Technology, Knowledge and Learning (2019) 24:599–619
https://doi.org/10.1007/s10758-019-09399-5

ORIGINAL RESEARCH

A Pedagogical Perspective on Big Data and Learning Analytics: A Conceptual Model for Digital Learning Support

Sabine Seufert[1] · Christoph Meier[1] · Matthias Soellner[1] · Roman Rietsche[1]

Published online: 1 March 2019

Abstract
The increasing prevalence of learner-centred forms of learning as well as an increase in the number of learners actively participating on a wide range of digital platforms and devices give rise to an ever-increasing stream of learning data. Learning analytics (LA) can enable learners, teachers, and their institutions to better understand and predict learning and performance. However, the pedagogical perspective and matters of learning design have been underrepresented in research thus far. In our paper, we propose a general design framework that includes critical dimensions of LA and assists in creating LA services that support educational practice. On the basis of a two-dimensional framework (individual vs. social, reflection vs. prediction), we then identify four generic approaches to LA aimed at improving learning process and learning outcomes. To demonstrate the application, four use cases are outlined that are based on four previously elaborated generic approaches to LA. Finally, we discuss the validation of the model and close with an outlook on relevant future research.

Keywords Learning analytics · Social learning analytics · Digital learning support · Learning analytics taxonomy

1 Introduction

Currently, big data and analytics are burgeoning fields of research and development (Abdous et al. 2012; Ali et al. 2012; Dyckhoff et al. 2012). In education, several concurrent developments are taking place that have implications for big data and analytics in the field of learning. A wide range of promises and anxieties about the coming era of big data and learning analytics (LA) are in debate (Cope and Kalantzis 2016; Ifenthaler 2015; Ifenthaler et al. 2014). Overall, there is widespread consensus that the educational landscape itself is in transition and the changes are substantial, with expository instructional methods being replaced by more learner-centred approaches to learning. As more and more learning is either taking place online or is supported through technology, these active learners produce

✉ Sabine Seufert
sabine.seufert@unisg.ch

[1] University of St. Gallen, St. Gallen, Switzerland

an ever increasing stream of data—both inside learning management systems (LMS) and outside, in other IT-based environments (Pardo and Kloos 2011).

Learning analytics refers to the use of ”dynamic information about learners and learning environments to assess, elicit, and analyze them for modeling, prediction, and optimization of learning processes” (Mah 2016, p. 288). As Roberts et al. (2017, p. 317) states, the pedagogical potential to provide students “with some level of control over learning analytics as a means to increasing self-regulated learning and academic achievement”. Visualisation of information, social network analysis, and educational data mining techniques are at the methodological core of this newly emerging field (Greller and Drachsler 2012). Techniques for analyzing big data are such as machine learning and natural language processing based on the particular characteristics of these data for learner and teacher feedback, the possibility of real-time governance, and educational research (Cope and Kalantzis 2016, p. 2).

While this field is multi- or even interdisciplinary, the pedagogical perspective appears to be somewhat underrepresented (Greller and Drachsler 2012). Current research on big data in education revolves largely around (1) the potential of learning analytics to increase the efficiency and effectiveness of educational processes and (2) the ability to identify and support students at risk and to thereby reduce drop out-rates. Accordingly, the main problem is that the core focus of research is on *prediction*, while the potential for supporting *reflection* on processes of learning is being neglected. Therefore, the main purpose of this paper is to map out how LA can be carried out from a pedagogical perspective and to conceptualize a generic framework for the design of LA environments.

2 Research Questions and Methodology

In line with Kelly et al. (2015), the claim that we put forth in this paper is that “theory-led design has the potential to yield innovation in the development of LA tools and, in turn, that the development of LA tools and their use may contribute to learning theory” (p. 15). Our paper presents a framework for the theory-led design of LA environments with particular focus on digital learner support and students’ cognition.

The key research question we pursue in this paper is the following:

How can big data and learning analytics be employed in order to improve learner guidance, students’ learning processes and learning outcomes with regard to meta-cognitive abilities for self-regulated learning?

We pursue these issues by asking a range of more detailed questions:

- What are critical dimensions/aspects when designing LA services that are integrated in a pedagogic process? And what would a generic framework for designing such LA services need to look like?
- What generic strategies for developing LA services currently exist? And what form would the concept and set-up of a decision-support framework for devising LA strategies need to take?
- Which skills are required by learners in their roles as data subjects and/or data clients in order to make competent use of LA services?

The research project we report on here was based on a methodological combination of systematic literature analysis and model development.

The goal of the proposed framework is to provide relevant stakeholders—in particular designers and teachers of learning environments—with decision guidelines from a pedagogical perspective. In order to obtain an overview of existing LA research, an initial systematic literature analysis was conducted. The focus of this analysis was on work that addresses basic conceptualisations of LA, reference models for LA, and methods applied in order to pursue LA. Building on the findings of this literature review, and by combining and expanding or extrapolating existing models, the generic framework for designing LA was created.

Our starting point was the framework provided by Greller and Drachsler (2012). This pedagogical model contains six dimensions: competences, constraints, method, objectives (distinguishing between reflection and prediction), data, and stakeholders. On this basis, we have proposed a design framework for a more holistic approach to learning analytics rooted in a pedagogical perspective and focusing on students' cognition resulting in four generic LA approaches we discussed and elaborated with stakeholders at our university.

For that reason, we conducted a needs analysis (e.g. in terms of relevant competences, data issues, etc.) at our university with 12 lecturers (diverse group, large-scale and small group lectures, different subjects, at least five years teaching experience, four lecturers with a programme manager role; all of them have experience with LA at least in one of the developed generic approaches).

We discussed the developed use cases and received feedback on the needs of important implementation factors. These interviews were helpful in order to (1) provide an understanding of the current state of the learning analytics field and (2) assist in identifying teachers for setting up an internal task force.

