Pre-print. Please cite-as follows:

Drachsler, H., & Kalz, M. (2016). The MOOC and learning analytics innovation cycle (MOLAC): a reflective summary of ongoing research and its challenges. Journal of Computer Assisted Learning, 32(3), 281-290. http://doi.org/10.1111/ical.12135

The MOOC and Learning Analytics Innovation Cycle (MOLAC): A reflective summary of ongoing research and its challenges

Hendrik Drachsler[*] & Marco Kalz[*†]

[*]Welten Institute—Research Centre for Learning, Teaching and Technology, Open University of the Netherlands

[†]Chair for Open Education, Faculty of Management, Science and Technology, Open University of the Netherlands

Correspondence: Hendrik Drachsler, Welten Institute, Open Universiteit, P.O. Box 29606401 DL, Heerlen, The Netherlands. Email: hendrik.drachsler@ou.nl

Introduction

The discussion about massive open online courses (MOOCs) and their role and impact is still intense and it is remarkable how much dispute is still ongoing about very basic questions.

The word *massive* refers to the large number of students who are simultaneously enrolled, from hundreds to thousands of participants. Massiveness is a challenge because of what David Wiley has called the teacher bandwidth problem (Wiley & Edwards, 2002), which is especially an issue in MOOCs if teaching is understood as more than lecturing. The common MOOC format (video, text, and forum discussions) is a scalable approach to communicate information, but for active knowledge building other support and feedback options are needed. Traditionally, teachers can only support a limited number of students with individual feedback. This problem has been solved by bringing relatively inexpensive teaching assistants into courses, but with enrolments in the hundreds of thousands there is no possibility to provide sufficient numbers of supporting staff simply due to limited resources and budget. Thus, in traditional education the handling costs per student increase when numbers of students that register increase as well. The Internet has proven to enable the opposite by significantly decreasing costs or change traditional businesses. As explained by Chris Anderson (2004), the Internet made it possible to offer much more products in cheaper ways like in the 'Amazon' example that revolutionised the traditional 'Wal-Mart' business model. Many felt that this same revolution could be replicated in education, erroneously assuming that education is little more than distributing products (Friedman, 2013).

MOOCs established the research field of *Open Online* education, they are in any case open in the sense of access (everybody with an internet connection can enrol) and in the sense of costs (no payment for access). Due to their digital distribution they are also open in terms of location and sometimes also in time. Most MOOCs are also open in the sense of formal entrance requirement; some are also openly sharing the course material as open educational resource. Therefore, they substantially take away many of the barriers that hinder learners from being involved in lifelong learning activities (Eurostat, 2012; Kalz, 2015) and are in that sense *open*. They MOOCs have quite some similarities with online courses offered by the Open Universities in the world.

And finally, MOOCs are *courses* and not just published resources. A course differs from open educational resources in the sense that there is an underlying teaching concept in a digital environment based on a coherent topic divided into subtopics and implemented based on a pedagogical theory or an instructional design approach.

While the question about the educational quality of current MOOC implementations is not answered yet (Kalz & Specht, 2013), MOOCs already have a direct impact on the potential for educational research. The massive number of participants offers a new playground for large-scale research interventions and also introduces new research questions and methods into the domain of technology-enhanced learning.

At the same time, there is an emerging market attracting venture capital, start-up companies and several platform providers. These platform providers claim to have reinvented distance education and they have if one considers the large number of participants MOOCs attract. But a large part of the public discussion about the impact of MOOCs is ill-defined and does not build on the large body of knowledge from the domain of distance education and technology-enhanced learning.

A good example of such an ill-defined problem is the discussion about the high numbers of dropouts in MOOCs. On the one hand, the dropout problem is something open universities have been dealing with since their beginnings in the 1960s. The contextual factors leading to dropout of learners are comparable, although the diversity of participants in the MOOC context is even higher. Learners in open education can vary a lot in terms of motivation and intention to enrol in a MOOC. In this sense, we cannot transfer the concept of dropout from formal (higher) education to the non-formal MOOC context. Often, participants who do not participate in all learning activities or who do not earn a certificate are regarded as students who have dropped out. We have proposed instead to define dropout in the MOOC context as the gap between initial intentions before the MOOC and realised intentions after the MOOC (Kalz et al., in press; Reich, 2014). Therefore we see the critique about dropout in MOOCs as ungrounded and underresearched.

MOOCs come in a wide variety of kinds and they differ in terms of underlying instructional design, support structures, tutoring approach and student assignments. To date, it has not been analysed systematically, which approaches are more suited for the teacher bandwidth problem than others.

