Big Data in Education: Supporting Learners in Their Role as Reflective Practitioners

Sabine Seufert & Christoph Meier, University of St.Gallen

Abstract

Recent discussions on the topic of big data in education currently revolve heavily around the potential of learning analytics to increase the efficiency and effectiveness of educational processes and the ability to reduce drop-out rates (with focus on prediction and prescription). This chapter refers to the pedagogical perspective in order to provide learners with appropriate digital tools for self-organization, and enable them to further develop their competences and skills. The normative orientation towards the reflective practitioner in the digital age highlights the necessity to foster reflection on big data approaches in education.

For this, a conceptual framework for digital learning support is introduced and illustrated via four case studies. This conceptual framework can be applied in two ways: First, it serves as a heuristic model for identifying and structuring the design questions that must be answered by developers of learning environments. Secondly, the conceptual framework provides guidance when it comes to generating and detailing relevant research questions that can then be transferred and processed in specific research designs.

Keywords

1) The Second Machine Age: Augmentation as a key challenge

Digitalization in an advanced form is about the expansion of the Internet from a web initially connecting information (e.g. webpages), then people (social software) and now objects (Internet of Things) (Seufert & Vey, 2016). It is, also, about processes and control systems that work mostly digitally, Big Data, elaborate predictive and prescriptive analytics and growing use of artificial intelligence and digital assistants in decision making. And, lastly, it is about the discovery of hidden connections in the enormous volume of data in the digital universe. Computing has moved from mechanical machines that undertook simple arithmetic to those that could be programmed digitally to intelligent machines that can learn and reason from completely unstructured data (machine learning).

With high hopes for a miraculous digital welfare economy on the one hand and fears of the end of jobs and prosperity on the other, we must not overlook the implications of human augmentation based on digital technologies. Currently, little attention is paid to this aspect in the educational debate – even though there are already important books out there on this aspect (e.g. Frank Pasquale's "The Black Box Society" and Eli Pariser's "The Filter Bubble"). There is a risk that even digitally skilled teachers are overwhelmed by the intensive and fast moving technological, social, and economic developments and not well prepared for the upcoming technological advances such as transformative Artificial Intelligence.

Discussions on big data in education currently focus mostly on the potential of learning analytics to increase the efficiency and effectiveness of educational processes (for example, how to reduce drop-out rates). Helbing et al. recently emphasized that our society is "at a digital crossroad" (Helbing et al., 2017): "If ever more powerful algorithms would be controlled by a few decision-makers and reduce our self-determination, we would fall back in a Feudalism 2.0, as important historical achievements would be lost." Helbing et al. argue for a shift from remote control based on data-driven top-down decision making to self-control. Educational policy makers have to develop a vision for the successful partnership of human and machine (in particular due to transformative Artificial Intelligence and emotional robotics), with the aim to win synergies through complementary competences.

Against this background, this contribution develops a general discussion about being competent in a digital world and about human augmentation as a key challenge in the fourth industrial revolution. The overarching research question is how to provide learners with appropriate digital tools to enable self-organization and further development of individual competences and skills: how to exploit the potential of big data in education to support learners as so called "reflective practitioners"?

As a starting point, we will elaborate what competent use of Big Data and Artificial Intelligence means. Based on this discussion, we propose the normative orientation of being a reflective practitioner in a digital society ("digital citizen"). This normative orientation provides the required foundational base for dealing with issues related to the competent use of big data for learning system

developers. We propose a conceptual framework for digital learning support on the basis of Big Data, Learning Analytics and Gamification and illustrate this on the basis of 4 cases. The conceptual framework allows for progress in two ways: First, it serves as a heuristic model for identifying and structuring the design questions that must be answered by training course providers. Second, it provides support for generating and detailing relevant research questions that can then be transferred and processed in specific research designs.

2) Competent use of Big Data and Artificial Intelligence

Big Data and analytics currently are a burgeoning field of research and development (Abdous & Yen, 2012; Ali et al., 2012; Dyckhoff et al., 2012). With regard to Big Data in education, what does "being *digital competent*" mean in light of these developments? The European commission (EU, 2006) provides the following definition: "Digital competence involves the confident and critical use of Information Society Technology (IST) for work, leisure and communication. It is underpinned by basic skills in ICT: the use of computers to retrieve, assess, store, produce, present and exchange information, and to communicate and participate in collaborative networks via the Internet." The EU framework of digital competences identifies the key components of digital competence in 5 areas: information, communication, content creation, safety and problem solving. To be competent requires knowledge and instrumental skills, advanced skills and appropriate attitudes. However, with advanced cognitive computing systems these key competence areas require higher-order skills in terms of complementary competences driven by augmentation. The following situation provides an example of what it means to be digitally competent – here understood as competent use of augmentation:

Imagine the following situation:

You feel really bad physically and decide to visit the emergency service at a local hospital. When it is your turn, two physicians enter, an elderly physician on duty together with his youngish assistant. The elderly physician says he commands 30 years of experience, he will find out what is wrong with you. The youngish assistant says he works with a computer database which comprises the knowledge of 600 years of western medical practice.

Who would you rather turn to?