In the process, we applied cognitive mapping techniques with the programme managers and lecturers participating in the task force (Ackermann et al. 2004). We used cognitive mapping as a communication tool between the analysts and the users for adapting the initial framework. Furthermore, we used cognitive mapping to decompose the model into finer detail by using elements of additional frameworks. We structured the use cases according to Greller and Drachsler (2012), and emphasized the learning objectives as well as skills required by learners as a core element for the competent use of LA applications.

3 Results

3.1 Literature Analysis

This study reviews literature selected with the primary focus on big data and learning analytics and their implications on higher education, educational technology, and instructional design. Google Scholar was used to search and locate academic papers from journals, conference proceedings, and professional magazines with the keywords "big data" and/or "learning analytics" and "framework" or "concept" or "model" or "applications" or "approaches". The search period was set from 2010 to 2017 and the papers reviewed include both qualitative and quantitative studies from researchers in the field of learning analytics worldwide. For the purpose of this study, the data collection process resulted in the identification of 45 articles. Ten of the articles provided frameworks that were too narrow, e.g., general principles or policy frameworks for the ethical use of data. Therefore, 35 articles were further analyzed and compared. The frequency with which these articles are cited by researchers bears witness to their relevancy and to the fact that they are a representative sample of the literature in the field. In addition to this search for original contributions, we conducted a literature analysis to identify current literature reviews on Learning Analytics. Of primay importance are the reviews by Papamitsiou and Economides (2014) who identified 40 articles; Sin and Loganathan (2015) who identified 45 articles; and Leitner et al. (2017), who identified 101 papers on learning analytics.

Starting from this body of research, the selection criteria for the overview presented in Table 1 were the following:

1. Holistic frameworks that describe or develop LA systems (e.g., static models vs. dynamic process models);
2. Generic approaches to a partial theory of LA with a focus on LA objectives and students' competences as this is our research focus.

   The analysis of the contributions in the body of research identified resulted in four categories: (1) research on prediction of performance; (2) research on formative individual feedback and assessment services; (3) research on social learning analytics; and, (4), research on competent use of LA applications.

In Table 1, below, the LA frameworks are clustered first in terms of their LA type and then according to the identified categories as shown in Table 1.

3.2 A Design Framework for Learning Analytics

As the literature analysis reveals, there are "softer" challenges that influence the acceptance of LA. These relate to issues of data ownership, ethical use and potential abuse of LA, and competences required to engage in meaningful LA activities. The pedagogic frameworks (e.g., Bakharia et al. 2016; Greller and Drachsler 2012; Gibson et al. 2014) for engaging in LA differ from other, more process-oriented frameworks (e.g., Clow 2012; Ferguson et al. 2014; Verbert et al. 2012). Building on holistic pedagogic frameworks, we aim at a descriptive framework that can later on be extended to a domain model or ontology. Depending on the (institutional) context, basic pedagogic principles and specific objectives, the workflow and process when engaging in LA may vary (Greller and Drachsler 2012).

The framework we propose (see Fig. 1 below) is similar to Greller and Drachsler (2012) and essentially represents a feedback loop. This conceptualization of the overall process as a feedback loop has been inspired by quality development frameworks (e.g., West et al. 2015) and dialogue with the multiple stakeholders involved is a key element. A particular pedagogic theory (or theory in use) and a specific learning design represent the starting point. From this consideration, both the particulars relating to the facilitation of learning as well as the specifics of LA are derived. The learning outcomes represent the feedback required in order to adjust and improve on the process and a particular pedagogic theory (in use) or learning design.

The design framework for LA comprises four dimensions:

- LA objectives
  These may relate to supporting reflection and/or prediction with regard to learning. Likewise, the LA objectives may relate to supporting individual students in their learning or to supporting interaction among students and/or facilitators. The framework of Greller and Drachsler (2012) distinguishes mainly between "reflection" and "prediction" as LA objectives. However, "individual learning" and/ or "social learning" need

Table 1 Comparison of conceptual frameworks for developing LA applications

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
Holistic Model for integrative LA applications Static Model to conceptualise LA systems	Framework to support enquiry-based evaluation of learning designs	Bakharia et al. (2016)	Dimensions: temporal analytics, comparative analytics, cohort dynamics, tool specific analytics, and contingency and intervention support tools
	Pedagogic framework with focus on cognitive operation	Gibson et al. (2014)	Framework COPA provides a basis for mapping levels of cognitive operation into an LA system (based on Bloom's taxonomy)
	Framework for a multilayer data warehouse	Ifenthaler and Widanapathirana (2014)	The LA framework is based on Greller and Drachsler (2012), and combines data directly linked to individual stakeholders, their interaction with the social web and the online learning environment, as well as curricular requirements (multilayer data warehouse)
	Research focus: identification of research question in each dimension	Chatti et al. (2012)	Four dimensions of the proposed reference model for LA are: (1) What kind of data does the system gather, manage, and use for the analysis? (2) Who does the analysis focus on? (3) Why does the system analyze the collected data? (4) How does the system perform the analysis of the collected data?
	Pedagogic framework to design LA systems	Greller and Drachsler (2012)	Dimensions: LA objectives (reference to pedagogical theories), competences as possible limitations, constraints (e.g. ethical issues), method, data, stakeholders

Table 1 (continued)

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
	Quality development of LA systems and measuring the impact of LA	West et al. (2015)	Mapping six domains: (1) institutional context, (2) transitional institutional elements, (3) learning analytics infrastructure, (4) transitional institutional elements, (5) learning analytics for retention, and (6) intervention and reflection)
	Quality development of LA systems and evaluation of LA tools	Scheffel et al. (2014)	Framework of quality indicators for learning analytics: objectives, learning support, learning measures and output, data aspects, organisational aspects
Holistic Model for integrative LA applications Process Model to develop LA systems	Process models for LA: Focus on Prediction	Clow (2012)	Five-step model of learning analytics: Capturing data, Report (identifying and measuring the students' progress), Predict (identify predictors for student success, outcomes and for identifying at-risk students), Act (interventions), Refine for continuous improvements
	Particularly for online learning (e.g. MOOCs)	Ferguson et al. (2014)	MOOC Learning Analytics Innovation Cycle—MOLAC at the micro-, meso- and macro-levels
	Process Modell with special focus on (self-)reflection	Verbert et al. (2012)	4-step process model to make use of LA: awareness, (self-)reflection, sense-making, impact