So far online fora have been used to solve this issue. They act as the MOOC's social learning space and as the place to go and seek support from peers and offer support to peers. But fora easily become overwhelming, particularly if a single forum serves thousands of students. Although fora organise discussions in threads, they may become so numerous and make it difficult to separate relevant content from less relevant. This is particularly true if the forum also acts as the store of relevant content snippets. A recent study by Clow (2013) introduces the notion of a "funnel of participation" in MOOCs. The author presents data from several MOOCs that help to explain the high dropout rates for MOOCs. To put it differently, the hard work of organising the content and structuring the discussions that the professor and the teaching assistants do in offline learning environments is offloaded to the students themselves in a

MOOC. But in fora the sheer cognitive load of managing content and emerging discussions rapidly gets overwhelming and negatively affects the learning itself. What is needed are smart support tools to limit the number of discussants for the tutor and better match the questions to the potential answers. The massive number of participants calls for support services that are scalable without increasing the workload of tutors and lecturers. Those are not new research subjects, there has been quite a lot of research and developments conducted prior to the appearance of MOOCs that already address the problem of overwhelming information flows, personalise information, and smart ways to tailor information to specific groups (Manouselis, Drachsler, Verbert, Duval, 2012). Within Technology-Enhanced learning those issues have been addressed by the research communities Adaptive Hypermedia (Brusilovsky, 2007) and Learning Networks (Koper & Tattersall, 2004). Especially in the context of learning networks, a number of dedicated support and feedback services have been developed that dealt with exactly the same problems that MOOCs face these times. Examples of such services are tools that help learners to gauge their learning needs on the basis of their prior knowledge (Kalz, Van Bruggen, Rusman, Giesbers, & Koper, 2007), recommender systems that support students to discover content and learning activities (Drachsler, Hummel & Koper, 2009), or matching services that suggest ‘peer learners’ who are most likely to be able to be of assistance to a learner with a certain need (Van Rosmalen, Sloep, Brouns, Kester, Kone, & Koper, 2006). However, these existing interventions are rarely applied in MOOCs although MOOCs can provide even richer data for these services due to their enormous numbers of students. New research combining Learning Analytics and MOOCs are emerging as recently demonstrated at the Learning Analytics and Knowledge conference 2015 (Vogelsang & Ruppertz, 2015).

Learning Analytics in MOOCs

The first generation of MOOC research has been mainly dealing with case studies or educational theory behind MOOCs. In the literature review by Liyanagunawardena, Adams, and Williams (2013) the authors stress that the plethora of data generated in the MOOC context is widely underexploited. In fact, the massive amount of student data that are generated provide an unprecedented chance to study student behaviour and provide even better and more personalised support services to the students. In addition to the study from 2013, Sunar, Abdullah, White, and Davis (2015) provided an overview about personalisation tools and how they are applied within MOOCs so far. They claim that there are very few articles about the personalisation of MOOCs with data driven tools. They found a few tools that are applied on the course level such as: personalised learning path, personalised assessment and feedback, personalised forum thread and recommendation service for related learning materials or learning tasks.

Learning analytics (Greller & Drachsler, 2012) is currently the term that is used for studies aimed at understanding and supporting the behaviour - especially the study behaviours - of learners based on large datasets. Not long ago, gathering data was done using surveys or interviews with a selected representative number of students. The amount of data gathered was constrained by the cost, the time to collect them and worries about the scope and authenticity of the data. Learning in digital learning environments with very large numbers of participants has made data collection part and parcel of delivering educational content to the students. With the advent of learning analytics the mining of student data and their analysis no longer need to be limited to representative pilot studies, now the entire student population may be studied.

In this sense, research on MOOCs and research on learning analytics is naturally closely intertwined: MOOCs create the huge amounts of data that can feed the various learning analytics technologies. Surprisingly, most MOOCs still adopt an old-fashioned, top-down teaching approach, ignoring the potential for facilitating awareness, self-regulation, and personalisation. Reich (2015) argues that first-generation MOOC research had few implications for the change of teaching and learning practices. He pleads for the exploitation of big data sets with the purpose to advance MOOC research in three directions: (a) From engagement studies to studies about learning, (b) From research about individual courses to comparative research across courses and providers, and (c) From post-hoc analyses to experimental research design that provides hard evidences for learning science.

Learning analytics can provide different levels of insights as demanded by Reich (2015) either it is provided to a single course level or on a collection of MOOCs. It can contribute to a second-generation of MOOC research that provides additional insights into effects of different learning designs and other educational interventions supported with a high numbers of participants. Buckingham Shum (2012) thus introduced the notion of micro, meso, and macro levels (see Figure 1) to distinguish the role that learning analytics can play on different clusters of MOOC data. The micro level mainly addresses the needs of teachers and students and aims at a single course, the meso level addresses a collection of MOOCs or structured within a curriculum and provides information for course managers, the macro level, takes a bird's-eye view on a directory of MOOCs and can provide insights for a whole community by monitoring learning behaviour in MOOCs even from different scientific disciplines. Depending on which level the learning analytics takes place different objectives and information are of relevance and can be monitored.

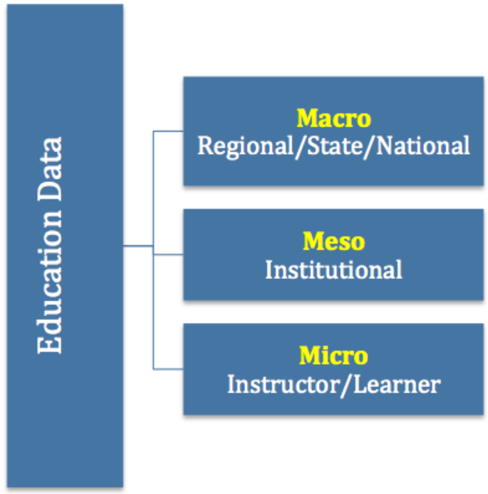

Figure 1: Different levels of learning analytics.

