Drachsler, Hendrik

Towards highly informative learning analytics

Heerlen : Open Universiteit 2023, 62 S.

Quellenangabe/ Reference:
Drachsler, Hendrik: Towards highly informative learning analytics. Heerlen : Open Universiteit 2023, 62 S. - URN: urn:nbn:de:0111-pedocs-267875 - DOI: 10.25656/01:26787

https://doi.org/10.25656/01:26787

Prof. dr. Hendrik Drachsler

Towards Highly Informative Learning Analytics

Open Universiteit

Open Universiteit

Prof. dr. Hendrik Drachsler

Towards Highly Informative Learning Analytics

Cite as:
Drachsler, H. (2023). Towards Highly Informative Learning Analytics.
Open Universiteit, Netherlands, Heerlen. https://doi.org/10.25656/01:26787

Editorial:
Dr. Joshua Weidlich I DIPF Leibniz Institute
John Arkenbout I Open Universiteit

Cover design:
Jeroen Berkhout I Open Universiteit

Layout:
Dennis Hendriks I ProefschriftMaken

Printing:
ProefschriftMaken I proefschriftmaken.nl

ISBN: 978-94-6469-372-0 DOI: https://doi.org/10.25656/01:26787

TABLE OF CONTENTS

List of abbreviations:

AI	Artificial Intelligence
AIED	Artificial Intelligence in Education
CSCL	Computer Supported Collaborative Learning
DeLA	Data enriched Learning Activities
ECD	Evidence Centred Design
EDM	Educational Data Mining
EFLA	Evaluation Framework for Learning Analytics
$FoLA^2$	Fellowship of the Learning Activity Learning Analytics
GDPR	General Data Protection Regulation
HILA	Highly Informative Learning Analytics
IRR	Inter-Rater Reliability
KPI	Key Performance Indicators
LA	Learning Analytics
LMS	Learning Management System
OpenLAIR	Open Learning Analytics Indicator Repository
SELAQ	Student Expectations of Learning Analytics Questionnaire
SIS	Student Information System
STEM	Science, Technology, Engineering, and Mathematics
TLA	Trusted Learning Analytics

List of figures:

Towards Highly Informative Learning Analytics

Rede

In verkorte vorm uitgesproken bij de openbare aanvaarding van het ambt van hoogleraar Learning Analytics bij de Open Universiteit

op vrijdag 12 mai 2023

door prof. dr. Hendrik Drachsler

1 INTRODUCTION

Among various trending topics that can be investigated in the field of educational technology, there is a clear and high demand for using artificial intelligence (AI) and educational data to improve the whole learning and teaching cycle. This spans from collecting and estimating the prior knowledge of learners for a certain subject to the actual learning process and its assessment. AI in education cuts across almost all educational technology disciplines and is key to many other technological innovations for educational institutions.

The use of data to inform decision-making in education and training is not new, but the scope and scale of its potential impact on teaching and learning have silently increased by orders of magnitude over the last few years. During the COVID-19 pandemic, many things that were not possible in the past due to data security, privacy, quality, or other higher goods were thrown overboard, and a vast amount of society experienced the full potential but also the shortcomings of digital education (Drachsler et al., 2021).

The release of ChatGPT was another driver to finally make everyone aware of the potential effects of AI technology in the digital education system of today. We are now at a stage where data can be automatically harvested at previously unimagined levels of granularity and variety. Analysis of these data with AI has the potential to provide evidence-based insights into learners' abilities and patterns of behaviour that, in turn, can provide crucial action points to guide curriculum and course design, personalised assistance, generate assessments, and the development of new educational offerings.

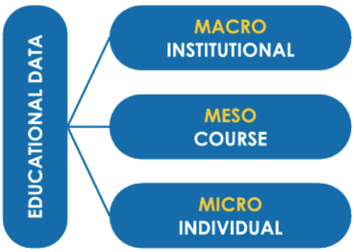
AI in education has many connected research communities like Artificial Intelligence in Education (AIED), Educational Data Mining (EDM), or Learning Analytics (LA). All these communities research questions from overlapping research domains such as psychology, education, computer and data science. The field of Learning Analytics (LA) emerged over 15 years ago and was strongly driven by higher education institutions. Recently, we have seen LA become an established field in the higher education sector as many universities established central LA units[1], as well as various policy documents[2] (Hansen et al., 2020) and study programmes are available[3].

LA is the term that is used for research, studies, and applications that try to understand and support the behaviour of learners based on large sets of collected data. A common definition that is used by the field itself was coined by the organisers of the first International Conference on Learning Analytics and Knowledge in 2011: "*LA is the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs.*"

1 https://cic.uts.edu.au/
2 https://www.jisc.ac.uk/guides/code-of-practice-for-learning-analytics
3 https://www.gse.upenn.edu/academics/programs/learning-analytics-online-masters

LA can provide different levels of insights as demanded by Reich (2015): either it is provided on the level of a single course, on the level of a collection of courses, or on the level of a whole curriculum. Buckingham Shum (2012) therefore introduced the notion of micro-, meso-, and macro-levels to distinguish the role that LA can play on different abstraction levels (see figure 1). The micro-level mainly addresses the needs of individuals, e.g., learners and instructors within a course; the meso-level addresses a collection of courses and provides information for course managers; the macro-level takes a bird's eye view on a directory of courses and can provide insights for a whole community by monitoring learning behaviour across courses and even across different scientific disciplines. Depending on which level LA is utilised, different objectives and information are of relevance and can be monitored. In the case of LA dashboards, for instance, the micro-level addresses the individual learners in a single course, whereas dashboards on the meso-level are often aimed at teachers and instructors. These dashboards inform them about a whole class's learning status and provide support in performing teacher's roles effectively in areas including class management, learning facilitation, provision of feedback, and evaluation and grading. The macro-level is also often connected to the strategic goals of a whole organisation and is reported according to the core Key Performance Indicators (KPIs) of an educational organisation. LA can extend traditional KPIs with new metrics that can be computed from data and reported next to already established KPIs. For instance, with process mining the most frequent learning paths of the learners can be visualised, grouped by specific cohorts of learners, and evaluated according to effectiveness and efficiency of learning outcomes. Those paths can show obstacles in a curriculum that might lead to structural problems for many children in a school and can therefore have an impact back to the meso- and micro-level.

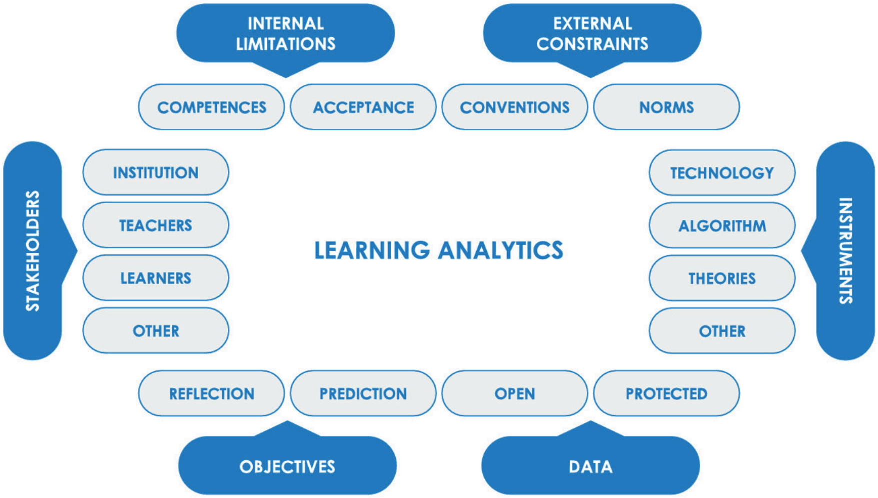
Figure 1. Three levels of educational data for Learning Analytics.



2 FOUNDATIONS OF TRUSTED LEARNING ANALYTICS

A comprehensive introduction to the different domains that are affected by LA has been provided by Greller and Drachsler (2012). They present the technological and educational aspects of LA in six dimensions. The mutual combination of these six dimensions forms the concept of Trusted Learning Analytics (TLA). In the following section, the TLA concept will be summarised according to its six dimensions (see figure 2) (Stakeholders, Objectives, Data, Instruments, External constraints, and Internal limitations) and guiding questions for the implementation of TLA will be provided.