This (admittedly hypothetical) scenario demonstrates the developments in medical diagnoses and why we need to come to terms with the changes in human-machine interaction (Holzinger, 2016). In the healthcare system, one person will soon generate 1 million GB of health-related data during her or his

lifetime – equivalent to about 300 million books (Karin Vey, IBM Research, personal communication). One example of augmentation is interactive machine learning in health informatics The following quote demonstrates how in this field humans and machines interact with complementary competences (Holzinger, 2016, p. 119):

> "The goal of Machine Learning (ML) is to develop algorithms which can learn and improve over time and can be used for predictions. Most ML researchers concentrate on automatic machine learning (aML), where great advances have been made, for example, in speech recognition, recommender systems, or autonomous vehicles. Automatic approaches greatly benefit from big data with many training sets. However, in the health domain, sometimes we are confronted with a small number of data sets or rare events, where aML-approaches suffer of insufficient training samples. Here interactive machine learning (iML) may be of help, having its roots in reinforcement learning, preference learning, and active learning. The term iML is not yet well used, so we define it as '*algorithms that can interact with agents and can optimize their learning behavior through these interactions, where the agents can also be human.*' This '*human-in-the-loop*' can be beneficial in solving computationally hard problems, e.g., subspace clustering, protein folding, or k-anonymization of health data, where human expertise can help to reduce an exponential search space through heuristic selection of samples. Therefore, what would otherwise be an NP-hard problem, reduces greatly in complexity through the input and the assistance of a human agent involved in the learning phase."
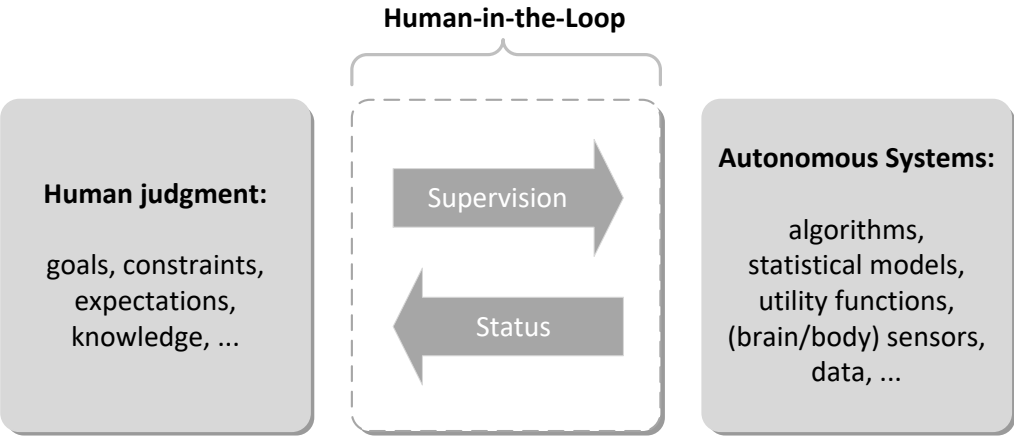

Figure 6.1: The phenomenon "Human-in-the loop" (HITL) (Rahwan, 2016)

Decisions on all management levels increasingly have to be made in consideration of computer-based data analyses as well as one's own gut feeling. Decision makers have to learn in what cases algorithms can help them to detect distortions in their thinking and when intuition in form of condensed knowledge needs to come into play. It is about being able to design flexible decision processes, understanding the role of digital tools and using them well-versed. A cognitive assistant that is

equipped with artificial intelligence can make statistically sound proposals on the basis of enormous data volumes. Nonetheless, these results are limited. The proposals are valid only for a specific area that we specify for the machine and a question that we trained with the system. The human on the other hand is able to make a holistic evaluation of the situation. A decision maker has to know about the different competences and limitations of machines on the one hand and humans on the other hand and be able to design adequate decision processes.

We argue that - as digitalization of knowledge work advances – not so much workforce substitution through automation but rather augmentation of knowledge work becomes the real new challenge. It is important to see work not as a zero-sum game where machines gain an ever increasing part. Many things that today cost a knowledge worker a lot of time. For example, time-consuming search for sources of information can be performed by computer systems in the future. However, significant improvement in research requires cooperation of machines and humans - so that the sources and knowledge collected will be usable not only in new and better ways but also in a considerably more economically manner. This allows better support of decisions. However, without humans providing direction, machines provide only fragmented or irrelevant results. Therefore, having humans in the loop (HITL) in augmentation is important (Rawahn, 2016).

Today, many apps already learn from human behaviour in order to improve their ability to offload routine work (e.g. sms systems, cognitive automation). The same is true for AI as for example in medical diagnosis. One of the main challenges in developing such systems is that the AI engineers training the systems using huge amounts of data (Big data) usually are not domain experts. Therefore, any biases or errors in the data will create models that reflect those biases and errors. That is the reason why Holzinger (Holziger, 2016, p. 3) demands a human lens for AI: "Human-in-the-loop machine learning (or interactive machine learning) is work that is trying to create systems to either allow domain experts to do the training or at least be involved in the training by creating machines that learn through interactions with experts. (…) At the heart of human-in-the-loop computation is the idea of building models not just from data, but also from the human perspective of data."