A comprehensive introduction to the different domains that are affected by learning analytics has been provided by Greller and Drachsler (2012). They presented the technological and educational aspects of learning analytics in six dimensions including the following perspectives (see Figure 2).

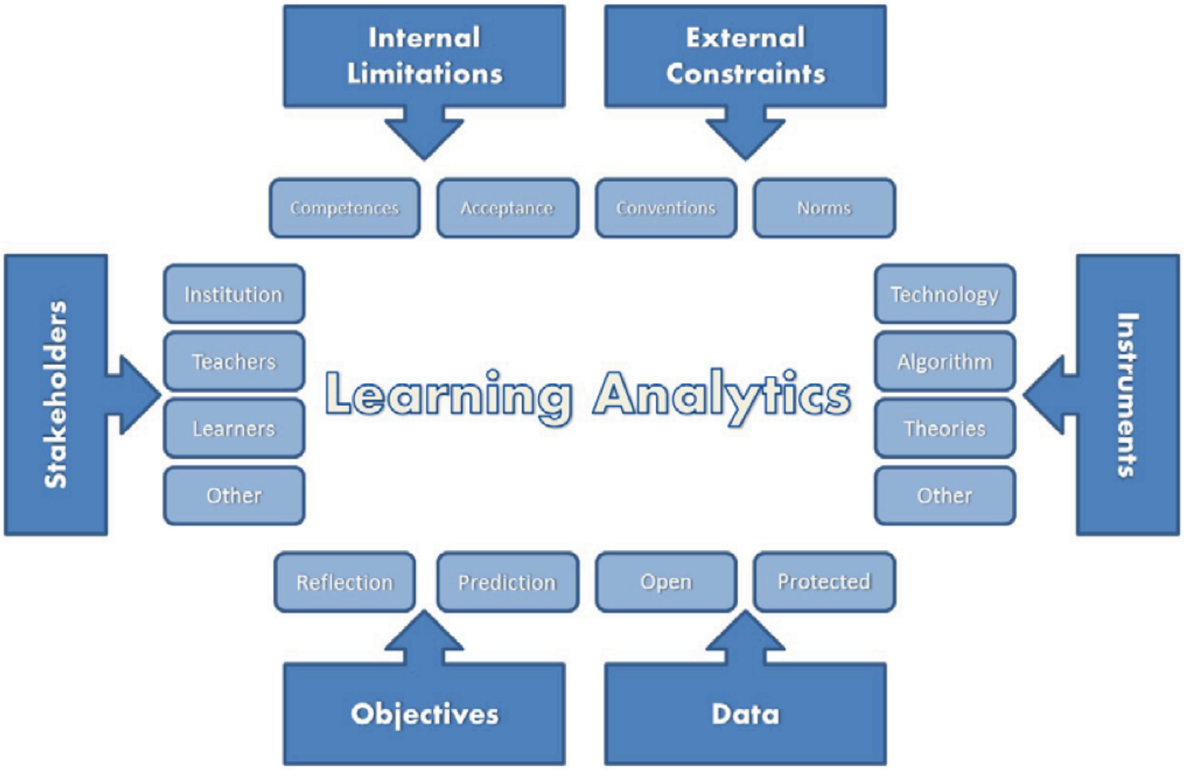

Figure 2: Learning analytics Framework by Greller and Drachsler (2012)

In the following description, we applied the six dimensions of learning analytics as defined by Greller and Drachsler (2012) especially to MOOCs.

Stakeholders: The contributors and beneficiaries of learning analytics.

The stakeholder dimension includes *data clients* as well as *data subjects*. Data clients are the beneficiaries of the learning analytics process who are entitled and meant to act upon the outcome (e.g., MOOC students & teachers). Conversely, the data subjects are the suppliers of data, normally through their browsing and interaction behaviour (e.g., MOOC participants). Those roles can change depending on the objective of the analytics on the meso and macro level.

Objectives: Set goals that learning analytics wants to achieve.

The main opportunities for learning analytics as a domain are to unveil and contextualise so far hidden information out of the educational data and prepare it for the different stakeholders. This new kind of information can support individual learning or teaching processes on the micro level. Here we mainly talk about supporting reflection and predictions as main objectives. But on the meso or macro level the objectives change and become more organisational knowledge management with the focus on benchmarking of pedagogical approaches and interventions.

Data: The educational datasets and the environment in which they occur.

Learning analytics uses datasets from different educational systems. Most of the data produced in institutions is protected. Nevertheless, to advance research for MOOCs on the meso- or macro level as demanded by Reich (2015), it would be supportive to have a metadata standard that allows to combine and compare the data that is collected in different scientific disciplines. Those metadata standards are emerging with the appearance of xAPI[1] and IMS Caliper[2].

Method: Technologies, algorithms, and theories that support and underpin learning analytics.