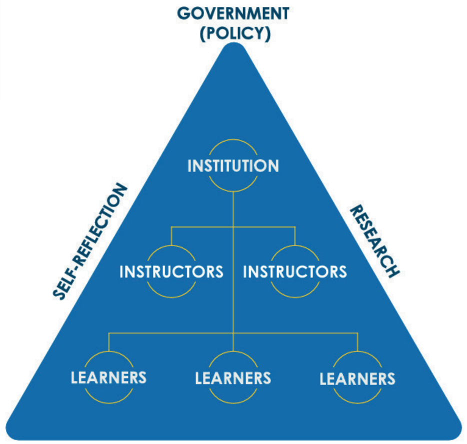
Figure 2: Holistic Learning Analytics framework according to Greller and Drachsler (2012).



2.1 Stakeholders: contributors and beneficiaries of learning analytics

The stakeholder dimension includes data clients as well as data subjects. Data clients are the beneficiaries of the LA process who are entitled and meant to act upon the outcome (e.g., learners & teachers). Conversely, the data subjects are the suppliers of data, normally through their browsing and interaction behaviour in digital learning environments. Those roles can change depending on the objective of the analytics at the meso- and macro-level. Information flow between different stakeholders can best be exemplified by the common hierarchical model taken from formal education (see figure 3). The diagram serves as an example of how benefits might be obtained from LA. The pyramid encapsulates the academic layers of educational and training institutions. Most directly, data analysis at the learner level, e.g., via a Learning Management System (LMS), can inform the above layer, in this case, the instructors. Instructors can then use the analytics information to plan targeted interventions or adjust their educational strategies. Similarly, institutions can retrieve benefits from learners' and instructors' data to provide staff development opportunities or to plan policies like quality assurance and efficiency measures. Finally (on top of the diagram), government agencies may collect cross-institutional data to assess the requirements of educational institutes and their constituencies.

Figure 3: Common hierarchical model for formal education institutions.



Educational technology requires a strong stakeholder needs analysis in order to be successful in the long run (Drachsler & Greller, 2012). This especially applies to LA solutions, as data sources are often highly diverse and the educational systems of various countries can be equally heterogeneous. A one-size-fits-all approach is therefore pointless; instead, a sophisticated methodology to identify KPIs is needed that can be transferred into technology innovations in learning and teaching. Examples of such requirement engineering research have been conducted by various researchers (Drachsler & Greller, 2012; Drachsler, Stoyanov & Specht 2014; Whitelock-Wainwright et al., 2017; Tsai et al., 2018; Scheffel et al., 2019; Kollom et al., 2020; Wollny et al., 2023).

Guiding questions for the LA stakeholders are:

- Micro-level:
  How can the LA solution help to meet the specific needs of individual stakeholders like teacher and students in their lectures?

- Meso-level:
  In what ways can university-based vocational training programs offering professional teacher education to educate to more agency in applying LA for teaching and learning?

- Macro-level:
  How can organisational KPIs be identified and translated into innovations in learning and teaching across educational institutions?

2.2 Objectives: set goals that LA applications aim to support

The primary potential for LA consists of revealing and contextualising previously concealed information from educational data and preparing it for the various stakeholders. This new type of information can enhance micro-level individual learning and teaching procedures. Here, we focus mostly on reflection and generating predictions as the two primary goals. Reflection happens when a student is taking a course and the LMS collects data on the progress, such as how much time is spent on each module, and the assignment results. LA can provide the student with feedback, visualisations and summaries of progress that can be used to reflect on the learning process and outcome. The same LMS can also use LA to predict how well the student is likely to perform in the course. LA can analyse the student data in comparison to the performance of other students who have taken the course and identify patterns and trends. For example, LA can see that students who spend more time on the assignment tend to perform better on the final exam. Based on this pattern, the LMS can predict the student is likely to perform well on the final exam because the student has spent more time on the assignments than the average student. LA can also identify students who have struggled in the past with similar assignments and suggest resources to help the student to master the assignment. This also leads to more individualised learning possibilities, including training and individualised learning routes leading to a learning objective. On a meso- or macro-level, however, the objectives shift and become increasingly focused on organisational knowledge management, with a particular emphasis on benchmarking educational techniques and interventions.

Guiding questions for the LA objectives are:

- Micro-level:
  How does the use of LA impact individual student reflection and prediction of their performance in a course?

- Meso-level:
  How can the use of LA support the generation of most effective learning paths through a study programme or curriculum?

- Macro-level:
  What are the benefits and challenges of implementing LA to support organisational knowledge management, particularly in benchmarking educational techniques and interventions across different institutions?

2.3 Data: educational datasets and the environment in which they occur

The technological ecology of any established educational institution can be rather varied and likely spans a myriad of technologies with varying and occasionally overlapping functions. Numerous organisations adequately secure data generated within the institution from external access or usage, especially after the establishment of the General Data Protection Regulation (GDPR) in 2018. As a result of the quick transition to remote instruction during the COVID-19 pandemic, many schools and their governing organisations made judgement calls about the use of platforms and technologies without adequately considering data privacy concerns. Consequently, they are currently seeking further ways to secure data from external access or use. Typically, LA uses data from several technical systems. In order to meaningfully integrate the data, it must be examined and linked (Berg et al., 2016). Alternatively, data can be merged by immediately streaming it from relevant source systems into a so-called "data lake". A data lake is a type of database that pulls important data from several other databases and makes it accessible to the organisation's authorised personnel (Ciordas-Hertel et al., 2019). In summary, while the technological landscape of educational institutions can be complex and diverse, data privacy concerns are increasingly important to consider, especially in the age of remote instruction. To effectively leverage learning analytics, it is necessary to examine and link data from various technical systems to integrate the data in a more streamlined manner.

Guiding questions for the LA data are:

- Micro-level:
  What digital learning technologies does your organisation employ in the daily classroom and does these provide relevant learning data?

- Meso-level:
  What are the best practices for integrating and linking data from various technical systems to create a comprehensive and accessible educational data lake for LA purposes?

- Macro-level:
  What policies and regulations can be implemented at an institutional or national level to ensure the ethical and responsible use of data in educational institutions?

2.4 Instruments: technologies, algorithms, and theories that carry

Different instruments can be applied in the development of educational services and applications that support the objectives of the different educational stakeholders (Drachsler et al., 2014). LA takes advantage of diverse technologies such as machine learning, social network analysis, or classical statistical analysis techniques in combination with visualisation techniques. Ciordas et al. (2019) investigated infrastructures to enable the harvesting of LA interventions, Scheffel et al. (2017) researched LA applications for group collaboration, Jivet et al. (2020) investigated LA dashboards for individual learners, Schneider et al. (2017) and Di Mitri et al. (2018) investigated tutorial systems for psychomotor learning like presentation skills.

Guiding questions for the LA instruments are:

- Micro-level:
  To what extent does the use of machine learning and visualisation techniques in LA prove to be effective in delivering tailored assistance for enhancing individual competencies?

- Meso-level:
  How can educational institutions effectively integrate LA technologies into their existing systems and infrastructure to improve student retention rates and academic success?

- Macro-level:
  How can policymakers and educational leaders leverage LA technologies and theories to develop evidence-based interventions to address issues of equity and access in education?

2.5 External Constraints: restrictions or potential limitations for anticipated benefits

The large-scale production, gathering, aggregation, and processing of data from educational programmes has raised ethical and privacy issues over the potential for harm to individuals and society. Examples such as the closure of inBloom in the United States because of privacy issues regarding LA and big data in education demonstrate how pertinent such concerns are (Singer, 2014). InBloom, a $100 million LA effort financed mostly by the Bill & Melinda Gates Foundation, sought to enhance American schools by providing a centralised platform for data sharing, learning apps, and courses.

Since then, ethics and privacy have been a continuing focus of study, with numerous articles published on the subject. An article by Drachsler and Greller (2016) explores the most prevalent fears and proposals for privacy and ethics and ends with a DELICATE eight-point checklist that academics, policymakers, and institutional management can use to support a trustworthy implementation of LA. The Open University UK[4] the University of Edinburgh[5], and the Goethe University Frankfurt have produced policies

4 http://www.open.ac.uk/students/charter/essential-documents/ethical-use-student-data-learning-analytics-policy
5 https://www.ed.ac.uk/information-services/learning-technology/more/learning-analytics

and guidelines concerning trustworthy LA[6] containing privacy, legal protection rights, and ethical consequences for data-driven organisations (Hansen et al., 2020). Although basic guidelines and solutions for ethics and privacy are presented, fundamental research problems remain, and novel technical solutions are required to answer the questions listed below (see figure 4).