Recently, Rahwan (Rahwan, 2016) emphasized the need for a scaled-up version of HITL in his blog: a "Society-in-the-Loop" approach for developing AI systems with wide societal implications. Similarly, Helbing et al. (Helbing et al., 2017) describe a future scenario: from programmed computer to programmed society and programmed citizens: "Everything started quite harmlessly. Search engines and recommendation platforms began to offer us personalized suggestions for products and services. This information is based on personal and meta-data that has been gathered from previous searches, purchases and mobility behavior, as well as social interactions. (…) The more is known about us, the less likely our choices are to be free and not predetermined by others. But it won't stop there. Some software platforms are moving towards "persuasive computing." In the future, using sophisticated manipulation technologies, these platforms will be able to steer us through entire courses of action, be it for the execution of complex work processes or to generate free content for Internet platforms, from

which corporations earn billions. The trend goes from programming computers to programming people."

Leaders, educational policy makers, and responsible educational developers (including teachers who develop competences for digital citizenship), must understand this connection and develop a vision for the successful partnership of human and machine – human values and big data/ artificial intelligence –, with the aim to win synergy through complementary competences. In the next section we will further elaborate on these complementary competences in order to clarify the implications for being digitally competent in the new domain of man-machine-interactions.

3) Digital competences as core competences: What is really new?

"Imagine a situation where the amount of data about our world determines how well we can see and understand it. It, then, becomes clear that we are moving from a time of darkness, where we did not see enough to make good decisions, into a digital age where we tend to be blinded by information, i.e. suffering from extreme information overload. To master this situation, we will need suitable filters, something like 'digital sun glasses'. Whoever builds these filters will determine what we see [1]. This creates possibilities to influence people's decisions in such subtle a way that they would consider these decisions their own, while they have been actually remote controlled." (Helbing, 2017)

In the last few decades, computers have posed a daunting challenge for us. In particular, in order to achieve better results, we had to learn how to adapt to the functioning of the machine. Now we are experiencing a radical change. The interaction with the system becomes increasingly natural. We can easily communicate with the systems – through our language and our gestures. Nevertheless, there are important differences in the communication with machines compared to the communication with humans (Seufert / Vey, 2016). The former is purely objective and specific in depth. A person, in contrast, would initiate a richer, more extensive exchange – for example, introduce more context, associations, and metaphors. Moreover, dialogue between people includes three further levels: self-disclosure, relationship level, and appeal character (Schulz von Thun, n.d.).

Big data and artificial intelligence challenge us to identify and develop our core competences. This is about raising our cognitive-emotional skills to a higher level. For us humans, it will be important in the future to be able to distinguish between accessibility through language expression and the restrictions mentioned above with respect to communication levels. We will be able to interoperate with data in a new way, compensate for local data space, and navigate in hybrid worlds. For example, we will make decisions in groups in immersive data spaces. This in many ways new interaction with digital content requires new skills. AI challenges us to identify and develop our core competences. It is about raising our *cognitive-emotional skills* to a higher level (Augmentation skills). Highly developed skills, such as abstraction ability, generalization, creativity, and empathy are increasingly in demand.

Highly developed skills such as the ability for abstraction, generalization, creativity and empathy are increasingly called for (OECD, 2016) as shown in the following figure:
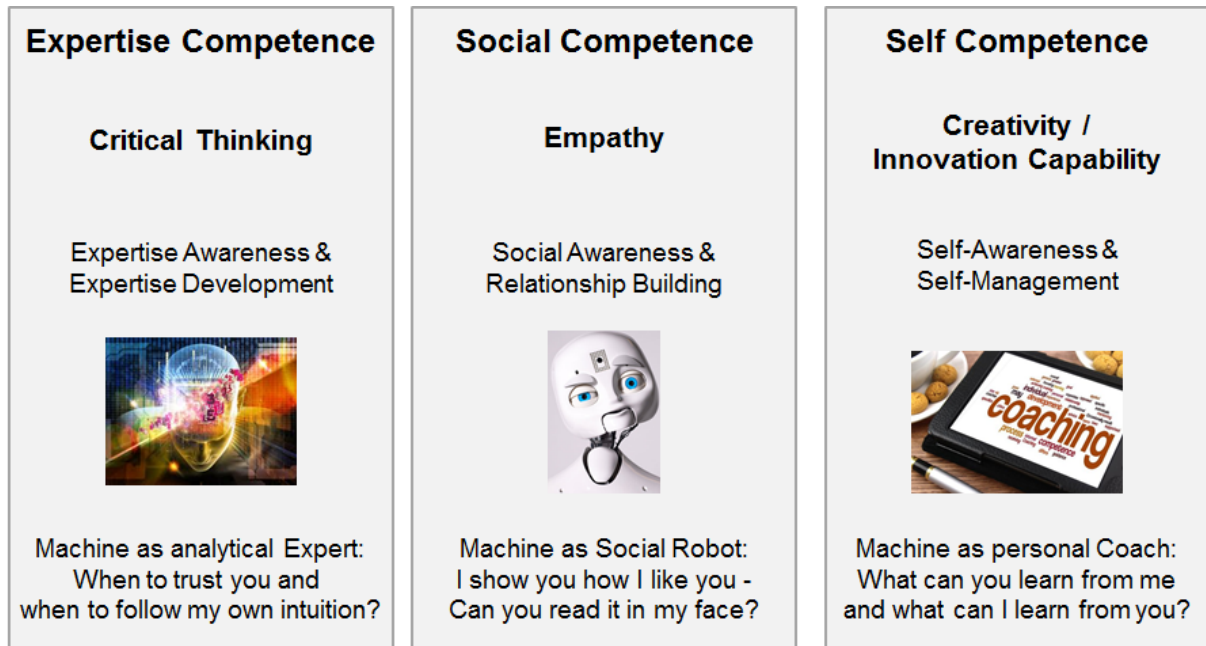

Figure 6.2: Complementary Core Competences of digital competences (as transversal competences) Source: Seufert 2017

The core competences represented above are interlinked with each other as all three areas relate to changing interactions with machines, cognitive systems and wearables of all kinds. However, for each of these three core competences several important sub-competences can be pointed out:

Expertise Competence: "Critical Thinking"
- Rethinking Research: Finding the right information in huge amounts of data in an efficient manner (e.g. by asking adequate questions based on a sound epistemological foundation).
- Decision Planning: Comprehensive presentation of alternatives and recommendations, with confidence levels and transparent sources (evidence-based).
- Discovery: Finding and identifying hidden connections, or recombining data from huge data spaces to create something new.