Different technologies can be applied in the development of educational services and applications that support the objectives of the different educational stakeholders. Learning analytics takes advantage of so-called information retrieval technologies like educational data mining (EDM), machine learning, or classical statistical analysis techniques in combination with visualization techniques. The output of those technologies changes depending on the level they

[1] http://tincanapi.com/

[2] http://imsglobal.org/caliper/index.html

are applied. In general one could think about specific technical requirements for MOOCs and ways to present this information in a general dashboard to the different stakeholders of a MOOC.

Constraints: Restrictions or potential limitations for anticipated benefits.

The large-scale production, collection, aggregation, and processing of information from MOOCs have led to ethical and privacy concerns regarding potential harm to individuals and society. Until now, there have been few papers published relating to ethics and privacy. But first policies and guidelines regarding privacy, legal protection rights and ethical implications are announced like the recent policy published by the Open University UK[3]. In a recent article, Prinsloo and Slade (2015) investigated the data usage conditions of various MOOC platforms towards ethical and privacy approaches. The authors identified that MOOC participants have very little to no control about their user data collected by the MOOC environment.

Competences: User requirements to exploit the benefits.

In order to make learning analytics an effective tool for MOOCs, it is important to recognise that learning analytics do not end with the presentation of algorithmically attained results. Those results need interpretation by the MOOC stakeholders, hence the exploitation of learning analytics requires some high-level competencies, such as interpretative and critical evaluation skills. Those skills are to date not a standard competence for the stakeholders.

The MOOC Learning Analytics Innovation Cycle (MOLAC)

To bring the different domains, objectives, levels of analysis and processes for learning analytics and MOOCs into a joint picture we have developed the *MOOC Learning Analytics Innovation Cycle* (MOLAC). The cycle works on three different levels. On the micro level, data from a single course is collected to foster predictions and reflection for individual learners or

[3] http://www.open.ac.uk/students/charter/essential-documents/ethical-use-student-data-learning-analytics-policy

teachers. On the meso level, educational institutions combine several MOOCs and enable the sharing and analysis of data beyond a single course via metadata standards. The combined data from different MOOCs can be used for classification of learners and contributes to the heavily debated notion of learner types and learning styles in a more informed and data driven approach.