The GDPR went into effect in the European Union in May 2018 and is progressively being adopted by other state laws. It advocates the data protection-by-design and privacy-by-default principles. For educational contexts, the following aspects are central:

- Right to restrict processing
- Right to data portability
- Right to object
- Right related to automated decision making and profiling
- Accountability and governance
- Breach notification
- Transfer of data
- Data protection by design and by default

The DELICATE checklist (see figure 4) aims to stimulate a trusted way of applying LA by considering privacy and ethics in LA at the outset of any LA project (Drachsler & Greller, 2016).

Guiding questions for the LA constraints are:

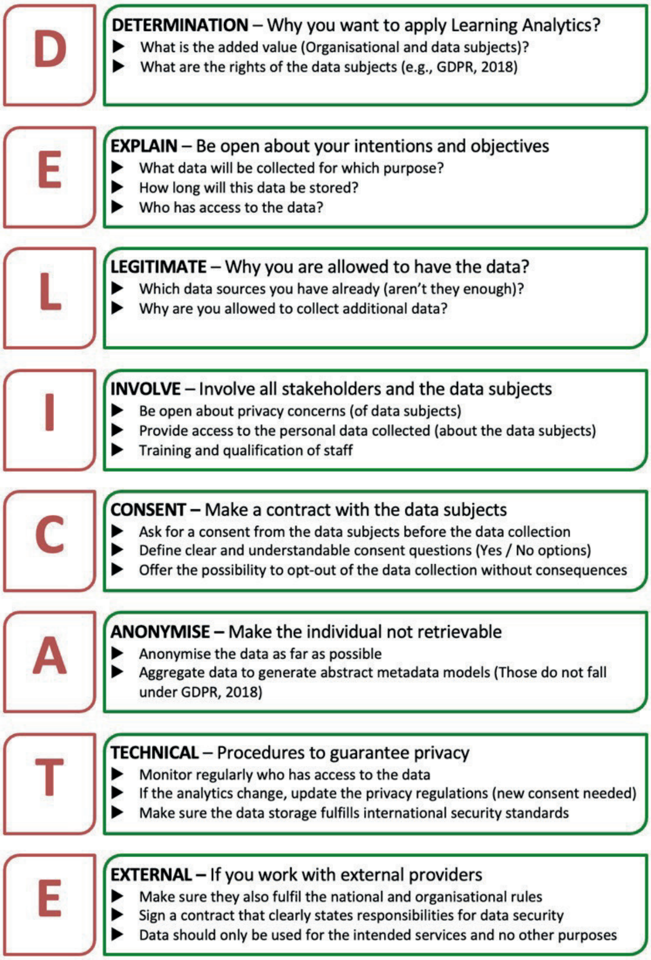
- Micro-level:
  What are the ethical and privacy concerns of students and teachers when using LA tools in the classroom?

- Meso-level:
  What potential biases exist in the data we are using, and how can we manage these biases to ensure fair treatment of heterogenous student groups?

- Macro-level:
  How can institutions ensure that their LA practices adhere to ethical and privacy principles, such as those outlined in the GDPR and the DELICATE checklist?

6 https://www.researchgate.net/publication/340183667_Verhaltenskodex_fur_Trusted_Learning_Analytics

Figure 4: DELICATE checklist to establish TLA according to GDPR adjusted from Drachsler and Greller (2016).



2.6 Internal Limitations: user requirements to exploit the benefits

To make LA a useful tool for educational institutions, it is crucial to recognise that LA does not end with the presentation of algorithmically obtained findings. These outcomes must be interpreted by educational stakeholders. Consequently, the exploitation of LA necessitates a number of advanced competencies, including information-, feedback-, and data literacy. The stakeholders must be informed on how to correctly interpret and utilise the information. These competencies are not yet represented in teacher education programmes of many European nations. It is thus a challenge for the adoption and implementation of LA to enhance the skills of educational stakeholders', primarily teachers, teacher trainers, students, and policymakers. Consequently, LA should also play a key role in teacher training in order to enhance these competencies. In addition to interpreting and acting on data, educational stakeholders should be involved in the design of individualised LA apps. In order for educational stakeholders to use the LA tools effectively in the future, the makers of the tools must supply them with the information they desire. Using instruments such as the Evaluation Framework for LA (EFLA[7]) created by Scheffel et al. (2017), or the Student Expectations of LA Questionnaire (SELAQ[8]) developed by (Whitelock-Wainwright et al., 2020; Wollny et al., 2023) can thus be a useful source of information to gain this user perspective. Finally, the FoLA[9] method can be used for the collaborative design of LA with practitioners in the field.

Guiding questions for the LA limitations are:

- Micro-level:
  How can you train teacher and students to enhance the information-, feedback-, and data-literacy competencies for effective utilisation of LA?

- Meso-level:
  What are the factors that hinder the adoption and implementation of LA in your institutions, and how can these factors be effectively addressed?

- Macro-level:
  How can policymakers ensure that the very dynamic evolvement of LA tools incorporates national and institutional policies?

3 FROM TRUSTED TO HIGHLY INFORMATIVE LEARNING ANALYTICS

The origin of LA was driven by the vision of exploring learning by tracing the footprints learners leave in digital learning environments. Because, unlike assessments, LA is not separate from normal learner behaviour, the information retrieved is highly authentic as it reflects actual, continuous learner behaviour. Greller and Drachsler (2012) suggest that TLA is more akin to data collected via observation than surveys or assessments. It may even have the potential to (partly) replace traditional assessments with the analytics information, by predicting learning success from data.

7 https://core.ac.uk/download/pdf/95759389.pdf, on page 138
8 https://www.researchgate.net/publication/338587740_Assessing_the_validity_of_a_learning_analytics_expectation_instrument_A_multinational_study
9 https://www.fola2.com/

The concept of TLA as provided in Section 2 has been intensively studied in the last five years and has made significant progress on social, technical, and educational frontiers.

On the social dimension, ethical guidelines (Drachsler & Greller, 2016; Hansen, et al., 2020), policies (Scheffel, et al., 2019), empirical studies with educational stakeholders (Biedermann et al., 2019; Kollom et al., 2020; Jivet et al., 2020; Wollny et al., 2023) and instruments (Tsai et al., 2018) for the ethical use of TLA in Europe have been made.

On the technical dimension, infrastructures (Ciordas-Hertel et al., 2021; Ciordas-Hertel et al., 2022a; Ciordas-Hertel et al., 2022b; Biedermann et al., 2023b; Wollny et al., 2021; Karademir et al., 2021) to record, process, and analyse data about how students learn in digital learning environments have been developed and deployed. Based on these outcomes, a mature data infrastructure has been established that opens the door for more rigorous empirical studies on the effects of TLA.

On the educational side, several field studies have been done to see how data from TLA research affects teacher adoption (Kollom et al., 2020) and how data-driven feedback models affect students (Jivet et al., 2020; Di Mitri et al., 2021; Gombert et al., 2022). Based on what we have learned from the TLA research programme, we have come up with three focus areas for the current research programme on "Highly Informative Learning Analytics" (HILA).

1 Focus on the micro level where learning and teaching happens

In the field of LA, there are frequent calls to remember that LA should be deeply grounded in the learning sciences (Greller & Drachsler, 2012; Gasevic et al., 2015; Motz et al., 2022). Gasevic et al. (2015) therefore refers to other well-established research communities like information seeking that have already reached a stage of scientific maturity: "As a developing field in information seeking, Wilson (1999, p. 250) noted that "*many things were counted, from the number of visits to libraries, to the number of personal subscriptions to journals and the number of items cited in papers. Very little of this counting revealed insights of value for the development of theory or, indeed, of practice*. *Significant progress in research and practice only really commenced when information seeking was framed within robust theoretical models of human behaviour*" (Wilson, 1999, p. 250)" Gasevic et al. (2015 p.6).

Jivet et al. (2017, 2018) also demanded that the field of LA needs to focus more on "learning" rather than putting the emphasis on "analytics". Because focusing on the analytics part of LA often results in working with the data that is readily available in some log files, instead of designing meaningful and rich data to support the actual learning goals and competence development of the students. Most recently, a pre-registered study on that focus already shows in their results that, in an extensive sample of research publications from the proceedings of two recent LAK conferences, 70.5% of papers do not provide any measure of learning and 91.4% do not attempt an educational intervention based on analytics (Motz et al., 2022). The authors conclude that there is not a clear direction towards supporting learning in LA research right now.