Social Competence: "Empathy"
- The capacity to place oneself in another's position. Empathy is seeing with the eyes of another, listening with the ears of another and feelings with the heart of another.
- Identification / evaluation of moral competence in social robots as an emerging category of collaborative machines (see e.g. http://www.hansonrobotics.com/robot/sophia/)
- Willingness and ability to enhance empathy through the interaction with social robots.

Self-Competence: "Creativity, Innovation Capability"
- Higher order learning competences, experimentation and reflection as metacompetences.
- Lateral thinking, creative thinking, divergent thinking, playful thinking.
- Dealing with uncertainty, risk taking, and rule breaking.
- Deliberate practice in the active maintenance of superior domain-specific performance (in spite of general age-related decline).

- New learning strategies in dealing with cognitive computing systems (e.g. interactive machine learning [HITL])

The competent use of big data in education depends on both, the competence of developers on the one side and the competence of users (learners) in the educational systems on the other side. In this context, information literacy is a major prerequisite for self-directed learning. The Stanford History Education Group (Stanford History Education Group, 2016) has analysed civic online reasoning via online assessments, focusing on "the ability to judge the credibility of information that floods young people's smartphones, tablets, and computers" The researchers described the information competence of the students just with one word – "bleak". They state: "we worry that democracy is threatened by the ease at which disinformation about civic issues is allowed to spread and flourish" (Stanford History Education Group, 2016, p. 5).

Helbing et al. (Helbing et al., 2017) come to the conclusion to practice fundamental principles in order to take the right decision at the digital crossroad. For that reason the researchers propose several fundamental principles, for example to support informational self-determination and participation and to promote responsible behavior of citizens in the digital world through digital literacy and enlightenment (Helbing et al., 2017).

For us, the challenge is how to transfer these fundamental principles to the educational system. For a start, we propose to clarify the normative orientation of learning. In doing so, we will employ an "old model", the concept of 'reflective practitioners'. Subsequently, we will clarify generic approaches to big data and learning analytics in order to promote the role model of a reflective practitioner as a role model in the digital world.

4) Normative Orientation: The reflective practitioner in the digital age

The concept of 'reflective practitioner' was proposed by Schön (Schön, 1983) and has since become well established. It highlights reflective abilities of individuals. Such reflective abilities are relevant, for example, in the process of preparing, realizing / delivering and controlling work activities. Beyond that, reflection may also pertain to oneself and also to the environment of work activities. Being reflective, therefore refers to conscious, critical and responsible evaluations of 1) one's own activities and competences (e.g., personal knowledge and results of task performance), 2) the personal process of competence development (e.g., formal and informal learning and the outcomes) and 3) the conditions for performing work (e.g., organizational structures or tools and resources available for task performance) (Dehnbostel, 2003, S. 42).

In the context of digital transformation, being a reflective practitioner refers to someone who reflects on her or his own competences (e.g., digital literacy), on relevant work environments (e.g., conditions and tools for remote collaboration) and work results (Dehnbostel, 2003).

In many corporate sectors, specific knowledge and specific skills have become a determining factor for the success of companies and organizations (North, 2011). Due to the dynamic nature of these changes, it is virtually impossible to predict which requirements and demands companies will face in the future, and which skills will be crucial for success. In consequence, traditional normative guidelines for the design, structure and promotion of learning and competence development need to be reformulated (Seufert & Diesner, 2010). It is important that critical reflection necessarily be ethical in the context of a radically technologized twenty-first century. While digital technology and the online world provide significant opportunities to education, these same opportunities can be leading to new risks and less self-control (Helbling et al., 2013).

The new points of departure for educational management can be developed on the basis of contrasts between "old work/old learning" in a remote-controlled society following the role model of a technocratic problem solver versus "new work/new learning" in a self-controlled society based on critical reflection and balanced ethics (see the following table):

Remote-controlled Society Technocratic Problem Solvers	Self-controlled Society Reflective Practitioners in the Digital Age
"Old Work/Old Learning"	**"New Work/New Learning"**
Work & Work Environment	
Stable and predictable environment with clearly defined organizational units	Unstable, dynamic and unpredictable environment and permeable organizational boundaries
Physical presence in the defined work environment	Real and virtual work environments as well as work in distributed / virtual teams
Paternalistic and transactional leadership	Transactional and transformational, meaningful leadership
Employees = learners who are passive ("consume" learning); "learning delivery"	Employees = learners who shape and co-produce; co-creation in work and learning
Modalities of Learning	
Learning by individuals	Learning by individuals, teams and organizations
Formally organized, "off-the-job" learning processes	Formal and informal learning "off / near / on-the- job"
Externally-controlled learning	Autonomous and self-directed learning
leveling of heterogeneity in the learning process	responding to and use of heterogeneity in the learning process
Point of Departure and Objectives for Learning	

