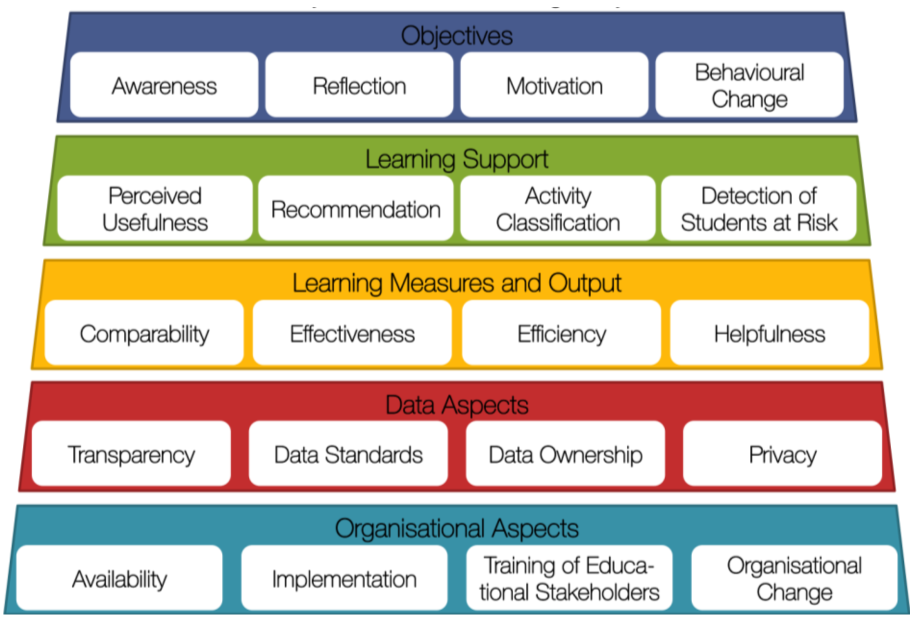
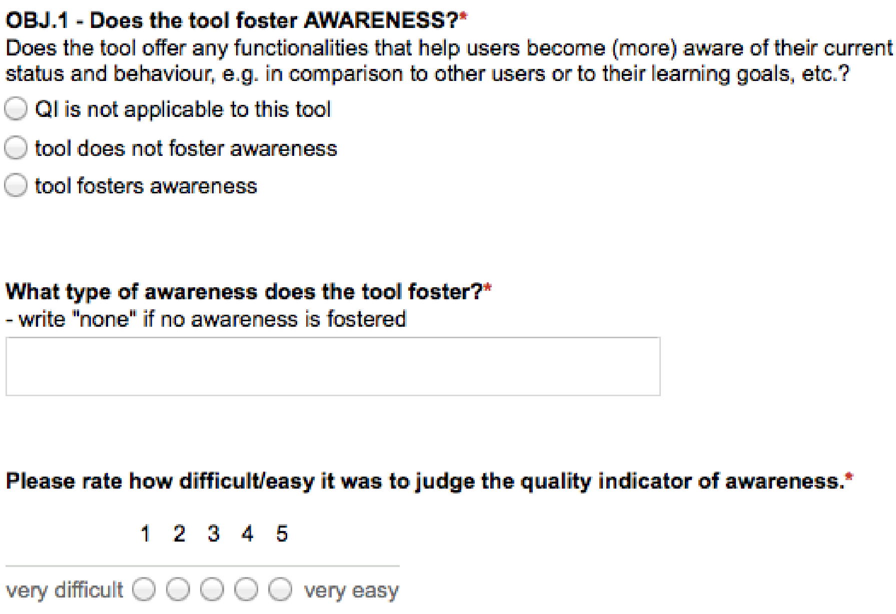


Figure 6 - First version of the learning analytics quality indicator framework.

Framework Study 2

Methodology

In order to find out whether the framework is applicable for the evaluation and comparison of learning analytics tools or whether it needs improvements, we conducted an evaluation study. We therefore turned the framework into a questionnaire, i.e. an applicable tool. For every quality indicator we asked the same three questions about (1) the presence or absence of the indicator, (2) the mode or style of the indicator, and (3) the ease or difficulty of judging the indicator. Figure 7 shows an example of such a question block.



OBJ.1 - Does the tool foster AWARENESS?*
Does the tool offer any functionalities that help users become (more) aware of their current status and behaviour, e.g. in comparison to other users or to their learning goals, etc.?

- QI is not applicable to this tool
- tool does not foster awareness
- tool fosters awareness

What type of awareness does the tool foster?*
- write "none" if no awareness is fostered

Please rate how difficult/easy it was to judge the quality indicator of awareness.*

	1	2	3	4	5	
very difficult	○	○	○	○	○	very easy

Figure 7 - Example of the questions for the evaluation.

We randomly chose eight learning analytics tools presented during the Learning Analytics and Knowledge conferences or created in research projects for the evaluation of the framework. The evaluation was done by eight members and associated partners of the LACE project. Each of the eight participants evaluated two of the eight tools, which in turn meant that each of the eight tools was evaluated twice.