With HILA research programme, we want to close this gap and focus on giving students and teachers, in particular, meaningful feedback for their learning. With this, HILA shifts the focus away from the meso- and macro-levels to focus entirely on the micro-level of LA to provide effective, efficient, and enjoyable learning experiences to students and teachers. This focus is also urgently needed as many education systems are under high pressure due to a lack of teaching staff. With the help of machine learning, TLA can make some teacher tasks like assessment and feedback less time-consuming and more efficient, giving teachers more time to create a rich and interesting learning environment for their students. The teacher can then further use HILA to provide immediate and personalised feedback on the student's outcomes. As a result, students feel motivated and excited to learn, and teachers are able to provide effective feedback and create an enjoyable learning environment.

2 Create data-enriched learning activities that can be applied across disciplines

Instead of focusing on one particular activity to provide LA interventions, the HILA approach wants to address the most common learning activities that happen on a daily basis in schools and higher education to make a significant impact on today's learning and teaching practices. The HILA programme shares these ambitions with other research groups that review the accomplishments of the last decade of LA research to make a practical impact on teaching and learning today. Saqr et al. (2022), for instance, explore whether and to what extent commonly used indicators of success are transferable to a homogeneous group of courses. The results showed that all the indicators had a statistically significant combined correlation coefficient with grades and could play a role in developing predictive models. It was found that indicators based on forum posts and course browsing were good predictors of student success and had high prediction intervals. Furthermore, the study shows how reliable and repeatable the indicators of overall activity (like the number of events, sessions, and time spent online) and the indicators based on forum contributions are. But these indicators are very general, and the authors cannot sufficiently address specific learning goals or skills with them.

Within the HILA research programme, we therefore aim to develop interventions that address common activities like reading, writing, calculating, programming, modelling, and group interaction that appear in almost any study program. We do this in order to provide the most comprehensive coverage of learning activities and to be able to make a HILA offer in almost every discipline. We therefore develop so-called data-enriched learning activities that enable us to provide highly informative feedback to students.

3 Build up a body of knowledge on the effects of data-enriched learning activities

Although there has been research on LA for at least 13 years, the empirical base in terms of randomised controlled field trials is still very thin. Already in 2016, there was a call for a body of knowledge about the effects of LA approaches that was based on more evidence-based research approaches. The EU project "Learning Analytics Community Exchange" (LACE) and the ambition to establish a LA evidence hub arose from this need. The LACE evidence hub aimed to distil an overview of effective and less-effective LA approaches from the yearly LAK conference and related journals. The evidence hub gave an overview of effective and ineffective LA studies based on four criteria: whether they improve and support learning outcomes, improve learning support and instruction, are widely used, and are used in an ethical way (Ferguson & Clow, 2017). Although, the LACE project made promising progress with new workshop types like the "LAKfailathon" to discuss mistakes in LA and integrate central information for the LACE evidence hub into the LAK conference review system. The LACE evidence hub ultimately failed due to the resources required to handle the massive number of new publications.

Nevertheless, there is still a need for a comprehensive and validated overview of different LA approaches towards specific LA projects and objectives. In practice, we still need to work on the situation that when a new LA project is started, the project team seldom looks back to previous LA studies and successful LA indicators and metrics applied in previous research. Furthermore, instead of focusing on relevant indicators for the learning process, most teams use the log data with that particular project's technological infrastructure. Thus, many LA projects are still more technology-driven than education-driven. In order to facilitate rich learning experiences, it is thus necessary to critically consider already applied LA approaches and study them in different settings and with different stakeholder groups to receive an evidence-based inventory of LA approaches' effects on desired learning outcomes.

In order to fulfil the three focus areas, we have set for HILA, we have to provide meaningful data to provide highly informative feedback. For this mission, we have to combine insights from the following areas to progress with the HILA research programme: a. Psychometric: Consider the learning goals, outcome, and context of learning like in the field of psychometrics, b. Feedback: Build upon findings from feedback research and also investigate the uptake of HILA feedback on the student side, c. Learning Design: By incorporating LA already at the design stage of learning to contextualise LA and make its feedback highly informative, d. Technology: Develop data-enriched learning activities that can be broadly applied and sufficiently controlled for research purposes, e. Research: Gain more insights from ecologically valid experimental settings and contribute the findings to a body of knowledge on what is effective under which conditions. In the following sections, we will explain each of the five areas in more detail.

3.1 Psychometrics and Trusted Learning Analytics

A comprehensive introduction to the differences and commonalities of Psychometrics and TLA has been provided by Drachsler & Goldhammer (2020). The following section summarises the outcomes of this publication and outlines its relevance for the HILA research programme.

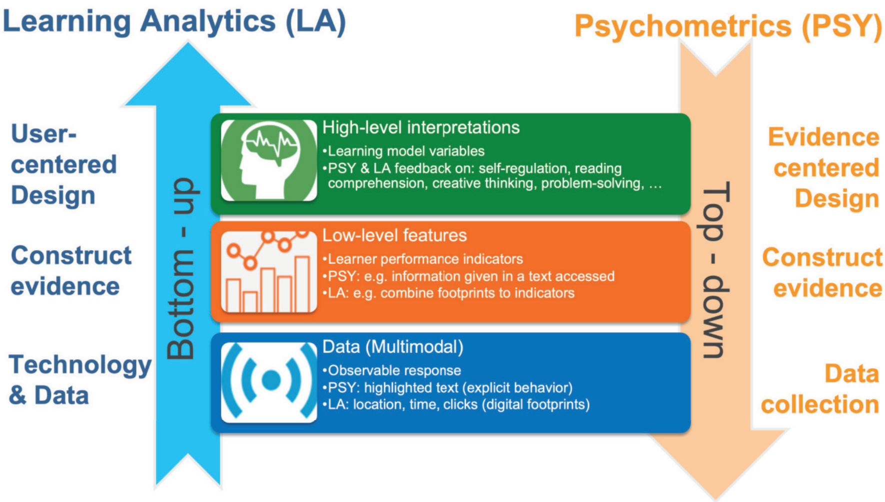
Modern approaches to assessing learning outcomes (e.g., reading and writing proficiency, but also computational thinking, for example) view "*assessment as a process of reasoning from the necessarily limited evidence of what students do in a testing situation to claims about what they can do in the real world*" (Zieky, 2014, p. 79). Thus, psychometrics is mostly about creating and improving psychological measurements, yet it is also broadly applied to measure different skills and abilities of learners using technology. Psychometricians and subject matter experts generally develop highly standardised conditions and activities to elicit the desired behaviour that serves as proof of the targeted construct, like motivation, computational thinking, or academic achievement (Mislevy et al., 2003). Using measurement models, the data obtained from a large number of test items is synthesised to infer differences between people or groups in the intended construct, such as knowledge, talents, or other learner characteristics. The psychometric community has developed a range of valid, reliable, and repeatable methods for measuring these discrepancies.

In 2011, the purpose of LA was to investigate learning by examining the digital footprints of learners; it was not initially intended as an additional method of assessment. Due to the fact that LA, unlike assessments, occurs during authentic learner activities, the information obtained is also ecologically valid in that it reflects actual and continuous learner behaviour. Consequently, LA is more analogous to observational data collection than to invasive acquisition by direct procedures such as questionnaires and examinations (Greller & Drachsler, 2012). However, as LA, similar to psychometric assessment, is about inferring a student's learning state or knowledge level to improve learning, there are natural synergies between these approaches. The core distinction is that LA uses process and product data as well digital learner profiles in authentic settings instead of more controlled assessment situations. Another important distinction is that the feedback cycle of LA is more continuous than that of psychometrics because it involves analysing actual learner behaviour and, ideally, delivering the results to a student or teacher in real-time. Therefore, it has a particularly tight relationship with formative assessment.

Despite the fact that psychometrics and LA share similar objectives while employing different methodologies and ideas, it is astonishing that the communities remain so divided to this day (Drachsler & Goldhamer, 2020). The two disciplines and scientific communities have existed roughly concurrently for the past decade. The reasons for this have more to do with domain- and community-specific notions, terminologies, traditions, and incentive systems than with the spectrum of investigated phenomena. However, it is also evident that integrating the principles and practises of one of the communities with those of the other has considerable promise. This is especially true for the topic of '*formative assessment and feedback*,' which is intensively researched by both communities. Both communities intend to draw inferences about how to support learning based on process and product data and do formative assessment and feedback. However, a clear difference is that process and product data is gathered in different contexts. While educational assessment focuses on processing data from digital educational assessment instruments, LA is primarily about the analysis of interaction data collected continuously in digital learning environments.