Remote-controlled Society Technocratic Problem Solvers	Self-controlled Society Reflective Practitioners in the Digital Age
"Old Work/Old Learning"	**"New Work/New Learning"**
Learning targeted exclusively to current needs	Learning targeted to current as well as future demands and needs
Learning and knowledge sharing according to organizational guidelines (conveying answers)	Learning and knowledge exchange as part of the culture of a learning organization (enabling problem-solving)
Available content (e.g. courses / expertise) as starting point for training	Complex and real problems as starting point
Knowledge transfer and development of knowledge pool (behaviorist view of learning)	Developing skills and competence (cognitive-constructivist perception of learning)
Learning at specified times	Learning as and whenever needed
Learning effected by 'teachers';'Teachers' as "intermediaries"; instruction	Learning supported by 'teachers', managers, colleagues and media (instruction and construction)
Measurement of learning success	Measurement of the success of knowledge transfer and the impact of learning
Focus on utilization of existing knowledge / existing skills	Balance of utilizing existing knowledge / existing skills with the exploration of new knowledge / new skills

Table 6.1: Points of Departure for new working and learning environments (Seufert, 2013)

5) Analytics in Education: a framework and four cases

In the context of a more general shift from "old work / old learning" to "new work / new learning", learners – as reflective practitioners – need to take over more responsibility in managing the process of learning (Seufert et al., 2017). Accordingly, greater emphasis is placed on the importance of self-organization, self-control and self-determination. And there is a wealth of digital media available that support learners in planning, organizing, and controlling their own learning - ranging from task management via mind mapping and note taking to gathering community support for personal learning achievements (e.g. stikk.com).

While much of the interest in Big Data and Learning Analytics is currently focused on prediction, reflection (i.e., monitoring and understanding) may in fact become more widely relevant (Siemens & Gasevic, 2012; Gaviria et al., 2011). All the more so, as learners take on more responsibility in managing their own learning processes. However, those employing Learning Analytics (LA) applications need to be aware that "competing methods, technologies and algorithms applied to the same set of data, will

result in different outcomes, and thus may lead to different consequences in terms of decision making based on these outcomes" (Greller & Drachsler 2012, 50).

Employing LA in order to support learning, therefore, requires specific competences – on the part of institutions, teachers / facilitators, and learners. The following table provides an overview of relevant LA competences modeled on the taxonomy of cognitive process dimensions (Anderson & Krathwohl, 2001):

The Cognitive Process Dimension	Main Goal Reflection by learner	Main Goal Prediction (and prescription) by system
Remember Retrieve relevant information	Identify strategies to retain access to information and datasets	Information retrieval and integration of data from different sources
Understand Construct meaning from instructional messages	Interpret (visualized) data sets	Visualization of data sets
Apply Carry out / employ a procedure in a given situation	Use of techniques that match one's own strengths	Instructional messages, including feedback; Resource suggestions; Answers to frequently asked questions; Suggestions for novices
Analyze Break material into constituent parts and determine how parts relate to another and to an overall structure or purpose	Deconstruct own biases	Learner stratification Deviations from suggested paths
Evaluate Make judgments based on criteria and standards	Engage in self-assessment (e.g., reflect on personal progress)	Computer-based Assessment; Judgements based on learner profiles

Create Put elements together to create a new whole, reorganize into a new pattern or structure	Generate an innovative learning portfolio using personal data-sets	Generation of personalized learning paths; Algorithms for learner profiles

Table 6.2: A taxonomy for learning analytics activities

With regard to employing Big Data and Learning Analytics as a means to support (digital) learning, we propose a set of generic approaches based on a 2x2 matrix (Figure 3). One dimension is set up via the distinction of reflecting on past learning activities on the one hand versus predicting next / future learner activities on the other hand. Reflection here refers to critical self-evaluation on the basis of different data sets (Greller & Drachsler, 2012, p. 41):

- OWN datasets created in the process of learning or supporting learning (in the case of teachers respectively facilitators);
- Datasets created by OTHERS (e.g., a teacher reflecting on his or her own teaching style based on datasets generated by the students).

Prediction refers to anticipating learner activities (e.g., further reducing investment in classwork or discontinuing with classwork altogether). Prediction is a precursor of prescription and interventions that aim at dealing with a predicted event (e.g. a student discontinuing classwork) (Siemens, 2011).

The other dimension is set up via a distinction between learning activities by individual learners on the one hand and social learning activities on the other. Much work in LA is oriented towards supporting and determining individual achievement, for example by analyzing the data generated through summative assessments. Buckingham Shum and Ferguson (Buckingham Shum and Ferguson, 2012) 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 in school) "must account for connected and distributed interaction activity". Buckingham Shum & Ferguson therefore propose social learning analytics as a domain in its own right (Buckingham Shum and Ferguson, 2012).

Similar, gamification or gameful design for learning is considered as an on own domain (Deterding et al., 2013) using LA in social systems, for example to provide visible status and progress, social comparison and reputation (e.g. with badges).

The focus on individual learners is focused on the goal of personalization and individualization. In order to provide pedagogically valuable feedback assessment systems have to become intelligent and connected with higher-order learning skills. Adaptive learning systems (individual and prediction) represent an own, quite new research field based on interactive machine learning.