Results

Table 1 shows how many *yes*, *no* and *not applicable* every quality indicator received. The highest scoring instance for each of the three is highlighted.

	yes	no	not applicable
Awareness	***15***	-	1
Reflection	12	2	2
Motivation	9	4	3
Behavioural change	14	1	1
Perceived usefulness	14	-	2
Recommendation	8	6	2
Activity classification	6	8	2
Detection of students at risk	12	3	1
Comparability	12	1	3
Effectiveness	9	1	6
Efficiency	4	4	***8***
Helpfulness	14	-	2
Transparency	9	5	2
Data standards	5	6	5
Data ownership	1	***10***	5
Privacy	9	2	5
Availability	7	3	6
Implementation	6	3	7
Training of stakeholders	7	1	***8***
Organisational change	8	5	3

Table 1 – Presence of quality indicators in a tool.

Figure 8 below visualises the same information as Table 1: blue is for *yes*, yellow for *no*, and red for *not applicable*.

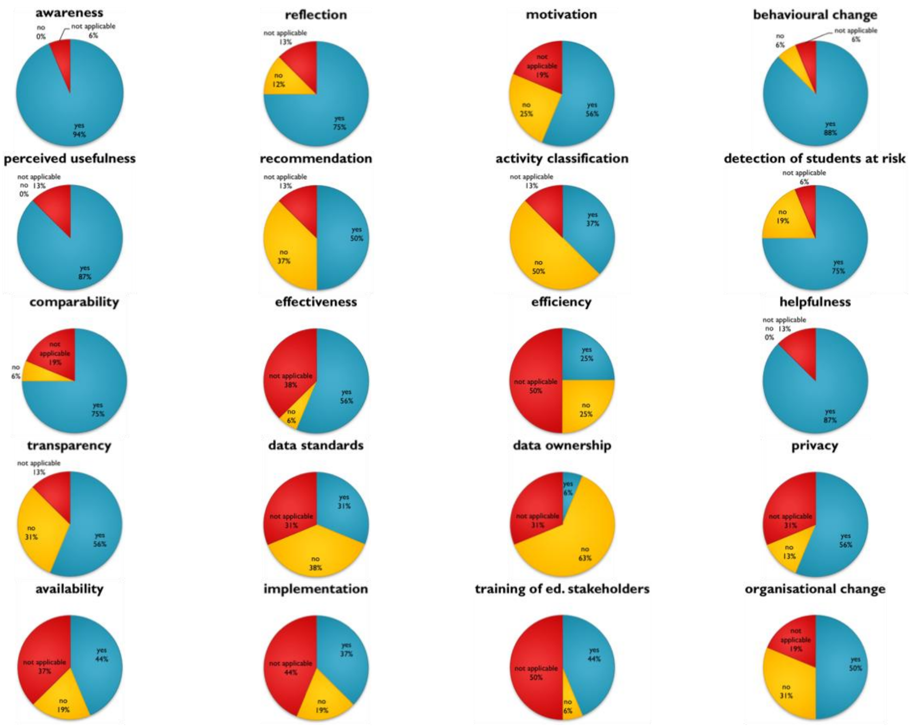


Figure 8 - Visualisation of data in Table 1.

Table 2 summarises the rating values of all quality indicators and also lists their average rating. The highest average rating is achieved by the quality indicator *awareness*, i.e. 4.3, while the lowest average is achieved by *efficiency*, i.e. 2.6. These two indicators are also the ones with the lowest (*awareness*) and highest (*efficiency*) non-applicability.

	1	2	3	4	5	avg.
Awareness	-	1	1	7	7	***4.3***
Reflection	1	-	2	6	7	4.1
Motivation	1	3	3	4	5	3.6
Behavioural change	-	3	5	5	3	3.5
Perceived usefulness	2	-	1	7	6	3.9
Recommendation	1	1	3	4	7	3.9
Activity classification	4	3	3	1	5	3.0
Detection of students at risk	-	1	3	6	6	4.1
Comparability	-	6	2	5	3	3.3
Effectiveness	2	5	4	4	1	2.8
Efficiency	4	3	5	3	1	***2.6***
Helpfulness	2	4	2	5	3	3.2
Transparency	-	4	6	5	1	3.2
Data standards	3	2	2	5	4	3.3
Data ownership	3	3	3	6	1	2.9
Privacy	-	3	2	8	3	3.7
Availability	2	1	1	3	9	4.0
Implementation	2	1	2	2	9	3.9
Training of stakeholders	2	-	1	8	5	3.9
Organisational change	2	-	1	12	1	3.6

Table 2 – Overview of all ratings plus the average rating for all quality indicators.