One of the key contrasts between the two areas is the starting point from which learning behaviour and consequences are investigated. While the psychometric area normally follows a top-down strategy, beginning with theory and proceeding to data collection, the LA field takes a bottom-up approach, beginning with data analysis and drawing conclusions about potential higher-level skills. Figure 5 below from (Drachsler & Goldhammer, 2020) compares and contrasts the two approaches of the research communities.

Figure 5. Contrasting Learning Analytics with Psychometrics adapted from Mislevy (2019, p.35), Drachsler and Goldhammer (2020).



As illustrated in figure 5, the notion of Evidence-Centred Design (ECD) has a significant impact on the area of psychometrics. Mislevy et al. (2003) introduced ECD as both a design and evaluation tool for (educational) examinations. It begins by determining which constructs should be assessed in terms of learner knowledge, abilities, and other traits (Figure 5 – Green: High-level interpretations layer). These factors cannot be directly seen, hence it is necessary to identify behavioural evidence and performance indicators that assess these so-called latent variables (constructs) (Figure 5 – Orange: Low-level feature layer). Lastly, the behavioural data that serve as the foundation for such indicators must be collected (Figure 5 – Blue: Data layer). To collect this information, objects that elicit the required response behaviour are created. The item structure might range from straightforward multiple-choice questions to intricate simulations. PISA (OECD, 2016) and PIAAC are two of the most well-known worldwide large-scale evaluations of this type (OECD, 2013). One of the benefits of ECD is that it uses a theory-driven method that uses a targeted construct to figure out the activities and data needed to come to a good conclusion about an individual or group. But these benefits don't come for free, as the ECD method takes a lot of time and requires a lot of work when putting the assessment into action. In addition, exam elements might become obsolete or well known by students and therefore have to be replaced with new assessments. In addition, the interpretations of psychometric test results must be validated, for instance, to support the reasoning from the empirical picture of a person in a given circumstance at a certain moment. Consequently, findings from standardised examinations may be more susceptible to variation from irrelevant causes than observations from spontaneous learning contexts. Therefore, "stealth assessment" (Shute, 2011) is an urgent research area in the field of psychometrics, to which LA might potentially contribute.

LA is less driven by a particular research methodology than psychometric research with the ECD. However, the field is dominated by computer scientists who have been evaluating student behaviour using data from digital learning environments. Consequently, the LA field is less theory- and evidence-driven at its core than the psychometrics field. LA collects data from digital learning settings and employs machine learning techniques on this data in order to find patterns that were previously concealed (Figure 5 – Blue: data layer). Using visualisation tools in so-called LA dashboards, these patterns can be reported (Jivet et al., 2017). (Figure 5 – Orange: Low-level feature layer). Lastly, these patterns have been researched and analysed by stakeholders, who have utilised this information to provide input on the learning procedure (Figure 5 – Green: Higher level interpretations layer).

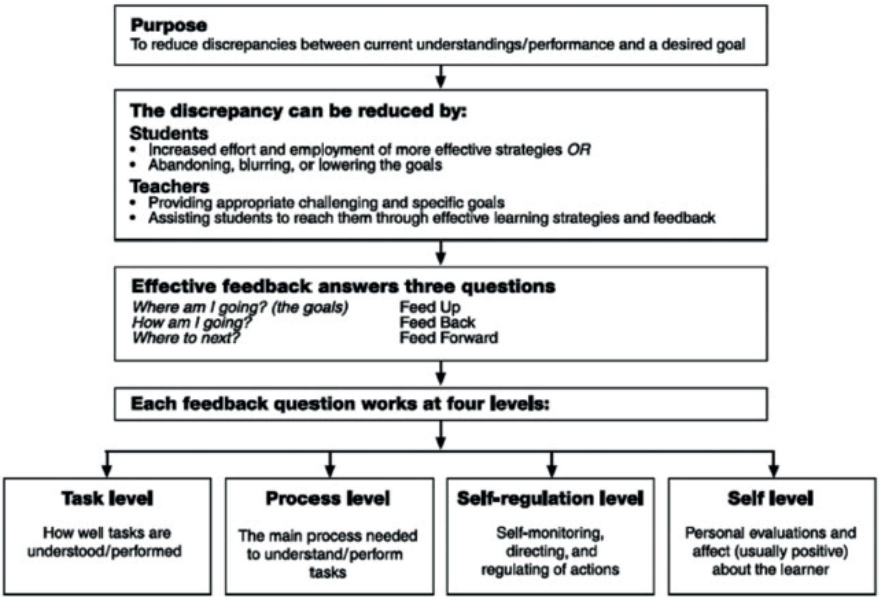
A benefit of LA is that, if the learning takes place in a digital environment, a more comprehensive assessment of the learner's condition and performance may be obtained using relevant data that contributes to measuring goal achievement. In the early days of LA, the field had some quick wins by unveiling so far unknown patterns and providing new insights into learning and learning behaviour that traditional learning science and psychometrics could not show so far (Arnold & Pistilli, 2012). It enabled new kinds of monitoring tools for learning, such as student drop-out warning systems, without any costly human intervention, just by identifying patterns from existing data and indicating if these patterns appear for certain learners or not. Despite the apparent benefits of LA, the field is currently stuck at a particular degree of intervention and critically reflects on the causal effect of its previous endeavours (Weidlich et al., 2023). To find out more about evaluating and researching how people actually learn the user-centred approach needs a deeper understanding of the learning context, such as the course design or curriculum, as well as the learners and their degree of prior knowledge and abilities to be acquired. Psychometrics may give important insights to LA and vice versa to open the next level of assessment and feedback, which can result in highly informative feedback as envisioned by the HILA research programme.

Thus, the combination of psychometrics with LA creates new potentials. The multidisciplinary investigation of the learning process using psychometric methodologies grounded in theory and LA technology driven by data are promising sources that provide highly informative feedback to students.

3.2 Feedback and Trusted Learning Analytics

Feedback has been called "*one of the most powerful influences on learning and achievement*" (Hattie & Timperley, 2007, p. 81). This makes feedback a key element in designing effective learning experiences with learning analytics. Hattie & Timperley (2007) reported in an extensive meta-analysis that feedback has the potential to influence students' performance if it uses the right learning strategies and visualised it in figure 6.

Figure 6: Feedback model according to Hattie & Timperley (2007).



Within the HILA research programme, we also use Hattie and Timperley's (2007) effective feedback model to mainly provide formative feedback. Formative feedback is the input students receive during a learning activity in order to shape and improve their learning processes and outcomes (Shute, 2008). The core of the feedback model consists of Feed-Up, Feed-Back, and Feed-Forward stages, as shown in figure 6. The Feed-Up stage involves setting clear goals and expectations for the learner. The learner needs to know what is expected of them and what they will be able to achieve if they meet those expectations. Feed-Back stage involves providing specific and timely feedback on the learner's performance. The feedback should be focused on the learner's progress towards the learning goals set in the Feed-Up stage. It should be descriptive, informative, and actionable, highlighting what the learner did well and what they need to improve. The Feed-Forward stage involves helping the learner develop a plan for improvement based on the feedback received. The focus should be on what the learner can do to achieve the learning goals. This can involve setting new goals, identifying areas for improvement, and developing strategies for improvement. Overall, the Feed-Up, Feed-Back, and Feed-Forward stage model shows how important it is for the teacher to give clear instructions and set clear goals for the students. It also shows that LA needs to consider the context in which the learning is happening in order to be highly informative.

Following Hattie and Timperley's (2007) findings, highly informative feedback is precious for poor and medium-performing students. However, it should go beyond simple information like points, grades, passing, and failing. Highly informative feedback enables students to self-reflect, improve, and meet educational goals. According to Hattie (2009), feedback strongly affects learning success, with a mean effect size of d = 0.75. Wisniewski, Zierer, and Hattie (2020) even report a mean effect of d = 0.99 for highly informative feedback.

While the power of feedback is widely accepted in the educational research community, some more recent studies draw a more differentiated picture of the effect of feedback on learners. Lipnevich and Panadero (2021) further decomposed feedback into four dimensions to be effective:

A. Cognitive: do I understand the feedback?
B. Affective: do I know how to deal with it?
C. Emotional: do I like the feedback?
D. Useful: do I find the feedback useful?