Objective for Learning Analytics

		Reflection	Prediction
Main Context and Target Group	Social	e.g. Social network analysis of students discussing in a forum (moderator tool) Illustration: Identify network connections between students and identify isolated students in order to facilitate their participation in the discussion.	e.g. Gameful design and data-driven rule sets for gaining reputation in a class Illustration: Identify visible status for social comparison and engage in an online community with data-driven incentive system.
	Individual	e.g. Digital formative assessment systems Illustration: Evaluate learning progress for self-reflection, visualize learning statistics, provide rapid feedback, and assist learners in developing meta-cognitive strategies.	e.g. Anticipatory & adaptive learning systems Illustration: Analyze learner profiles for automated decisions on facilitation activities, personalized learning pathways, and adaptive provisioning of learning resources.

Fig. 6.3: Generic strategies for approaches to learning analytics

In the following sections we will illustrate how this matrix framework can be translated into specific use cases:

- Social learning analytics for reflection

- Individual learning analytics for reflection

- Social learning analytics for prediction

- Individual learning analytics for prediction and prescription

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 et al. (Dawson, 2008; MacFayden & Dawson, 2010). This implies a shift in attention away from summative assessment of individuals towards learning analytics of social activity (Buckingham Shum & Ferguson 2012, p. 5). Here, it is relevant to distinguish between social analytics sui generi (e.g., social networks analysis or discourse analytics) from socialized analytics that are based in personal analytics while also 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 to this) (Buckingham Shum & Ferguson, 2012, p. 10-11).

The following example exemplifies the first type of social analytics sui generis:

Dimension	Exemplification
Objective	Reflection: Analyze student interactions in a forum discussion, identify network connections between students, and identify isolated students as a prerequisite for remedial action (aimed at helping these students to create links to others)
Digital competences required / to be developed	Interpretation: Do teachers / facilitators have the necessary competences to interpret and act upon the information available? Critical thinking: Are teachers / facilitators able to critically evaluate the data basis (e.g., missing data) when interpreting and / or devising a path of corrective action?
Contribution to fundamental principles of the digital agenda (Helbing et al., 2017)	- support informational self-determination and participation; - improve transparency in order to achieve greater trust in strategies for teaching / facilitation; - support social and economic diversity (i.e. success by diverse students); - improve interoperability and collaborative opportunities; - create digital assistants and coordination tools for the teacher; - support collective intelligence on the basis of visualizations of contributions and interactions.
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?

Table 6.3: Exemplary detailing of use case 1

Use case 2: Individual learning analytics for reflection

This use case is about LA with a focus on reflection at the individual level – for example about assessment results. As Evans (Evans, 2013) found out in a thematic analysis of the research evidence on assessment feedback in higher education (over 460 articles from 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 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).

Dimension	Exemplification
Objective	Reflection: Evaluate objective and subjective assessments; Identify knowledge gaps in order to support improved learning strategies (e.g., preparation for an exam); Provide opportunities for active learning during/ after lectures in order to evaluate the impact of teaching.
Digital competences required / to be developed	Students: self-assessment competences; metacognitive learning strategies. Teachers: scaffolding competences (help students to interpret the data).
Contribution to fundamental principles of the digital agenda (Helbing et al., 2017)	- increasingly decentralize the function of information systems; - support informational self-determination and participation; - improve transparency in order to achieve greater trust in strategies for teaching / facilitation; - reduce the distortion and pollution of information; - enable user-controlled information filters; - create digital assistants (for students); - promote responsible behavior.
Constraints	Privacy: Is anonymity (hiding of 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?

	Time scale. Should the analyses be carried out in-class or outside of class (trade-off with time required for teaching time)?

Table 6.4: Exemplary detailing of use case 2

Use case 3: Social analytics for prediction

The more environments for working and learning are becoming digital, the more data is generated in the course of 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 has 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 & Thomas, 2013). By this we refer to the use of game design elements in non-game contexts (Dixon, Khaled & Nacke, 2011, p. 10). 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 some 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; Sailer et al., 2013).

The following use case focuses on gamified learning designs:

Dimension	Exemplification
Objective	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 engagement and activity in learning in general. Predict which learners will respond to incentives by displaying more desired behaviors (e.g., engagement / activity in the course).
Contribution to fundamental principles	- improve transparency in order to achieve greater trust in strategies for teaching / facilitation (on the part of teachers / facilitators);

of the digital agenda (Helbling et al., 2017)	- improve collaborative opportunities (through targeting students at risk and – hopefully – retaining them in class); - create digital assistants and coordination tools (for teachers / facilitators).
Digital 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: Realistic estimates of ability and motivation of learners when creating a gamified learning design. Teachers / Facilitators: Ability to interpret (visualizations of) levels of student activity.
Constraints	Privacy: What are relevant authentication & data security issues when points earned for gamified activities are feed into final grades? Ethics: What are dangers of abuse/ misguided use of a data-driven rule system? Norms: Course gamification could be misused for selling old designs in new terminology, for example, by renaming assignments to quests and scores to 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)?