Table 1 (continued)

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
Partial Theory: (1) Prediction of Performance to support timely intervention and to prevent students from failing a course	Prediction of student success by using variables	Barber and Sharkey (2012)	Variables as predictors for student success (e.g. attendance per week, % cumulative course points per week, discussion post count per week, late assignments, etc.)
	Prediction of final grade by using online questions	Abdous et al. (2012)	Using data mining for predicting relationships between online question theme and final grade
	Prediction of student success by tracking of events	Romero-Zaldivar et al. (2012)	Tracking of events (such as work-time, commands, compile, etc.) and analyzed the gathered data with multiple regression for the estimation of the variance of performance
	Prediction by using post-hoc analysis	Pardo and Kloos (2011)	Discovery with models, post-hoc analysis of tutor logged data and sensor-free detectors of affect (based on classification algorithms)
	Performance index as metric for serious games analytics	Loh et al. (2015)	Expert performance index as a metric for serious games analytics (it can rank play-learners according to their competency levels in the serious games)

Table 1 (continued)

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
(2) Formative individual feedback and Assessment Services (e.g. Visualisation of information, representation of feedback)	Individualised formative feedback for students and teachers	Buckingham Shum and Deakin Crick (2012)	Model contains learning dispositions and transferable skills: The 7-dimensional construct of learning power
	Assessment Services: Visualising of Information	Duval (2011)	Dashboards for students and teachers: Visualization of information
	Generating informative feedback for students and teachers using LA	Tempelaar et al. (2015)	Collection of both dynamic, longitudinal user data and semi-static data, such as prior education: Entry test data and the combination of mastery data and use intensity data generated by the e-tutorial systems are a second best alternative for true assessment data
	Multiple representation of feedback types	Ali et al. (2012)	In order to provide instructors with pedagogically meaningful information and to help them extract such information on their own, the researchers embedded multiple representations of feedback types

Table 1 (continued)

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
(3) Social Learning Analytics social analytics only make sense in a collective context	Social learning analytics based on social media and associated recommendations	Buckingham Shum and Ferguson (2012)	A conception of social learning analytics as a distinctive class of analytics: 1) social network analytics (interpersonal relationships define social platforms and link learners to contacts, resources and ideas; 2) Discourse Analytics—language is a primary tool for knowledge negotiation and construction
	Tool-based Social Network Analysis	Schreurs et al. (2014)	An SLA tool was developed to visualize discussion activities in real time in order to help stimulate, monitor and evaluate interactions. Tool helped to visualize the learning relationships between users, based on their contributions to the discussion forums
	Gamification and gameful designs	Deterding et al. (2011)	Rules-based service system (driven by software) that provides feedback and interaction mechanisms with an aim to facilitate and support the learning process. Game design element constitute social and experiential dimensions of games in learning environments
	Content Analysis in discussion forums	Lin et al. (2013)	Content analysis for threaded discussion forums for monitoring capabilities. Facilitating the automated coding process within a repository of postings in an online course, in order less monitoring of the discussion to be needed by the instructor

Table 1 (continued)

Type of framework	Overall goal of the framework	References	Framework: Structure and core elements
(4) Competent Use of LA Applications students learning skills	Mapping multiliteracies to learning analytics techniques and applications	Dawson and Siemens (2014)	Multiliteracies are mapped to learning analytics techniques and applications: Experimentation (to Modelling, Knowledge domain mapping), Products and Creation (to personalisation, structured mapping, prediction), Network agility and citizenship (to relationship mining, modelling), task effectiveness and efficiency (to structured mapping, prediction)
	Providing a continual learning cycle, providing generic principles	Wise et al. (2016)	Student Tuning Model as a continual cycle in which students plan, monitor and adjust their learning activities (and their understanding of the learning activities) as they engage with LA

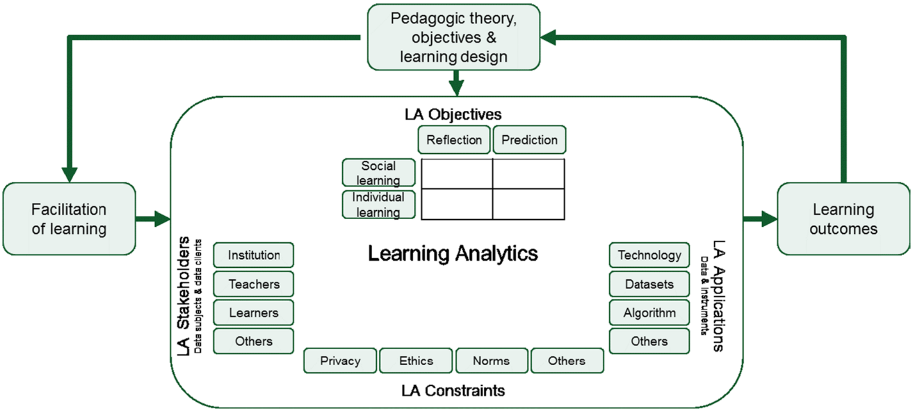

Fig. 1 Design framework for LA

to be differentiated as well. From a pedagogical perspective it is a design criterion for LA applications whether you focus on individual learning (e.g., individualized feedback, assessments, tracking learning progress, etc.) or on social learning in a collective context (e.g., social comparison activities, rewards from others as motivational factor for student engagement, etc.).