The authors could demonstrate a correlation between the cognitive and affective dimensions of understanding feedback (Lipnevich & Lopera-Oquendo, 2022) and developed a revised student feedback model (Lipnevich & Smith, 2022). Furthermore, Harks et al. (2014) and Rakoczy et al. (2013, 2019) found a relationship between performance and interest in feedback. Among the positive effects of feedback, Nachtigall et al. (2020) emphasise that the combination of poor feedback and no study success can initiate unproductive and demotivating learning processes that can leave the student hopeless. Thus, feedback has different effects on different learners, and sometimes even negative effects (Wisniewski et al., 2020). This is partly due to the learners, who influence the processing of feedback through their specific dispositions (Panadero & Lipnevich, 2022).

Therefore, the LA community cannot be satisfied with delivering information from data analytics processes on a dashboard or through personalised messages. This would be a simplification of the complexity of education to match the potentials of new technologies. This reduction of educational reality to fit technological requirements is also known as 'simplification trap' or 'technology determinisms' (Winstone & Carless, 2019; De Bruyckere & Kirschner, 2015, 2019). A prominent example of educational simplification and technology determinisms' is the myth of learning styles that never had any empirical evidence but is often used by computer scientist because learning styles are easy to model learners (Kirschner, 2017).

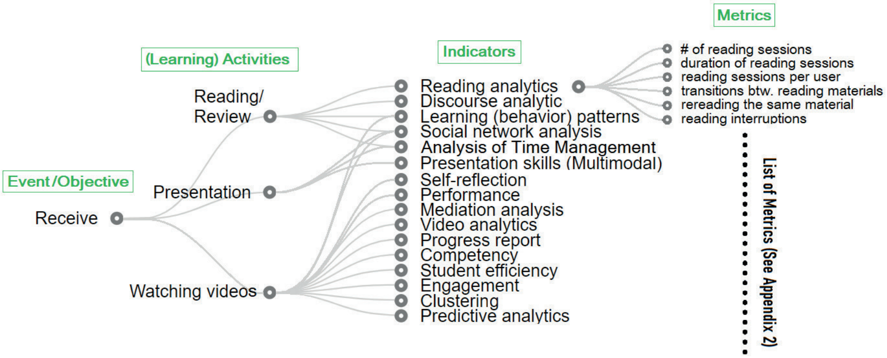
Technology determinism in LA takes away the innovative potential of LA and brings us back to old-fashioned feedback models dressed up in modern technology (Winstone & Carless, 2019). The LA community has acknowledged this tendency to prioritise technological advancements over educational expertise (Greller & Drachsler, 2012; Gasevic et al., 2015; Buckingham Shum et al., 2019; Motz et al., 2022). To increase the effectiveness of LA-based feedback, we must shift our attention from *providing feedback* to *receiving feedback* (Lui & Andrade, 2022). This necessitates taking into account both the characteristics and perspectives of students as well as having rich and relevant data that can be utilised by LA and teachers to provide HILA feedback. Each of these aspects must be considered in order to support optimal learning outcomes.

3.3 Learning Design and Trusted Learning Analytics

In order to make the highly informative feedback more suited to learning activities, it is vital that the learning objectives of a given course are in line with the learning activities, the intended feedback, and the final examination. In other words, constructive alignment (Biggs, & Tang, 2011) is crucial to making highly informative feedback work. To achieve these requirements, we have to think about learning and education as design disciplines (Goodyear & Dimitriadis, 2013). Instructional design (Smith & Ragan, 2004) or learning design (Koper & Olivier, 2004) are two terms that have been used to combine characteristics of learners with learning activities, courses, and curricula. Thus, the design of learning is not a new idea but is currently experiencing a renaissance due to the need for LA to give highly informative feedback. A first step towards HILA is the design of a digital learning environment that provides the process and textual data needed for highly informative feedback. Therefore, the digital learning environment needs to be part of the learning process and supported with customised data-enriched learning activities. It is insufficient to use a system that functions as a content management system, mainly allowing students to access and download lecture learning materials. Instead, close alignment of learning activities, learning environment design, and overall course goals is needed. This means that after defining the course's learning outcomes, not only the assessment needs to be designed, but LA indicators that provide valuable insights into the learner's performance in a learning activity need to be defined, too (Schmitz et al., 2022). Examples of well-designed learning environments that provide learning activities that directly send relevant indicators of learning are manifold. Scheffel et al. (2017) report a LA tool for group collaboration, and Tabuenca et al. (2015) for students' time management, to mention a few. For all these examples, the student's learning process is inherently tied to using a digital learning environment for receiving, producing, and exchanging information. This allows for the observation of relevant learning behaviour that can be used to make inferences about the learner's status and create indicators for providing highly informative feedback.

In a special issue of the Journal of Learning Analytics on Learning Design and LA, Macfadyen et al., (2020) collect many examples of LA-supported design, such as a framework and technique (Law & Liang, 2020) and a model (Mangaroska et al., 2020). Additionally, LA scholars have proposed tools, such as Inspiration Cards (Vezzoli et al., 2020) and the LA-Deck (Alvarez et al., 2020), to capture what educational stakeholders want to look at in order to improve learning processes through LA. Ahmad et al. (2022) conducted a review of the LA literature from an instructional point of view by grouping LA indicators for commonly applied educational activities as reported in LA publications. The outcomes of this review are freely accessible and continuously updated in the *Open LA Indicator Repository* (OpenLAIR[10]), which enables users to explore LA indicators for certain learning activities (see figure 7).

Figure 7: A tree-view of the Learning Event (Receive) in OpenLAIR, followed by typical Learning Activities, their LA indicators, and metrics.

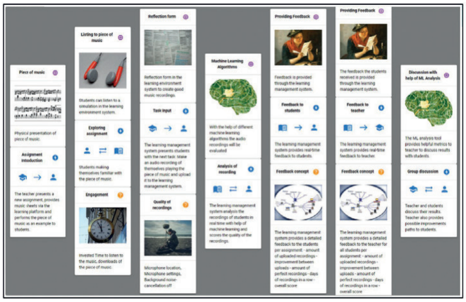


Schmitz et al. (2022) have responded to the need to design learning and its analytics in order to provide highly informative feedback. The authors created the FoLA2 (Fellowship of Learning Activities and Analytics) method to enable the practical application of LA in current learning designs (see figure 8). FoLA2 is based on a game board with consecutive movements by multiple players (roles) to create a learning analytics-supported learning design. It is also made available as a digital working environment[11]. The approach enables participants with varying responsibilities to engage with a deck of cards to develop a LA-supported learning design collectively. FoLA2 can be used to develop, collect, and systematise design aspects, as well as to integrate LA in a systematic manner. It assists teachers in integrating LA into their learning designs, allowing them to employ LA inside their own setting and context while integrating their educational vision and philosophy. It can also act as a knowledge base within a teaching team or institution to facilitate the transmission, analysis, and improvement of their instructional models. FoLA2 has been widely

10 https://edutec.science/products/
11 https://fola.s.studiumdigitale.uni-frankfurt.de/

applied by several institutions in the Netherlands, Germany, and Italy in various disciplines (STEM, psychology, and education) and at different school levels (secondary school, higher, and further education) (Kubsch et al., 2022).

Figure 8: Examples of FoLA² Learning Designs, on the left a psychology course designed on the FoLA² board game, and on the right a music course designed in the digital version of FoLA².



3.4 Technology for HILA

With the combination of the FoLA² method and OpenLAIR tool, we are able to design Data-enriched Learning Activities (DeLAs) that collect relevant data for the learning process as well as the learning outcome of a learning activity within the HILA research programme. In contrast to the rather nonspecific log-files from LMS, DeLAs collect data that is aligned and meaningful with respect to the actual learning task. As such, data sourced from DeLAs are the most accurate representation of learning in digital learning environments one can get with LA.