Table 6.5: Exemplary detailing of use case 3

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 standard expository teaching techniques in classrooms with about 30 learners (Bloom, 1984). The idea of individualized tuition for large numbers of learners is currently pursued in the context of research and development of adaptive or intelligent tutorial platforms (Romero et al., 2008) which in turn 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 the (repetitive) solution of problems as

was the case in prior generations of so-called intelligent tutorial systems (Bagheri, 2015). Adaptive Learning Systems closely track student activities and student performance and, based on machine learning algorithms and predictive models, provide students with adequate learning pathways and adaptive learning resources (Butz, Sigaud & Gerard, 2003). However, more substantial empirical research is needed to investigate, in particular, the appropriateness of such algorithms in disciplines other than the typical mastery learning subjects (e.g. biology, mathematics, and information science) and the effectiveness for reaching higher learning outcomes.

The following use case focuses on adaptive learning designs:

Dimension	Exemplification
Objective	Prediction based on student model / learner profiles and prescription of next learning activities in order to facilitate comprehension and retention. Achieve learning outcomes more efficiently (and possibly also outcomes at higher cognitive levels) through continuous analysis and guidance in the learning processes.
Contribution to fundamental principles of the digital agenda (Helbing et al., 2017)	- improve transparency in order to achieve greater trust in strategies for teaching / facilitation (on the part of both learners and teachers); - support collective intelligence.
Digital 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	Problems may be caused by poor models. Sensitivity, spurious correlations, meaningless patterns, noise and classification errors (all very common problems in Big Data analytics) Data manipulation. 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 abilities regarding monitoring and planning their own learning?

Table 6.6: Exemplary detailing of use case 4

6) Conclusion and outlook

In this chapter, we started out taking a wider perspective on big learning data and learning analytics. Against a backdrop of alternative scenarios for a second machine age and the possibilities of software obliterating management as a profession and a field for education, we have pursued the issue of what it means to be competent in digital learning – specifically in the use of big learning data and learning analytics. We have taken note of scenarios for the use of AI that may lead from "programming computers to programming people". And we have pointed out how it is important for humans to be in the loop and to provide judgement when it comes to working with algorithms and AI-systems. We have pointed out that in collaborating with powerful information systems and intelligent machines we need to focus on our core competences as humans: critical thinking, empathy and creativity / capability for innovation. And we have alerted to a set of principles that should guide how we design and work with information systems – such as big data and analytics systems in education.

The discussion on big data in education is mostly focused on the potential of learning analytics to increase the efficiency and effectiveness of educational processes. A classic case is the endeavor to identify and support students at risk in order to reduce drop-out rates. Accordingly, prediction is in focus while the potential for supporting reflection on learning is neglected. We proposed the concept of 'reflective practitioner' as a guiding normative principle for both educators and (continuous) learners. And we pointed out that in the context of a general shift from "old work / old learning" to "new work / new learning" learners – as reflective practitioners – need to take over more responsibility in managing the process of learning. Building on both these concepts, 'reflective practitioner' and 'new work / new learning', we provided a rough overview on what competences are required when developing learning analytics for reflection and prediction at different levels of cognitive processes.

In a next step we proposed a 2x2 matrix for learning analytics, differentiating 'reflection' and 'prediction' as relevant objectives and also the use in 'individual' and 'social' learning. The four fields set apart in this matrix we have subsequently illustrated through four use cases: 1) social learning analytics for reflection; 2) individual learning analytics for reflection; 3) social learning analytics for prediction; 4) individual learning analytics for prediction and prescription. For each use case, we have set out the objective, the digital competences required, the contribution to principles of designing / working with information systems, and, last but not least, relevant constraints.

When it comes to supporting learners as reflective practitioners through analytics, it is not only important to enable them in their reflection on the different cognitive process dimensions we have pointed out (remembering, understanding, applying, etc.) – and the implications for their own study behaviors and strategies. As developed at the beginning of this chapter, it is also important to alert learners to larger issues related to machine learning, augmentation and autonomy. Via small "nudges" – but on massive scale – we as citizens (and learners) are steered towards healthier, safer and more environmentally friendly behavior in many domains: when selecting a menu as well as when driving cars (Helbing, 2017). Learners, i.e. all of us, need to be aware of the possible impact such nudging may have on our acting as reflective practitioners. Big learning data and learning analytics should help us find the way to our own (learning) goals. This is the support that learners should expect.

In order to deal with these issues, future research should focus on empirical evaluation methods of learning analytics tools (Ali et al., 2012; Scheffel, Drachsler, Stoyanov & Specht, 2014) and on competence models for digital learning (Dawson & Siemens, 2014). The LA taxonomy proposed in this paper provides a (small) starting point for modelling required skills and attitudes as the needed implementation requirements to guarantee successful exploitation of learning analytics. 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 & Siemens, 2014), 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).

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.

Anderson, L. & Krathwohl, D. A. (2001). *Taxonomy for Learning, Teaching and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. New York: Longman.

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

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). doi: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., & 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 Butz, M.V.; Sigaud, O. & P. Gerard (Eds). *Anticipatory Behavior in Adaptive Learning Systems*. Volume 2684 of the series Lecture Notes in Computer Science (pp 86-109). Berlin, Heidelberg: Springer.

Dawson, S. (2008). A study of the relationship between student social networks and sense of community, *Educational Technology & Society, 11*(3), 224-38.

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). Retrieved from http://www.irrodl.org/index.php/irrodl/article/view/1878/3006

Dehnbostel, P. (Eds.). (2003). *Perspektiven moderner Berufsbildung. E-Learning, didaktische Innovationen, modellhafte Entwicklungen.* Bielefeld: Bertelsmann.