- LA stakeholders
  Stakeholders in LA activities are those that either are subjects of data analysis services or clients of data analysis services. Students and teachers, for example, may be subjects of data analyses in that data resulting from their learning activities are aggregated and analysed. Students, teachers, and institutional representatives, for example, may be clients of data analyses in that such analyses aim at supporting their activities and decisions.
- LA application
  Learning analytics applications comprise, among other things, technologies, platforms, data sets, and algorithms employed in carrying out analytics activities. The configuration of these elements may vary depending on the specific given context.
- LA constraints
  These constraints comprise rules and regulations concerning privacy and ownership of data, ethical considerations, as well as cultural norms and values. Again, these constraints may depend on the context at hand, for example, whether the educational institution pursuing LA is a primary school, an institution of higher education, or a commercial provider of learning and development services.

Taking this overall design framework for LA as a starting point, we propose in the following section a systematisation of one dimension of this framework: the learning objectives. The matrix derived later on serves as a basis for the use cases derived which focus on learning process and students' cognition.

3.3 A Framework for Learning Analytics Objectives

With regard to employing LA as a means to support and improve on (digital) learning, we propose a set of generic approaches based on a 2 × 2 matrix (see Table 1). This matrix includes

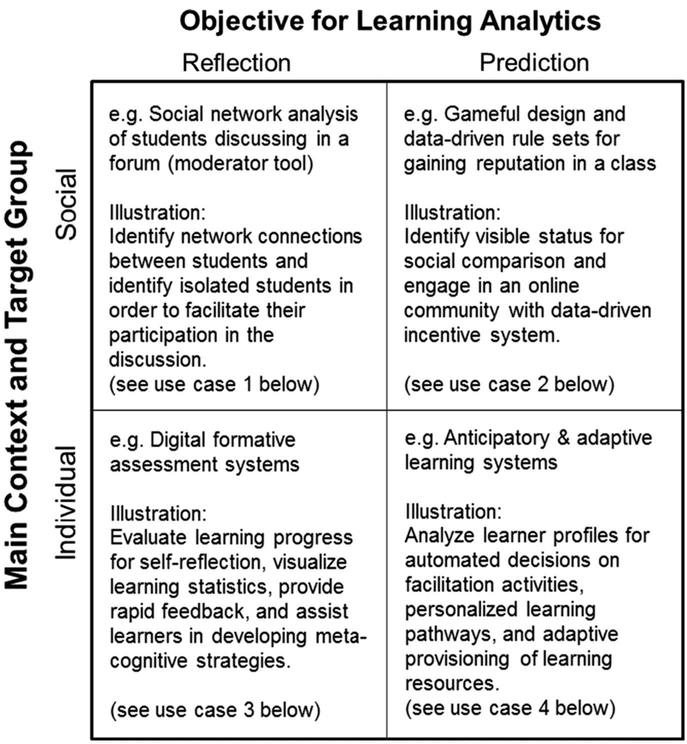

Fig. 2 Generic approaches to learning analytics with focus on Students' Cognition

the main pedagogical objectives to improve students' cognition and learning processes in either an individual or social learning context.

3.3.1 Student Cognition: Reflection and/ or Prediction

One dimension is set up via the distinction between reflecting on past learning activities versus predicting next/future learner activities. Reflection in this context refers to critical self-evaluation on the basis of (1) *own* data sets created in the process of either learning (students) or supporting learning (teachers/facilitators) and (2) data sets created by *others* (e.g., a teacher reflecting on his or her own teaching style based on data sets generated by the students) (Greller and Drachsler 2012, p. 41). Prediction refers to anticipating learner activities (e.g., further reducing investment in classwork or discontinuing with classwork altogether) and interventions that aim at preventing this (Siemens et al. 2011).

3.3.2 The Context of Learning Activities: Individual LA Systems and/ or Social LA Systems

The other dimension is set up via a distinction between individual learning activities versus social learning activities. Much work in LA is oriented towards supporting and determining individual achievement, for example by analysing the data generated through summative assessments. The focus on individual learners is related to the goal of personalization and

Table 2 Exemplary detailing of use case 1

Dimension	Exemplification
Pedagogic theory and learning design	Based on a socio-constructivist understanding of learning, (1) it is hypothesised that active participants in a discussion show better learning outcomes; (2) social network analyses of students discussing in a forum are conducted in order to discover effective ways of supporting participatory online learning
Objective	Reflection: Analyse student interactions in a forum discussion, identify network connections among students, and identify isolated students as a prerequisite for remedial action (aimed at helping these students create links to others)
Stakeholders	Data subjects: a group of learners Data clients: Teachers, tutors, discussion moderators
LA model	Partial Theory “Social Learning Analytics (social analytics only make sense in a collective way)” (see 3) in literature review)
LA application: data	Protected data set: student interactions and posts in the discussion forum of the LMS; Relevant indicators: posts published, post replied to; Time frame: period of time set for a specific discussion task
LA application: instruments	Technology: social network analysis (SNA), statistics provided by SNAPP tool; Presentation: network diagram visualisation, statistics tables
Competences required/to be developed	Interpretation: Do the data clients have the necessary competences to interpret and act upon the information available? Critical thinking: Are data clients able to critically evaluate the data basis (e.g., missing data) when interpreting and/or devising a path of corrective action?
Constraints	Privacy: Is the analysis in accordance with privacy arrangements and are the students properly informed? Ethics: What are the dangers of abuse/misguided use of the data? Norms: Are there legal data protection or IPR issues related to this kind of use of student data? Time scale. Is the analysis post-hoc or just-in-time? Will students still be able to benefit from the analytics outcome?

individualization. In order to provide pedagogically valuable feedback, assessment systems have to become intelligent and connected with higher-order learning skills. Adaptive learning systems (focused on individual learning and prediction) represent a distinct, quite new field of research based on interactive machine learning.