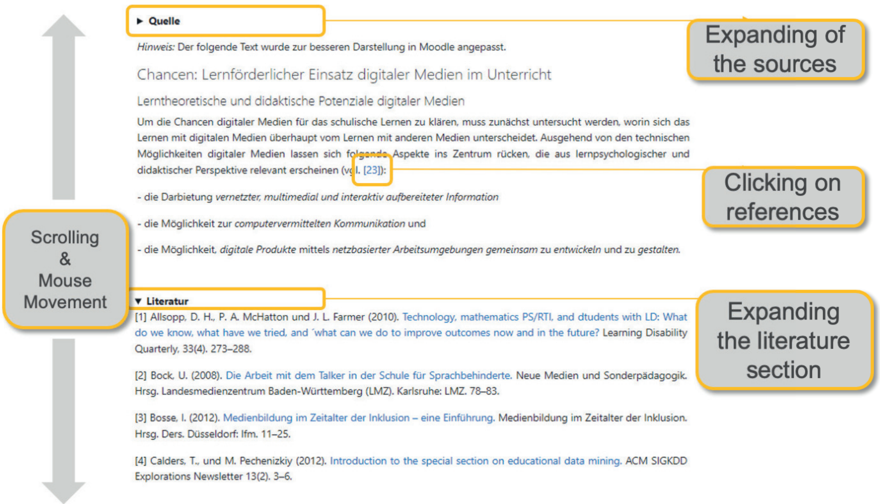
On the one hand, the DeLAs are highly flexible, as they can be applied to a variety of learning scenarios in different scientific disciplines. On the other hand, they also offer fairly stable and reliable conditions that make them well suited for experimental settings to test and evaluate the effects of LA in the field. They ultimately contribute to building a body of knowledge and provide insights into effective and less effective indicators for different learners' dispositions. Within the HILA research programme, we have been focusing on the most common learning activities that students can have within a broad range of studies, such as 1. reading, 2. writing, 3. collaborating, and 4. modelling. In the following section, we will introduce four types of DeLAs that have been developed for the LMS Moodle.

3.4.1 DeLA for reading

Reading is a central part of our culture, and students still often work with written text. Many study programmes include reading difficult literature as a key component (Wolf & Barzillai, 2009). The Reading-DeLA (Biedermann et al., 2023a) is based on Hahnel et al. (2019) insights into what constitutes academic reading skills. It records events like mouse movement, marking text, expansion of source and reference, and data about how long certain section of text shown (see figure 9).

Most LA applications apply the time-on-task indicator to estimate if a text was sufficiently read or not (Cocea & Weibelzahl, 2011; Mills & D'Mello, 2015; Kovanović et al., 2015). The use of time-on-task is appropriate because a text's characteristics, such as its length and difficulty, predict how long it will take to read. As a result, reading difficult texts requires spending more time (Goedecke et al., 2015). The benefit of using time-on-task as a text completion indicator is that it can be constructed from process data alone, without additional efforts. Despite its usefulness, time-on-task as a measure of student engagement is vulnerable to inaccuracies caused by factors such as prolonged inactivity on reading materials, which can overestimate the actual time spent on a task. For instance, if a student opens a text and then leaves it untouched for an hour, the time-on-task indicator would show 60 minutes spent on the reading assignment, although the actual time spent reading was significantly less (Biedermann et al., 2023a).

Figure 9: The Reading-DeLA Widget in the Moodle LMS (Biedermann et al., 2023a).



Within the Reading-DeLA we use scrolling data as a feature to analyse reading behaviour in addition to the conventional time-on-task indicator. When reading a document that is bigger than the space on the screen, you have to scroll to move the area you can see. The "viewport" refers to this region that is viewable (see figure 10). Like a sort of eye-tracking (Catrysse et al., 2018) viewport data can be used to estimate what a user is looking at. This can be accomplished by analysing the viewport scrolling data to determine how long a student spent on different sections of the document, as opposed to time-on-task systems, which just measure the time spent on the entire page. Neither the time-on-task nor the viewport approaches can indicate whether or not students read carefully and comprehended what they were reading. Yet, we obtain a more accurate estimate if particular portions of text are displayed on the screen for at least enough time to provide a comprehensive reading. By utilising this capability, we may evaluate whether or not a reader has likely completed a text. In addition, there is a great deal of untapped potential in studying the function of text interaction elements such as marking or commenting on a text. Based on the outcomes of Reading-DeLA, the learners receive feedback on their academic reading competence and potential actions for improvement (see figure 11: Reading Dashboard).

Figure 10: Visualisation of the view port technique to analyse the reading behaviour (Biedermann et al., 2023a).

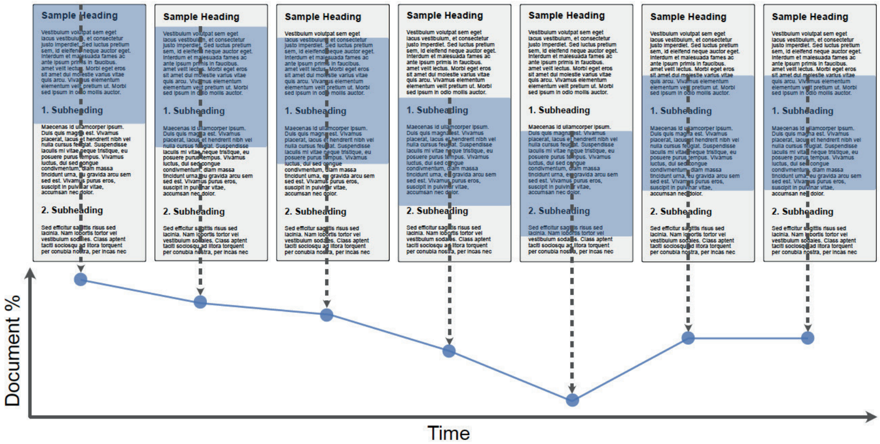
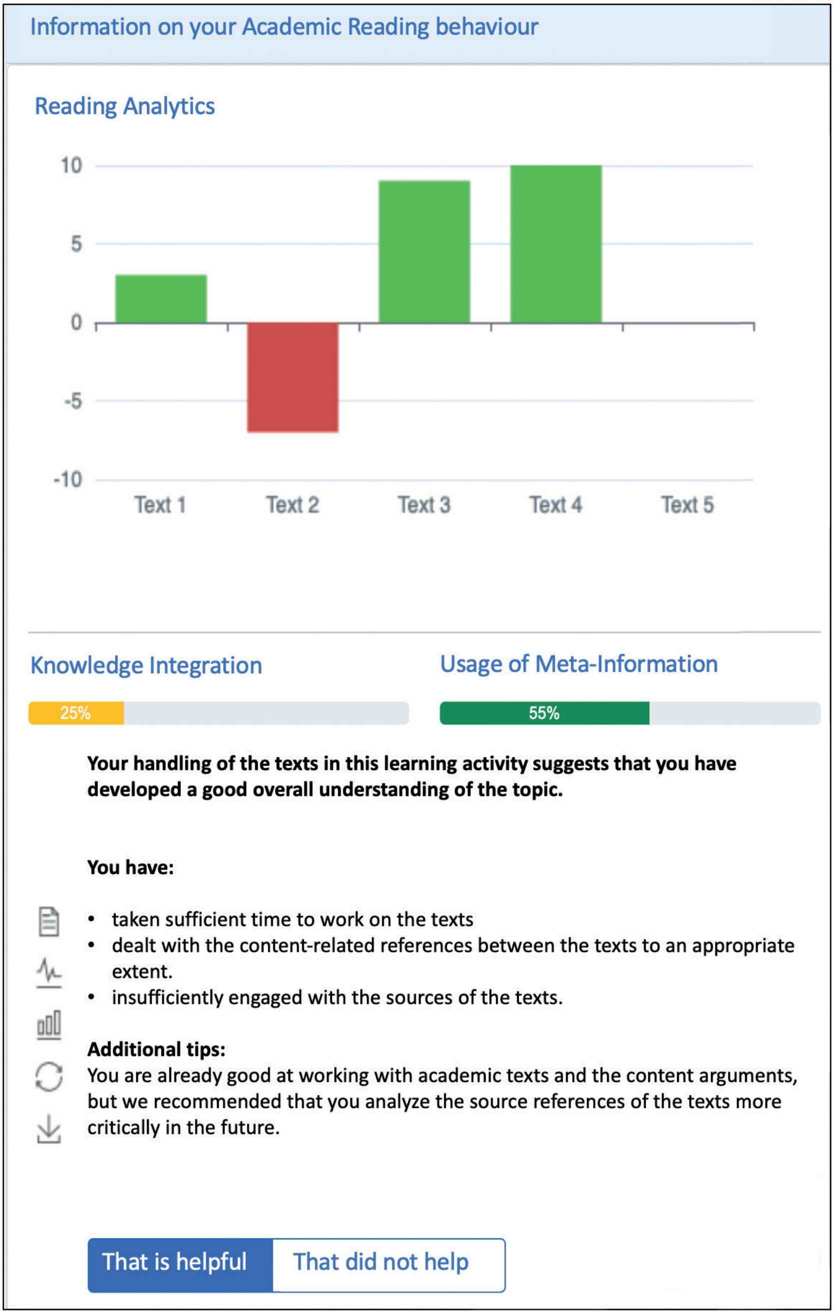