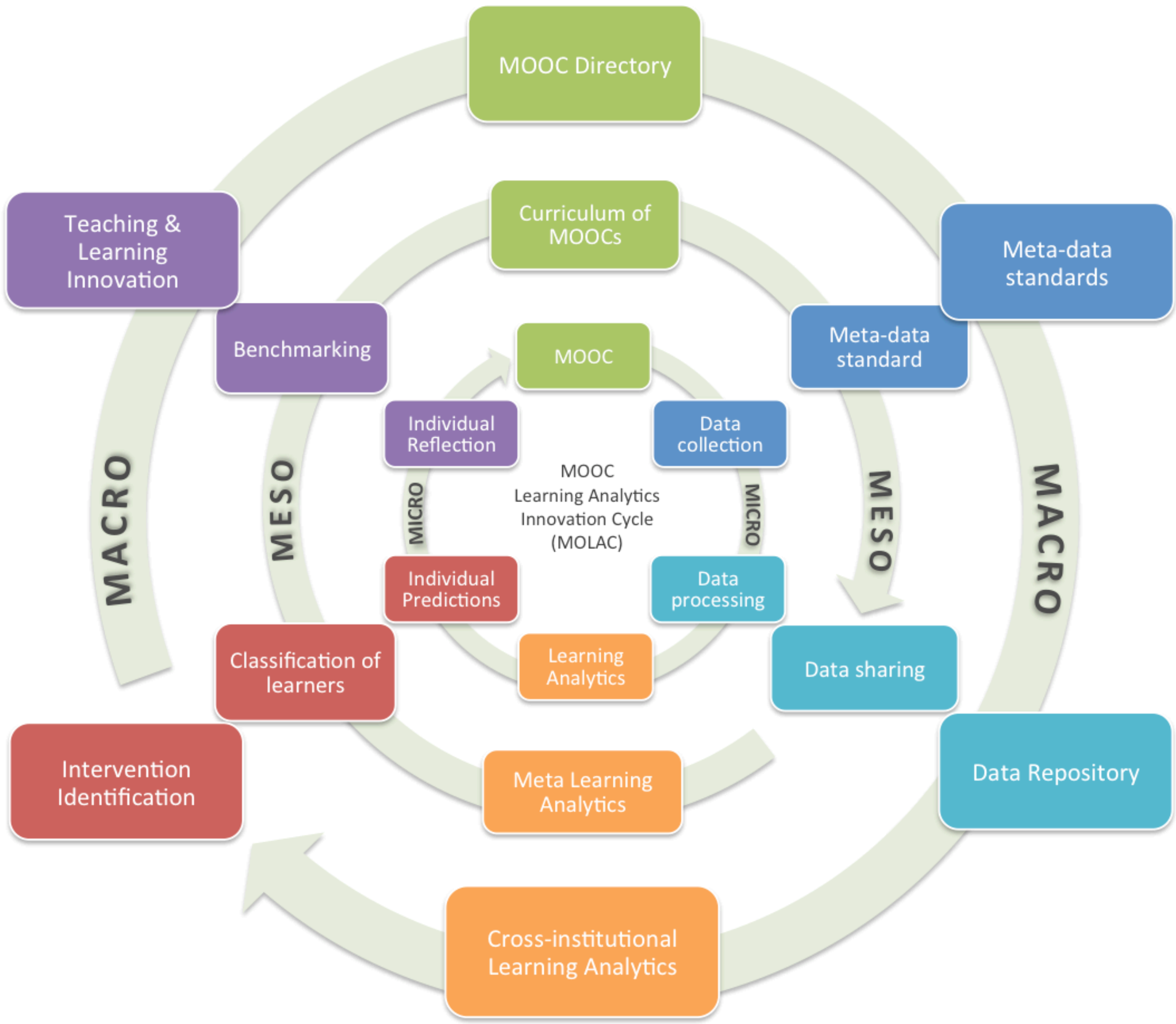

Figure 3: The MOOC Learning Analytics Innovation Cycle - MOLAC.

On the macro level, the analysis is conducted across MOOC providers and curricula and data is shared between providers via a data-repository. This type of cross-institutional learning analytics targets the identification of interventions that contribute to the innovation of learning and teaching for the individual institution but also for a wider group of stakeholders like the learning science community at large.

In this sense, the combination of MOOCs and learning analytics provide an innovation environment for educational institutions that allows the testing of interventions and new concepts outside of the current educational system of the institution. Most initiatives start at the micro level but ideally also recognise the potential to generalise the research interest to a higher level by conducting cross-institutional initiatives. The European MOOCKnowledge initiative is an example of cross-institutional cooperation (Kalz et al, in press) with the goal to inform institutions about strategic value of their current open education strategies and to inform (European) policy-makers with regard to socio-economic impact of open education but also barriers for institutions to make the European Higher Education system more accessible and flexible.

In the following section, we discuss the articles in this special issue in relation to the MOLAC Innovation Cycle. Almost all of the articles report on studies that were conducted on the micro level of the Cycle; however, in two of the articles there is clear potential to generalise the evidence found at the micro level to higher levels of investigation. Considering those developments, we are confident that in a rather short timeframe (2 to 3 years) a richer body of evidences will be available for the meso and macro level of the MOLAC Innovation Cycle as demanded by Reich (2015).

Applying MOLAC to the Special Issue Articles

The article by Alario-Hoyos, Muñoz-Merino, Pérez-Sanagustín, Delgado Kloos, and Parada G. (2015), entitled “Who are the top contributors in a MOOC? Relating participants’ performance and contributions” is situated at the micro level of the MOLAC framework. Learners with the potential to act as co-facilitators within an open course are identified based on the contributions in five selected social tools. Based on an analysis of activities in these tools, the study reports a moderate positive correlation between the number of posts submitted to the five social tools and the overall performance of participants. The authors argue that top contributors can play a special role in the MOOC, for example through partially taking over tasks traditionally done by the teacher. The identification of the top contributors is thus seen as a research challenge in which the authors employ methods from social learning analytics. This research clearly deals with the teacher bandwidth problem and follows a technological approach to replace tutor capabilities with other experts in the course. Earlier research has shown that one-dimensional approaches to tutor selection are problematic (Van Rosmalen et al., 2008) and that more factors need to be taken into account than activity patterns within a course. Results of the contribution show that activities in social tools are not a perfect predictor for knowledge since the top-contributors are not always the top-performers. The authors propose additional factors that need to be taken into account in future research. While the use of peers for knowledge building and support is a very interesting research direction in MOOCs to solve the teacher bandwidth problem, it is also important to keep the learner-bandwidth in mind. Therefore, peer-tutor selection tools should also take into account that the ideal peers should also not be overloaded through too many requests, otherwise the learning experience by these learners will be disregarded (Van Rosmalen et al., 2006). For future work, not only the potential to act as peer-supporters should be taken into account but also psycho-social dispositions including their motivation, willingness and intentions and their actual behaviour. To scale the approach up, data needs to be collected across MOOCS, ideally with a diverse audience to allow cross-comparison.

The study conducted by De Barba, Kennedy, and Ainley (2015) and described in their article "The role of students' motivation and participation in predicting performance in a MOOC" was also carried out at the micro level of the MOLAC framework. The research specifically focuses on intrinsic motivation and the influence of the factors individual interest, mastery-approach goals, and utility value beliefs. The goal of the study was to predict achievement by motivation and persistence in a MOOC. The authors introduce a research model for the study with three components: (a) motivation, (b) participation, and (c) performance. The authors make a distinction between initial motivation and maintained motivation and use levels of participation from video hits and quiz attempts as indicators. The study confirms that motivation has a direct and indirect influence on performance and it states that the number of quiz attempts is the strongest predictor for performance. Here it is very likely that test taking is not only a predictor for performance but that repeated testing has also a positive impact on knowledge gain for learners. In future research, this "testing effect" (Van Gog & Kester, 2012; Dirkx, Kester & Kirschner, 2014) should be analysed more thoroughly in the context of MOOCs. To scale up this research, a larger and more diverse sample is needed to permit the drawing of valid conclusions that go beyond the unit of a single MOOC. In addition, the complexity of motivation poses a challenge to identify the fine-grained differences that can influence participants to enrol in a MOOC and to also be persistent. The model proposed by the authors does not take into account individual factors (e.g., socio-economic status, skills), environmental factors (e.g., support by family or job context) and also no other types of motivation (e.g., extrinsic motivation). While it is understandable that the complexity of the research model is initially limited, in the future these other factors need to be taken into account. The authors could follow-up on their study by taking into account the application of the theory of reasoned action (Ajzen, 2011) in combination with the locus of causality (Ryan & Connell, 1989) to systematically analyse participant behaviour and goal achievement in a cross-provider data collection.