Buckingham Shum and Ferguson (2012, p. 4) have argued that “new skills and ideas are not solely individual achievements, but are developed, carried forward, and passed on through interaction and collaboration”. In consequence, LA in social systems (e.g., in the context of a classroom at a school) “must account for connected and distributed interaction activity”. Buckingham Shum and Ferguson therefore propose social learning analytics as a domain in its own right (2012). Similar, gamification or gameful design for learning is considered as an on own domain using LA in social systems, for example to provide visible status and learning progress, social comparison and reputation (e.g., based on badges). Rule-sets and game design elements implemented in a learning environment can provide systematic support for learning and may contribute to student engagement. They may function as “nudges” that influence

Table 3 Exemplary detailing of use case 2

Dimension	Exemplification
Pedagogic theory and learning design	Based on a view of active learning as a constructive process, self-assessments are seen as a way to enhance reflection processes and learner engagement. Feedback is most effective when highly related to clearly formulated learning goals
Objective	Reflection: Evaluate objective and subjective assessments; identify knowledge gaps in order to support better developed learning strategies (e.g., preparation for an exam); provide opportunities for active learning during/after lectures in order to evaluate their impact on student experience
Stakeholders	Data subjects: students; Data clients: learners/students (for self-reflection), and teachers (for scaffolding process)
LA model	Partial Theory "Formative individual feedback and Assessment Services" (see 2) in literature review)
LA application: data	Protected data set based on students' assessment results Relevant indicators: e.g., objective and subjective assessments; algorithm: % difference of discrepancy over a given period of time; Time scale: period of time defined for assessment activities and comparison of objective/subjective assessments
LA application: instruments	Technology: assessment tool and statistics (quantitative analysis); Presentation: visual feedback, written communication (feedback) with individual preferences
Competences required/to be developed	Students: self-assessment competences; metacognitive learning strategies Teachers: scaffolding competences (help students to interpret the data)
Constraints	Privacy: Is anonymity (not disclosing student names) required for effective self-assessment? Ethics: Is the potential for misinterpreting data hindering the scaffolding process by teachers? Norms: Is social comparison inducing motivation or demotivation in students in the first semester? Time scale: Should the analyses be carried out in class or outside of class (trade-off with time required for teaching time)?

student behavior in a predictable manner without having to resort to prohibitions, commandments, or extrinsically motivating incentives (Fig. 2).

The matrix developed here elaborates one dimension of the proposed framework and emphasizes the need to tackle LA objectives from a pedagogical perspective in order to support students' learning skills. The matrix provides a starting point for generating use cases in an LA systematic.

4 Use Cases

The following section illustrates how the framework comprising generic approaches can be translated into specific use cases. Starting from the use cases provided by Greller and Drachsler (2012), we elaborate the pedagogical perspective by exemplifying the

Table 4 Exemplary detailing of use case 3

Dimension	Exemplification
Pedagogic theory and learning design	Based on ideas of behaviourism ("behavioural economics, "big nudging"), learning design includes: the use of game elements in learning environments and for particular types of learners in order to achieve i) motivation for student engagement and ii) better learning outcomes (ideally on higher levels of cognitive processes)
LA objective	Prediction: The LA application based on a data-driven rule system and a gameful design provides an incentive system for different types of learners in order to increase student engagement in a social context (e.g., community or class)
Stakeholders	Data subjects: community, an entire class/cohort and individual learners; Data clients: learners/students, learning designers implementing rule systems (closely interacting with students)
LA model	Partial Theory "Prediction of Performance to support timely intervention and to prevent students from failing a course)" (see 1) in literature review)
LA application: data	Protected data set: student activities (e.g., contributions in forums, peer rating, quizzes and points awarded, team competitions)
LA application: instruments	Game design elements (e.g., visual status, badges, awards, avatars as personal identities) and a system of rules (implemented on a separate platform or in an LMS)
Competences required/to be developed	Students: readiness for (more) autonomy in learning and for self-regulation based on system feedback; ability to navigate gamified environments; ability to interpret dashboard information. Learning designers must consider ability and motivation of learners when creating a gamified learning design
Constraints	Privacy: What are the data security issues when used as part of the grading? Ethics: What are dangers of abuse/misguided use of a data-driven rule system? Norms: Course gamification could be merely misused by masking the terms; for example, by calling assignments "quests" and scores "experience points" without contributing to the students' learning goals Time scale: What is the overall dramaturgy of the design and how much time is required for different phases (e.g., onboarding, scaffolding, mastery)?

pedagogical theory. Additionally, we spell out relevant aspects to consider in the design of learning activities.

4.1 Use Case 1: Social Learning Analytics for Reflection

The first use case relates to conducting a social network analysis of students discussing in a forum, for example using the SNAPP tool developed by Dawson (2008). This implies a shift in attention away from the summative assessment of individuals towards learning analytics of social activity (Buckingham Shum and Ferguson 2012, p. 5). In this context, it is relevant to distinguish between social analytics sui generi (e.g., social networks analysis or discourse analytics) and socialised analytics that are based on personal analytics while also

Table 5 Exemplary detailing of use case 4

Dimension	Exemplification
Pedagogic theory and learning design	Based on ideas of behaviourism and cognitivism, learners are presented—in a highly adaptive manner—with materials and problems that enable them to develop new knowledge and concepts and to provide immediate feedback to performance on problem solutions
Objective	Prediction based on student model/learner profiles and prescription of next learning activities in order to facilitate comprehension and retention
Stakeholders	Data subject: learners/students; Data client: teachers, educational institutions concerned about student drop-out rates
LA model	Partial Theory "Prediction of Performance to support timely intervention and to prevent students from failing a course)" (see 1) in literature review)
LA application: data	Data from different sources; algorithms for student modelling
LA application: instruments	Adaptive learning systems, intelligent tutoring systems
Competences required/to be developed	Students: basic understanding of how such systems work and acceptance of permanent monitoring as well as suggestions by system; Learning designers/institutions: deep understanding of how such systems model the domain, the students, and the tutoring process and where they differ in order to select/configure appropriate solutions
Constraints	Privacy: What data are generated in closely monitoring students' activities and who has access to these in what manner? Ethics and norms: Is there a risk that students guided by such systems will develop less metacognitive ability regarding monitoring and planning their own learning?

being relevant in a social learning context (e.g., analytics of user generated content, analytics of personal dispositions, or analytics of contexts such as mobile computing and the networking opportunities related thereto) (Buckingham Shum and Ferguson 2012, p. 10).