For the *Objectives* criterion we can see that the quality indicators are often present in or supported by a tool. The non-applicability is quite low. This in turn means that the quality indicators are applicable and thus quite suitable for the evaluation of learning analytics tools. Motivation is the most diverse/controversial one. Overall the study participants found this criterion rather easy to judge. Some of them mentioned that a there is a distinction between actually fostering something and having the intention to do so. They also remarked that there is a difference between direct and indirect fostering. Both suggestions should be taken into account when using the framework to evaluate learning analytics tools. Participants also suggested indicating the intended user of a tool, e.g. a learner or a teacher.

In the *Learning Support* criterion the non-applicability is also quite low. But this does not mean that the quality indicators are actually present in or supported by the tools as can be seen from the number of *nos*. Again participants stressed that the user type of a tool should be taken into account when evaluating and comparing learning analytics tools. They also pointed out that in this criterion there are several types of indicators: some are goals of a tool while some are functionalities. They suggest only having one type of indicator within a criterion. In some cases participants wished to have had a scaling between yes and no as the amount might make a difference (e.g. a tool might intervene too much which could be bad). Participants also suggested rephrasing or redefining some indicators in order to avoid confusion.

The *Learning Measures & Output* showed not too many *no* values. Either the indicators were present or not applicable. Overall participants found this criterion hard to judge and thought that the criterion was confusing. For some indicators they suggested a clear definition or a differentiation between indicators. They also pointed out again that the user type of a tool should be taken into account.

Of all criteria the one about *Data* Aspects had the most *no* values. Participants also pointed out that many aspects about data issues are hard to judge if one is not an actual user of a tool but has to rely on descriptions of a tool in publications. They therefore suggested introducing an “I don’t know”-option into the questionnaire. Again, they also stressed the importance of knowing the user type as well as clearer differentiation and definitions of the indicators.

Similar answers were given for the *Organisational Aspects* criterion. It had the highest non-applicability rate of all criteria. This was again due to the tool descriptions not offering enough information. As with many other criteria, participants asked for clarifying differentiations between indicators.

From these results we were able to identify several issues that needed to be addressed in order to have a working framework of quality indicators for learning analytics. We divided these issues into four categories:

1. concept definitions
    - rephrasing or defining criteria or indicators more clearly
2. differentiations
    - often needed between pairs of indicators
3. framework structure
    - inter-criterion homogeneity (goal vs. functionality)
    - indicators should tend to be concept-driven, not feature-driven
4. questionnaire adaption

- different questionnaires for different user types
- intention of tool
- answer options (I don't know)

Conclusion

In this paper we have shown that it is possible to develop an evaluation framework of quality indicators for learning analytics that builds on the experience and insight of those working in the field. The concept mapping method which we used creates a framework by combining bottom-up collection of statements from participants with a classification of those statements by experts in learning analytics. A different selection of experts, different analytical tools, and different judgements in clustering the statements could result in a framework which is, to some degree, different from that which we propose. Consequently the results constitute a first version of a framework, which will need to be refined in future work and the breadth of its applicability assessed.

We have taken the first steps towards refinement of the framework by carrying out an evaluation study, which identified problematic issues with the framework and collected suggestions about how to overcome those issues. This improved framework will be offered as a resource for the learning analytics community. It will form the basis of further studies within the LACE project and, we hope, beyond, with the aim of facilitating well founded comparisons between the tools and methods of learning analytics.

References

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M., 2014. Quality Indicators for Learning Analytics. *Educational Technology & Society, 17*(4), 117-132. Available at: http://www.ifets.info/issues.php?id=65

Scheffel, M., Drachsler, H., & Specht, M., 2015. Developing an evaluation framework of quality indicators for learning analytics. In *Proceedings of the Fifth International Conference on Learning Analytics and Knowledge – LAK '15*. New York, New York, USA: ACM Press, pp. 16-20. Available at: http://dl.acm.org/citation.cfm?id=2723629

About ...

Acknowledgements
The author would like to thank all participants of the two studies for their time and effort.

This document was produced with funding from the European Commission Seventh Framework Programme as part of the LACE Project, grant number 619424.

About the Authors

Maren Scheffel is a researcher and PhD candidate at the Welten Institute of the Open University of the Netherlands. She is a computational linguist and has been working in the field of technology enhanced learning (TEL) for several years where she was involved in several national and international TEL projects. Currently, she focuses her research on learning analytics, reflection and awareness support, personalisation and educational data. In LACE she is part of the project management and also involved in research about cross-domain studies.

About this document

The Learning Analytics Review is published by the LACE project at the University of Bolton, Bolton, UK. Please cite this document including the issue number and the ISSN of the Learning Analytics Review, which is 2057-7494.

The persistent URL for this document is: http://www.laceproject.eu/learning-analytics-review/learning-analytics-quality-indicators/

About LACE
The LACE project brings together existing key European players in the field of learning analytics & educational data mining who are committed to build communities of practice and share emerging best practice in order to make progress towards four objectives.

> *Objective 1 – Promote knowledge creation and exchange*
> *Objective 2 – Increase the evidence base*
> *Objective 3 – Contribute to the definition of future directions*
> *Objective 4 – Build consensus on interoperability and data sharing*

http://www.laceproject.eu *@laceproject*