The exploratory study by Goggins, Galyen, and Laffey (2015), the authors of "Connecting performance to social structure and pedagogy as a pathway to scaling learning analytics in MOOCs: An exploratory study," *uses a data set that* is selected out of a data pool from many years that can be allocated to the meso level. The explorative study is based on a single course from the datasets, which brings the study back to the micro level. The study focuses on the design and evaluation of teaching analytics that relate social learning structure with performance measures in a massive open online course (MOOC) environment. The authors apply a post-hoc analysis of online learning trace data and qualitative performance measures for their study with three main outcomes: (a) Evaluates a novel, multi-dimensional performance construct, (b) Describes differences in small group dynamics and structure, and (c) Draws a connection between learning performance and group structure. Performance is operationalised using a combination of knowledge construction measurement from discussion boards, analysis of student work products and several indicators of small group identities. Interviews and observational data are used to develop an approach for deriving and validating a model of the social structure of students in the course using traces of interaction data. Implications for MOOC design, scaling MOOC analytics and a vision for developing social sensors in MOOC environments are presented that could contribute to the general technical requirements for a MOOC dashboard as mentioned in the learning analytics framework (Greller & Drachsler 2012).

The article by Pursel, Zhang, Jablokow, Choi, and Velogel (2015), "Understanding MOOC Students: Motivations and Behaviors Indicative of MOOC Completion," also addresses the micro level of MOOC research but has high potential to be applied also on the meso level. The authors contribute to the previously mentioned ill-defined dropout definition (Kalz et al., in press). The authors examined MOOC student demographic data, intended behaviours, and course interactions to better understand variables that are indicative of MOOC completion. The results of this study provide early insights into several variables, such as prior degree attainment and course interaction data that show some relationship to MOOC completion. Those variables contribute to better defining and predicting the chances for dropout. MOOC design teams can attempt to predict student completion, and devise methods to keep students engaged in the MOOC. The authors provided interesting results to better define the MOOC completion problem. The results can be further extended and brought to the meso level of MOLAC by investigating the same variables in other MOOCs. It is a promising approach to advance the body of knowledge about learner behaviour and MOOC completion rates in specific.

The article by Baker, Clarke-Midura, and Ocumpaugh (2015), "Towards general models of effective science inquiry in virtual performance assessments" is not a MOOC study in the former sense, but nevertheless it provides an interesting example how insights can be learned at the micro level and applied to the meso level later on. In the study the authors developed a model that assesses student inquiry in a virtual environment where students talk to other avatars, collecting samples, and conducting scientific tests with those samples in the virtual laboratory. They analysed log file data from nearly 2000 middle-school students using the Virtual Performance Assessment (VPA) tool to develop models of student interaction within VPA that predict whether a student will successfully conduct scientific inquiry. The authors identify behaviours that lead to distinguishing *causal* from *non-causal* factors to identify a correct final conclusion. The authors demonstrate then that those models can be adapted with minimal effort and applied to new VPA scenarios. The authors are positive about generalising their findings and use the models as a tool to better understand scientific inquiry competences and how to assesses those also in other courses.

The article by Rayyan et al. (2015), “A MOOC based on blended pedagogy” is the only article in the special issue that can be said to pertain to the meso level of the MOLAC cycle. It describes three versions of a MOOC on Introductory Physics that have been given to different target groups at Massachusetts Institute of Technology (MIT). The same MOOC was offered to a general audience, specific MOOC targeting teachers, and a large-scale MOOC on the edX.org platform. This approach allowed the authors to compare and contrast the same MOOC as given to three different target groups. Their results are therefore a suitable example how research can be conducted on the meso level of the MOLAC Innovation Cycle. The authors provided evidence of the effect of certain course designs on student behaviour and showed how modifications like reducing the class size and posting materials well in advance resulted in higher retention. The study therefore also contributes to addressing the dropout problem. By applying Item Response Theory to common homework problems, Rayann et al. were able to show that the MOOC participants had significantly higher ability than students in a traditional MIT course and that they maintained this advantage over the duration of the MOOC.