The following example illustrates the first type of social analytics sui generis (Table 2).

4.2 Use Case 2: Individual Analytics for Reflection

This use case is about LA with a focus on reflection at the individual level. As Evans (2013) discovered in a thematic analysis of the research evidence on assessment feedback in higher education (based on over 460 articles over a time span of 12 years), effective online formative assessment can enhance learner engagement during a semester class. Focused interventions (e.g., self-checking feedback sheets, mini writing assessments) can make a difference to student learning outcomes as long as their value for the learning process is made explicit to and is accepted by students and lecturers. The development of self-assessment skills requires appropriate scaffolding on the part of the lecturer working with the students so as to achieve co-regulation (Evans 2013) (Table 3).

4.3 Use Case 3: Social Analytics for Prediction

The more environments for working and learning become digital, the more data is generated in the course of activities relating to working and learning: accessing web pages, working on short knowledge tests, posting in an online forum, commenting on a forum post, etc. (Manouselis et al. 2010). Until recently, the availability of such data for analysis had been mostly confined to what is going on inside a particular learning management system (LMS). With the development of the xAPI specification for transfer of interaction data, a much wider range of data from both inside and outside an LMS can be made available for analysis (Berking et al. 2014).

These developments help to enable gamified learning designs (Berkling and Thomas 2013). By this we refer to "the use of game design elements in non-game contexts". Frequently, this takes the form of awarding points and badges for individual learning activities (e.g., posting in a discussion forum) and displaying top performers (or rather point generators) on leaderboards (Deterding et al. 2011; Mak 2013). While there is evidence that gamified designs (can) lead to higher student engagement and improved learning (Dicheva et al. 2015, p. 83), the opportunity to engage in a more systematic motivation design that also includes choices, social integration, team assignments, as well as characters and stories is often missed (Seufert et al. 2017).

The following use case focuses on gamified learning designs as an example of behavorial "nudging" (Table 4).

4.4 Use Case 4: Individual Analytics for Prediction and Prescription

More than 30 years ago, Leonard Bloom demonstrated that individual tuition leads to a 2-Sigma performance improvement in tests compared to then standard expository teaching techniques in classrooms with about 30 learners (Bloom 1984). The idea of individualised tuition for large numbers of learners is currently being pursued in the context of the research and development of adaptive or intelligent tutorial platforms (Romero et al. 2008). The research and development in this area is based on advances in artificial intelligence and cognitive computing (Verbert et al. 2012). Adaptive learning systems aim at supporting the development of conceptual structures in learners rather than merely supporting (repetitive) problem solving as was the case in prior generations of so-called intelligent tutorial systems.

Adaptive learning systems closely track student activities and student performance and provide students with adequate learning pathways and adaptive learning resources based on machine learning algorithms and predictive models (Butz et al. 2003).

However, more substantial empirical research is needed, in particular to investigate (Nour et al. 1995) the appropriateness of such algorithms in disciplines other than the typical mastery learning subjects (e.g., biology, mathematics, information science) and their effectiveness for reaching higher learning outcomes (Table 5).

5 Discussion

Learning analytics (LA) has the potential to enable learners, teachers, and their institutions to better understand and predict learning and performance. However, the pedagogical perspective, and in particular the focus on reflection instead of prediction, has been neglected in research so far. Therefore, the main contribution of the paper is to provide a generic framework for the design of LA environments from a pedagogical perspective and focusing on students' cognition.

The presented framework provides a matrix with two important dimensions from a pedagogical point of view: (1) Objective for LA: Reflection versus Prediction and (2) Main context and target group: Individual analysis versus social (network) analysis. Based on the proposed framework we developed use cases in order to define the overall generic strategy in more detail. The proposed conceptual framework serves as a heuristic model for identifying and structuring the research questions. A learning analytics plan for research could be tuned depending on the pedagogic goals.

However, we want to emphasize that the proposed generic framework has its limits as a helpful concept map for further research. The proposed two dimensions might be too narrow to pursue the pedagogical perspective in LA environments. The second limitation of our research is that the sample of teachers of our focus group was rather small with 12 lecturers. A further limitation is that the empirical validation of our developed framework is missing. Most important for the validation of the proposed framework is its perceived utility by the stakeholders, in particular the course designers, lecturers, as well as the students in the different use cases. In order to verify that the model does indeed provide actionable information, a pilot within an action research design to validate and revise the generic model and for every use case is planned with only a few experts of the initial task force. These more experienced teachers are looking at the model in terms of both its accuracy (does the information provided by the model align with what they learn by talking to the student?) and its utility (does it trigger contact with the right students and are those students then successful?). Once the pilot is completed, the utility will be evaluated and a decision will be made as to whether to implement the model into the production processes, making the results available to all teachers. The model will continue to be refined even after initial implementation.

6 Conclusion and Outlook

Current research and discussion on big data in education focuses largely on (1) the potential of learning analytics to increase the efficiency and effectiveness of educational processes, (2) the ability to identify and support students at risk, and (3) to inform efforts to reduce drop-out rates. Accordingly, the main focus is on prediction. Therefore, we emphasized the research question how big data and learning analytics can be employed in order to improve learner guidance, students' learning processes and learning outcomes with regard to reflection and meta-cognitive abilities for self-regulated learning.