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, NY: ACM.

Dicheva, D., Dichev C., Agre G., & Angelova G. (2015). Gamification in Education: A Systematic Mapping Study. *Educational Technology & Society, 18*(3), 75–88.

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.

European Union (2006). RECOMMENDATION OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 18 December 2006 on key competences for lifelong learning (2006/962/EC). Retrieved from http://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex%3A32006H0962

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

Gaviria, F., Glahn, C., Drachsler, H., Specht, M., & Gesa, R. F. (2011). Activity-based learner-models for learner monitoring and recommendations in Moodle. In C. D. Kloos et al. (Eds.), *Proceedings*

of the 6th European Conference on Technology-Enhanced Learning (pp. 111-124). Heidelberg, Berlin: Springer.

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society, 15*(3), 42–57.

Helbing, D.; Frey; B. S.; Gigerenzer, G.; Hafen, E.; Hagner, M.; Hofstetter, Y.; van den Hoven, J.; Zicari, R. V. & Zwitter, A. (2017). *Will Democracy survive Big Data and Artificial Intelligence?* Retrieved from https://www.scientificamerican.com/article/will-democracy-survive-big-data-and-artificial-intelligence/

Helbing, D. (2017). From remote-controlled to self-controlled citizens. *The European Physical Journal Topics, 226*, 313-320.

Helbing, D. (2015). *Thinking Ahead: Essays on Big Data, Digital Revolution, and Participatory Market Society*. Springer: Cham Heidelberg New York Dordrecht London.

Holzinger, A. (2016). Interactive machine learning for health informatics: when do we need the human-in-the-loop? *Brain Informatics*, 2016, 3 (2), pp. 119-131. doi:10.1007/s40708-016-0042-6

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. Retrieved from http://dx.doi.org/10.18608/jla.2015.22.3

Leonard, T. C. (2008). *Book Review: Richard H. Thaler, Cass R. Sunstein, Nudge: Improving decisions about health, wealth, and happiness.* http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.364.5470&rep=rep1&type=pdf

MacFayden, L.P., and Dawson, S. (2010). Mining LMS data to develop an "early warning system" for educators: A proof of concept, *Computer & Education 54*(2), 588-599.

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, Germany: Springer.

OECD (2016). *Skills for a digital world. 2016 Ministerial Meeting on the digital economy*. Background Report. OECD Publishing (OECD Digital Economy Papers, 250).

Pariser, E. (2011). *The Filter Bubble. What The Internet Is Hiding From You*. Penguin Books.

Pasquale, F. (2015). *The Black Box Society: The Secret Algorithms That Control Money and Information*. Harvard University Press.

Patterson, R. W. (2015). Can behavioral tools improve online student outcomes? Experimental evidence from a massive open online course (Report). *Cornell University, Department of Policy Analysis and Management*. Accessed via https://www.ilr.cornell.edu/sites/ilr.cornell.edu/files/cheri_wp165_0.pdf

Rahwan, I. (2016). *Society-in-the-Loop – Programming the Algorithmic Social Contract.* Blogpost retrieved from https://medium.com/mit-media-lab/society-in-the-loop-54ffd71cd802#.byd1hcygm.

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

Sailer, M., Hense, J., Mandl, H. & Klevers, M. (2013). Psychological Perspectives on Motivation through Gamification. *Interaction Design and Architecture(s) Journal, 19*, 28-37.

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

Schön, D. A. (1983). *The Reflective Practitioner. How professionals think in action*. London: Temple Smith.

Schulz von Thun, F. (n.d.). *Six tools for clear communication*. Available via http://www.schulz-von-thun.de/index.php?article_id=172

Seufert, S. (2017). *Digital Competences. Position Paper for the Swiss Sciences and Innovation Council*. St.Gallen: University of St.Gallen.

Seufert, S. & Diesner, I. (2010). Wie Lernen im Unternehmen funktioniert. *Harvard Business Manager,* August 2010, 2–5.

Seufert, S. & Vey, K. (2016). Hochschulbildung 2030. Humboldt im digitalen Zeitalter. *Neue Züricher Zeitung*, September 9, 2016.

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.

Seufert, S.; Meier, C.; Schneider, C.; Schuchmann, D.; Krapf, J. (2017): Geschäftsmodelle für inner- und überbetriebliche Bildungsanbieter in einer zunehmend digitalisierten Welt. In: Erpenbeck / Sauter (Eds.): Handbuch Kompetenzentwicklung im Netz. Bausteine einer neuen Lernwelt. Stuttgart: Schäffer-Poeschel, pp. 429–447.

Siemens, G. (2011). *Learning analytics: A foundation for informed change in higher education*. Retrieved from http://www.slideshare.net/gsiemens/learning-analytics-educause

Siemens, G., & Gasevic, D. (2012). Guest Editorial - Learning and Knowledge Analytics. *Educational Technology & Society, 15*(3), 1–2.

Stanford History Education Group (2016): *Evaluating Information: The Cornerstone of Civic Online Reasoning*. Executive Summary. Retrieved from https://sheg.stanford.edu/upload/V3LessonPlans/Executive%20Summary%2011.21.16.pdf

Thaler, R. H. & Sunstein, C. R. (2008). *Improving Decisions about Heath, Wealth and Happiness*. Yale: Yale University Press.

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.

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). Retrieved from https://olj.onlinelearningconsortium.org/index.php/olj/article/view/783