Conclusion

We believe that these high quality articles are promising approaches towards more advanced and data-driven research in MOOCs as demanded by Reich (2015). We hope that the MOLAC Innovation Cycle provides an inspiring vision to advance the body of knowledge about learning facilitated with data from MOOCs. In order to apply MOLAC in its full potential a number of aspects need to be established to enable meta-analysis of MOOC data . In the remaining part of the article we want to set out this future research agenda consisting of:

a.) A standardised way of describing the educational design of a MOOC

One could also expect that there are always clear learning objectives formulated, but a recent study by Margaryan, Bianco, and Littlejohn (2015) has shown that this is not always the case. The traditional categories like xMOOC and cMOOC are less supportive in order to collect evidences and data for specific educational designs. Without a clear approach of describing the educational design applied to a MOOC, a meta-analysis and impact evaluation of different instructional approaches is hardly possible due to the course diversity. By cross comparing data from MOOCs with the same educational design approach, we expect to gain more evidence about effective educational designs and potential.

b.) Data sharing facilities

In order to develop towards shareable data sets for specific instructional designs metadata standards such as IMS Caliper or xAPI need to be applied to store the online behavior data of a MOOC. In a second step, the community of MOOC researcher is in need of a data platform to make those data sets available online. Such data sharing platforms have been established already in other educational contexts as successfully shown by the LinkedUp (d'Acquin et al., 2014) and DataShop project (Stamper & Koedinger, 2011).

c.) Policy making and ethical guidelines.

Data sharing also raises issues like data anonymisation, data ownership and the need to have ethical and privacy guidelines in place that allow to use the data of the MOOC participants to conduct further research with it. The MOOC providers therefore need to adjust their terms-of-content descriptions (Prinsloo, Slade, 2015) that allow the further use of the data for research purposes or enable the students to opt-out out of a data sample after the MOOC has been completed.

d.) Standardised evaluation approaches

Finally we have to work towards an standardised evaluation framework that supports the MOOC researcher to compare the effects of the different MOOCs in a standardised and comparable way. Promising work has been done by Drachsler et al., (2014) who have develop a first evaluation framework for educational tools that have been submitted to various data competitions (Drachsler et al., 2014). The identified variables and indicators of the evaluation framework can be a first promising step towards a standardised approach for the evaluation of MOOCs in the future.

References

Ajzen, I. (2011). The theory of planned behaviour: reactions and reflections. *Psychology & Health*, *26*, 1113-1127.

Alario-Hoyos, C., Muñoz-Merino, P. J., Pérez-Sanagustín, M., Delgado Kloos, C., & Parada G., H. A. (2015). Who are the top contributors in a MOOC? Relating participants' performance and contributions. *Journal of Computer Assisted Learning, 31*.

Anderson, C. (2004). The Long Tail. [Electronic Version]. Wired Magazine. Retrieved 25.05.09 from http://www.wired.com/wired/archive/12.10/tail.html

Baker, Clarke-Midura, and Ocumpaugh (2015). Who are the top contributors in a MOOC? Relating participants' performance and contributions. *Journal of Computer Assisted Learning, 31*.

Brusilovsky, P. (2007). The Adaptive Web (Lecture Notes in Computer Science ed. Vol. 4321). Berlin Heidelberg: Springer

Buckingham Shum, S. (2012, November). *Learning analytics* (UNESCO IITE policy brief). Retrieved from UNESCO Institute for Information Technologies in Education website: http://iite.unesco.org/pics/publications/en/files/3214711.pdf

Clow, D. (2013). MOOCs and the funnel of participation. In D. Suthers, K. Verbert, E. Duval, & X. Ochoa (Eds.), *Proceedings of the Third International Conference on Learning Analytics and Knowledge (LAK' 13)* (pp. 185–189). New York, NY: Association for Computing Machinery.

d'Aquin, M., Dietze, S., Herder, E., Drachsler, H., & Taibi, D. (2014). Using Linked Data in Learning Analytics. eLearning Papers, 36. Online available at: http://www.openeducationeuropa.eu/en/download/file/fid/33993

De Barba, P. G., Kennedy, G. E., & Ainley, M. D. (2015). The role of students' motivation and participation in predicting performance in a MOOC. *Journal of Computer Assisted Learning, 31*.

Dirkx, K., Kester, L. & Kirschner, P. (2014). The testing effect for learning principles and procedures from texts. The Journal of Educational Research. Vol. 107, Iss. 5, 2014.

Drachsler, H., Hummel, H. G. K., & Koper, R. (2009). Identifying the goal, user model and conditions of recommender systems for formal and informal learning. *Journal of Digital Information, 10*(2). Retrieved from https://journals.tdl.org/jodi

Drachsler, H., Stoyanov, S., d'Aquin, M., Herder, E., Guy, M., & Dietze, S. (2014, 16-19 September). An Evaluation Framework for Data Competitions in TEL. In C. Rensing, S. de Freitas, T. Ley, & P. J. Munoz-Merino (Eds.), Proceedings of the 8th European Conference on Technology Enhanced Learning, LNCS: Vol. 8719 (pp. 70-83). Graz, Austria.

Eurostat. (2012). Lifelong learning statistics. Retrieved from http://ec.europa.eu/eurostat/statistics-explained/index.php/Lifelong_learning_statistics

Friedman, T. L. (2013, January 26). Revolution hits the universities. *The New York Times*, p. SR1.