Competency development on the part of the data clients (students, teachers/tutors, institutions) is a key requirement for progress in this area. On the basis of the survey data available, Greller and Drachsler (2012, p. 51) have pointed out that the large majority of students currently do not have command of the competences required to interpret LA results and to determine appropriate next activities.

In our model (cf. Figure 1), we include critical evaluation skills among the key competences for LA (similar to Greller and Drachsler 2012). A superficial understanding of data presentation can lead to false conclusions. Furthermore, it is important to understand that data not included in the respective LA approach may be equally if not more important than the data set that is included. To judge a learner's performance merely on one aspect, such as quantitative data provided by a LMS, is like looking at a single piece taken from a larger jigsaw puzzle. Lifelong learning takes place across a wide range of schooling, studying, working, and everyday life situations. In addition to competency requirements, acceptance factors influence the application or decision making that follows an analytics process. Lack of acceptance of analytics systems and processes can lead to blunt rejection of either the results or the suggestions on the part of relevant constituencies (data clients).

In order to deal with these issues, future research should focus on empirical evaluation methods of learning analytics tools (Ali et al. 2012; Scheffel et al. 2014) and on competence models for digital learning (Dawson and Siemens 2014). The conceptual framework can be further elaborated with the application of the four different use cases by adjusting and integrating partial theories for the competence development of students (e.g., mapping multiliteracies to learning analytics techniques and applications (Dawson and Siemens 2014). It is planned that these cases become four real case studies in which we analyse critically the outcomes, problems and implications of each case. This will be based on a Student Tuning Model as a continual cycle in which students plan, monitor, and adjust their learning activities (and their understanding of the learning activities) as they engage with LA (Wise et al. 2016).

Acknowledgements Funding was provided by Research Grant of the University of St. Gallen.

References

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

Ackermann, F., Eden, C., & Cropper, S. (2004). *Getting started with cognitive mapping*. Retrieved from http://www.banxia.com/pdf/de/GettingStartedWithCogMapping.pdf.

Ali, L., Hatala, M., Gašević, D., & Jovanović, J. (2012). A qualitative evaluation of evolution of a learning analytics tool. *Computers & Education, 58*(1), 470–489.

Bakharia, A., Corrin, L., de Barba, P., Kennedy, G., Gašević, D., Mulder, R., et al. (2016). A conceptual framework linking learning design with learning analytics. In *Proceedings of the sixth international conference on learning analytics & knowledge* (pp. 329–338). New York: ACM Press.

Barber, R., & Sharkey, M. (2012). Course correction: Using analytics to predict course success. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 259–262). New York: ACM Press.

Berking, P., Foreman, S., Haag, J., & Wiggins, C. (2014). *The experience API—Liberating learning design*. Report, eLearning Guild.

Berkling, K., & Thomas, C. (2013). Gamification of a software engineering course and a detailed analysis of the factors that led to its failure. In M. E. Auer & D. Guralnick (Eds.), *Proceedings of international conference on interactive collaborative learning* (pp. 525–530). https://doi.org/10.1109/icl.2013.6644642.

Bloom, B. S. (1984). The 2 Sigma Problem: The search for methods of group instruction as effective as one-to-one tutoring. *Educational Researcher, 13*(6), 4–16.

Buckingham Shum, S., & Deakin Crick, R. (2012). Learning dispositions and transferable competencies: Pedagogy, modelling and learning analytics. In *Proceedings 2nd international conference on learning analytics & knowledge* (pp. 324–335). New York: ACM Press.

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

Butz, M. V., Sigaud, O., & Gerard, P. (2003). Internal models and anticipations in adaptive learning systems. In M. V. Butz, O. Sigaud, & P. Gerard (Eds.), *Anticipatory behavior in adaptive learning systems. Volume 2684 of the series lecture notes in computer science* (pp. 86–109). Berlin: Springer.

Chatti, M. A., Dyckhoff, A. L., Schroeder, U., & Thüs, H. (2012). A reference model for learning analytics. *International Journal of Technology Enhanced Learning, 4*(5/6), 318–331.

Clow, D. (2012). The learning analytics cycle: Closing the loop effectively. In *Proceedings of the 2nd international conference on learning analytics and knowledge* (pp. 134–138). New York: ACM Press.

Cope, B., & Kalantzis, M. (2016). Big data comes to school: Implications for learning, assessment, and research. *AERA Open, 2*(2), 1–19.

Dawson, S. (2008). A study of the relationship between student social networks and sense of community. *Educational Technology & Society, 11*(3), 224–238.

Dawson, S., & Siemens, G. (2014). Analytics to literacies: The development of a learning analytics framework for multiliteracies assessment. *The International Review of Research in Open and Distributed Learning, 15*(4), 285. Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1878/3006.

Deterding, S., Dixon, D., Khaled, R., & Nacke, L. (2011). From game design elements to gamefulness. In Academic MindTrek 2011, ACM Digital Library. ACM Special Interest Group on Computer-Human Interaction. & ACM Special Interest Group on Multimedia. (Eds.), *Proceedings of the 15th international academic MindTrek conference envisioning future media environments. Defining „Gamification"* (pp. 9–15). New York: ACM Press.

Dicheva, D., Dichev, C., Agre, G., & Angelova, G. (2015). Gamification in education: A systematic mapping study. *Educational Technology & Society, 18*(3), 75–88.

Duval, E. (2011). Attention please! Learning analytics for visualization and recommendation. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 9–17). New York: ACM Press.

Dyckhoff, A. L., Zielke, D., Bültmann, M., Chatti, M. A., & Schroeder, U. (2012). Design and implementation of a learning analytics toolkit for teachers. *Educational Technology & Society, 15*(3), 58–76.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83,* 70–120.

Ferguson, R., Clow, D., Macfadyen, L., Essa, A., Dawson, S., & Alexander, S. (2014). Setting learning analytics in context: Overcoming the barriers to large-scale adoption. *Journal of Learning Analytics, 1*(3), 120–144.

Gibson, A., Kitto, K., & Willis, J. (2014). A cognitive processing framework for learning analytics. In *Proceedings of the fourth international conference on learning analytics and knowledge* (pp. 212–216). New York: ACM Press.

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

Ifenthaler, D. (2015). Learning analytics. In J. M. Spector (Ed.), *The SAGE encyclopedia of educational technology* (Vol. 2, pp. 447–451). Thousand Oaks, CA: Sage.

Ifenthaler, D., Adcock, A. B., Erlandson, B. E., Gosper, M., Greiff, S., & Pirnay-Dummer, P. (2014). Challenges for education in a connected world: Digital learning, data rich environments, and computerbased assessment—Introduction to the inaugural special issue of technology, knowledge and learning. *Technology, Knowledge and Learning, 19*(1), 121–126.

Ifenthaler, D., & Widanapathirana, C. (2014). Development and validation of a learning analytics framework: Two case studies using support vector machines. *Technology, Knowledge and Learning, 19*(1–2), 221–240.

Kelly, N., Thompson, K., & Yeoman, P. (2015). Theory-led design of instruments and representations in learning analytics: Developing a novel tool for orchestration of online collaborative learning. *Journal of Learning Analytics, 2*(2), 14–43.

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning analytics in higher education—A literature review. In A. Peña-Ayala (Ed.), *Learning analytics: Fundaments, applications, and trends* (pp. 1–23). Berlin: Springer International Publishing. https://doi.org/10.1007/978-3-319-52977-6_1.

Lin, C.-F., Yeh, Y., Hung, Y. H., & Chang, R. (2013). Data mining for providing a personalized learning path in creativity: An application of decision trees. *Computers & Education, 68,* 199–210.

Loh, C. S., Sheng, Y., & Ifenthaler, D. (2015). *Serious games analytics: Theoretical framework* (pp. 3–29). Berlin: Springer.

Mah, D.-K. (2016). Learning analytics and digital badges: Potential impact on student retention in higher education. *Technology, Knowledge and Learning, 21*(2), 285–305.

Mak, H. W. (2013). *The gamification of college lectures at the University of Michigan*. Retrieved from http://www.gamification.co/2013/02/08/the-gamification-of-college-lectures-at-the-university-of-michigan/.

Manouselis, N., Drachsler, H., Vuorikari, R., Hummel, H., & Koper, R. (2010). Recommender systems in technology enhanced learning. In P. B. Kantor, F. Ricci, L. Rokach, & B. Shapira (Eds.), *Recommender systems handbook* (pp. 387–415). Berlin: Springer.

Nour, M., Abed, E., & Hegazi, N. (1995). A proposed student model algorithm for an intelligent tutoring system. In *Proceedings of the 34th SICE annual conference. International session papers* (pp. 1327–1333), Hokkaido.

Papamitsiou, Z., & Economides, A. (2014). Learning analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Journal of Educational Technology & Society, 17*(4), 49–64.

Pardo, A., & Kloos, C. D. (2011). Stepping out of the box: Towards analytics outside the learning management system. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 163–167). New York, NY: ACM Press.

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give me a customizable dashboard: Personalized learning analytics dashboards in higher education. *Technology, Knowledge, and Learning, 22*(3), 317–333.

Romero, C., Ventura, S. Espejo, P. G., & Hervs, C. (2008). Data mining algorithms to classify students. In R. de Baker, T. Barnes, & J. Beck (Eds.), *Proceedings of the 1st international conference on educational data mining* (pp. 8–17). Retrieved from http://www.educationaldatamining.org/EDM2008/-uploads/proc/1_Romero_3.pdf.

Romero-Zaldivar, V.-A., Pardo, A., Burgos, D., & Kloos, C. D. (2012). Monitoring student progress using virtual appliances: A case study. *Computers & Education, 58*(4), 1058–1067.

Scheffel, M., Drachsler, H., Stoyanov, S., & Specht, M. (2014). Quality indicators for learning analytics. *Journal of Educational Technology & Society, 17*(4), 117–132.

Schreurs, B., De Laat, M., Teplovs, C., & Voogd, S. (2014). Social learning analytics applied in a MOOC-environment. *e-Learning Papers, 26,* 45–48.

Seufert, S., Preisig, L., Krapf, J., & Meier, C. (2017). *Von Gamification zum systematischen Motivationsdesign mit kollaborativen und spielerischen Gestaltungselementen. Konzeption und Anwendungsbeispiele* (scil Arbeitsberichte No. 27). St.Gallen: Institut für Wirtschaftspädagogik/scil.

Siemens, G., Gasevic, D., Haythornthwaite, C., Dawson, S., Shum, S. B., Ferguson, R., et al. (2011, July 28). *Open learning analytics: An integrated & modularized platform*.

Sin, K., & Loganathan, M. (2015). Application of big data in education data mining and learning analytics—A literature review. *ICTACT Journal on Soft Computing, 5*(4), 1035–1049. https://doi.org/10.21917/ijsc.2015.0145.

Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning Analytics in a data-rich context. *Computers in Human Behavior, 47,* 157–167.

Verbert, K., Manouselis, N., Drachsler, H., & Duval, E. (2012). Dataset-driven research to support learning and knowledge analytics. *Educational Technology & Society, 15*(3), 133–148.

West, D., Heath, D., & Huijser, H. (2015). Let's talk learning analytics: A framework for implementation in relation to student retention. *Online Learning, 20*(2), 12. Retrieved from https://olj.onlinelearningconsortium.org/index.php/olj/article/viewFile/792/202.

Wise, A. F., Vytasek, J. M., Hausknecht, S., & Zhao, Y. (2016). Developing learning analytics design knowledge in the "middle space": The student tuning model and align design framework for learning analytics use. *Online Learning, 20*(2), 3. Retrieved from https://olj.onlinelearningconsortium.org/index.php/olj/article/view/783.