Goggins, S. P., Galyen, K. D., & Laffey, J. M. (2015). Connecting performance to social structure and pedagogy as a pathway to scaling learning analytics in MOOCs: An exploratory study. *Journal of Computer Assisted Learning, 31*.

Greller, W., & Drachsler, H. (2012). Translating learning into numbers: A generic framework for learning analytics. *Educational Technology & Society, 15*(3), 42–57.

Kalz, M., Kreijns, K., Walhout, J., Castaño-Munoz, J., Espasa, A., & Tovar, E. (in press). Establishing a European cross-provider data collection about open online courses. *The International Review of Research in Open and Distributed Learning*.

Kalz, M. (2015). Lifelong learning and its support with new technologies. In J. D. Wright (Ed.), *International Encyclopedia of the Social & Behavioral Sciences*: Vol. 14 (2nd ed., pp. 93–99). Oxford, UK: Elsevier.

Kalz, M., & Specht, M. (2013). *If MOOCs are the answer, did we ask the right questions? Implications for the design of large-scale online courses* (Working paper 2013/25). Maastricht, Netherlands: Maastricht School of Management.

Kalz, M., Van Bruggen, J., Rusman, E., Giesbers, B., & Koper, R. (2007). Positioning of learners in learning networks with content, metadata and ontologies. *Interactive Learning Environments, 15*, 191–200.

Koper, R., & Tattersall, C. (2004). New directions for lifelong learning using network technologies. *British Journal of Educational Technology, 35*, 689–700.

Liyanagunawardena, T., Adams, A. A., & Williams, S. A. (2013). MOOCs: A systematic study of the published literature 2008–2012. *The International Review of Research in Open and Distance Learning*, *14*(3), 202–227.

Manouselis, N., Drachsler, H., Verbert, K., & Duval, E. (Eds.) (2012). Recommender Systems for Learning. Berlin, Springer, 2012, 90 p.

Margaryan, A., Bianco, M., & Littlejohn, A. (2015). Instructional quality of Massive Open Online Courses (MOOCs). *Computers & Education, 80*, 77–83.

Prinsloo, P., & Slade, S. (2015). Student privacy self-management: implications for learning analytics. In J. Baron & G. Lynch (Eds.), *Proceedings of the Fifth International*

Conference on Learning Analytics and Knowledge (LAK'15) (pp. 83–92), New York, NY: Association for Computing Machinery.

Pursel, B. K., Zhang, L., Jablokow, K. W., Choi, G. W., & Velegol, D. (2015). Understanding MOOC students: Motivations and behaviors indicative of MOOC completion. *Journal of Computer Assisted Learning, 31*.

Rayyan, S., Fredericks, C., Colvin, K., Liu, A., Teodorescu, R., Barrantes, A., ... Pritchard, D. E. (2015). A MOOC based on blended pedagogy. *Journal of Computer Assisted Learning, 31*.

Reich, J. (2014, December 8). MOOC completion and retention in the context of student intent. *EDUCAUSE Review*. Retrieved from http://er.educause.edu/articles/2014/12/mooc-completion-and-retention-in-the-context-of-student-intent

Reich, J. (2015). Rebooting MOOC research. *Science, 347*(6217), 34–35.

Ryan, R. M., & Connell, J. P. (1989). Perceived locus of causality and internalization: Examining reasons for acting in two domains. *Journal of Personality and Social Psychology, 57*, 749–761.

Stamper, J. C., & Koedinger, K. R. (2011, January). Human-machine student model discovery and improvement using DataShop. In *Artificial Intelligence in Education* (pp. 353-360). Springer Berlin Heidelberg.

Sunar, A. S., Abdullah, N. A., White, S., & Davis, H. C. (2015). Personalisation of MOOCs—the state of the art. In M. Helfert, M. T. Restivo, S. Zvacek, & J. Uhomoibhi (Eds.), *Proceedings of the Seventh International Conference on Computer Supported Education (CSEDU 2015)* (pp. 88–97). Setúbal, Portugal: Science and Technology Publications.

Van Gog, T., & Kester, L. (2012). A test of the testing effect: Acquiring problem-solving skills from worked examples. *Cognitive Science*, *36*, 1532–1541.

Van Rosmalen, P., Sloep, P., Brouns, F., Kester, L., Koné, M., & Koper, R. (2006). Knowledge matchmaking in learning networks: Alleviating the tutor load by mutually connecting learning network users. *British Journal of Educational Technology, 37*, 881–895.

Van Rosmalen, P., Sloep, P., Kester, L., Brouns, F., De Croock, M., Pannekeet, K., & Koper, R. (2008). A learner support model based on peer tutor selection. *Journal of Computer Assisted Learning*, *24*(1), 74-86.

Vogelsang, T. & Ruppertz, L. (2015). On the validity of peer grading and a cloud teaching assistant system. In *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge* (LAK '15). ACM, New York, NY, USA, 41-50. DOI=10.1145/2723576.2723633

Wiley, D. A., & Edwards, E. K. (2002). Online self-organizing social systems: The decentralized future of online learning. *Quarterly Review of Distance Education*, *3*, 33–46.