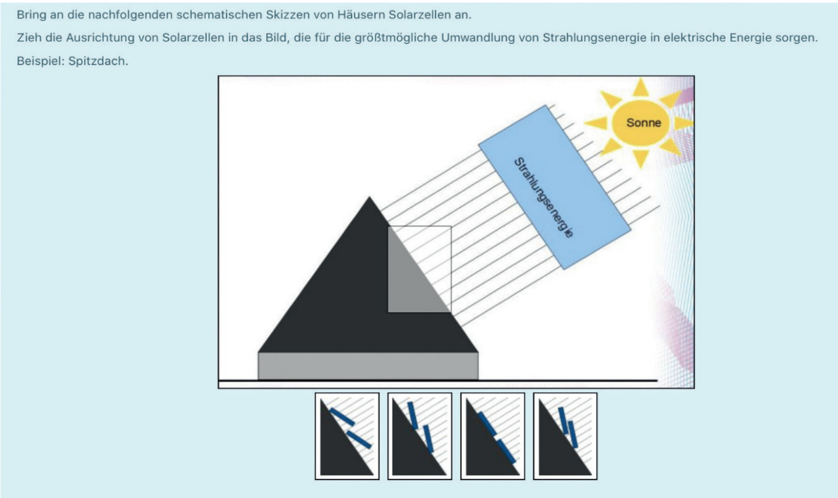
Figure 11: Example of Reading Dashboard from the DiFA project funded by the Leibniz Cooperative Excellence. The dashboard visualises the results of LA indicators for academic reading behaviour (Biederman et al., 2023a) and descriptive text with hints for working with text in the future.



3.4.2 DeLA for text

Even though multiple-choice assessments are the most common type of assessment item because they can be graded automatically, there is a notion to also use open-text fields to let students write their individual answers and demonstrate their knowledge in a written text (see figure 12). But open-text fields require much more efforts from teachers because, unlike multiple-choice items, the answers have to be read carefully and graded manually, what results in a higher time investment.

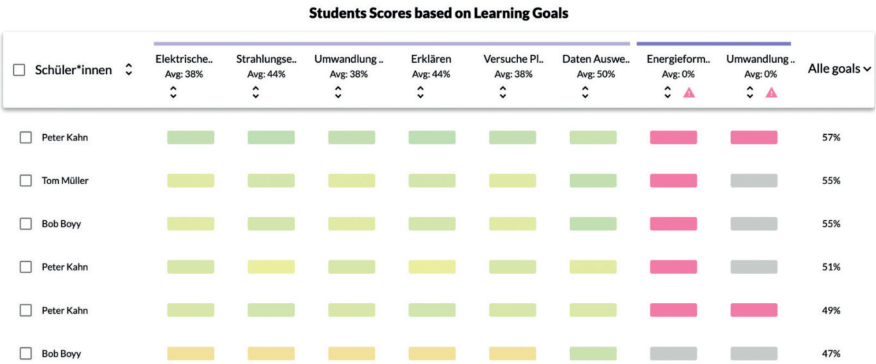
Figure 12. Example of an open text assignment in the Moodle LMS. The written explanation of the student is analysed by a language model to assess the correctness of the response.



Many studies have been undertaken on how to automatically evaluate open student replies, although it is one of the oldest and most generally acknowledged uses of natural language processing in education, also known as "automatic short answer grading", the topic is still not sufficiently solved (Burrows et al., 2015). While latent-semantic analysis has already shown the potential of natural language processing for prior-knowledge assessment (Kalz et al., 2014) and automated answer grading (Wild et al., 2005) years ago, the widespread success of neural networks, especially with transformer language models such as BERT (Devlin et al., 2019), resulted in a new level of accuracy in the analysis of the open-text responses. For the HILA approach, we take advantage of the transformer models advanced language technologies to provide an advanced analytical evaluation of students' responses.

Our Text-DeLA evaluates students' understanding of various kinds of topics by identifying arguments that are an indicator for a deeper understanding of a certain topic. The Text-DeLA was used successfully on short text responses to inform high school STEM teachers how knowledgeable their students were (Karademir et al., submitted; Gombert et al., 2022). Furthermore, the Text-DeLA also has been applied to higher education students in the field of educational science to provide them with highly informative feedback on written abstracts (Gombert et al., submitted). Similar to all DeLAs, the Text-DeLA can provide highly informative feedback on a dashboard (see figure 13).

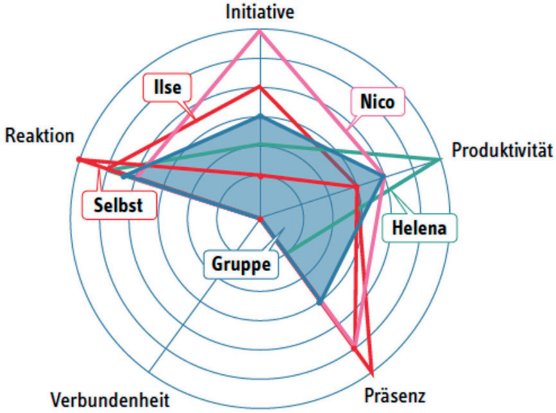
Figure 13: Example of a teacher dashboard that highlights the knowledge level on certain learning goals of each student on the basis of aggregated open text responses of the students (Karademir et al., submitted).



3.4.3 DeLA for collaboration

Collaborative learning has been a crucial educational method since the earliest days of education, and so it is for digital education today. Using synchronous and asynchronous communication tools helps virtual student teams work together to learn. Not only since the COVID-19 pandemic (Drachsler et al., 2021), it is known that facilitating all communication, coordination, and cooperation using online technologies can result in less cohesion and support for team members (Kreijns et al., 2003). This can diminish perceptions of social presence and provide barriers to successful collaborative learning. Making students more aware of how their groups performs is one way to improve collaborative learning experiences (see figure 14). LA can be used to automatically collect and analyse interaction data from the LMS. In the early days, this was done mainly with metadata collected from forum posts (Scheffel et al., 2017) or in combination with social network analysis (Bakharia & Dawson, 2011). Also here the arisen of the transformer models created new chances for analysing group collaboration and providing highly informative feedback to learners. Current approaches promote the use of emergent roles in the collaboration process (Dowell et al., 2019; Saqr & López-Pernas, 2022) as opposed to earlier Computer Supported Collaborative Learning (CSCL) scripts, which assigned roles that were fixed a priori (Weinberger et al., 2010). The combination of group communication analytics in combination with transformer models and time series analysis (Dowell et al., 2019), allowing for automatic measurement of variables such as involvement, responsiveness, cohesion, and novelty, and grouping them into emerging roles in the collaboration.

Figure 14: Example student dashboard for group feedback (Hansen et al., 2020; Scheffel et al., 2017) based on Collaboration-DeLA (Menzel et al., submitted).

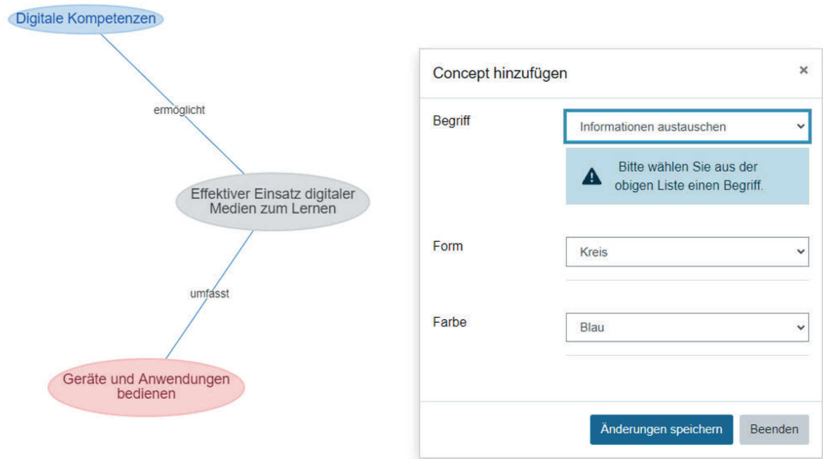


Within the Collaboration-DeLA (Menzel et al., 2022; Menzel et al., submitted), we have developed machine learning models to compute these emerging roles and can provide them with highly informative feedback based on their role membership in the collaboration process. The analytics approach is also used to research the effect of gender in CSCL scenarios of mixed student groups in STEM education (Kube et al., 2022; Kube et al., submitted).

3.4.4 DeLA for concept modelling

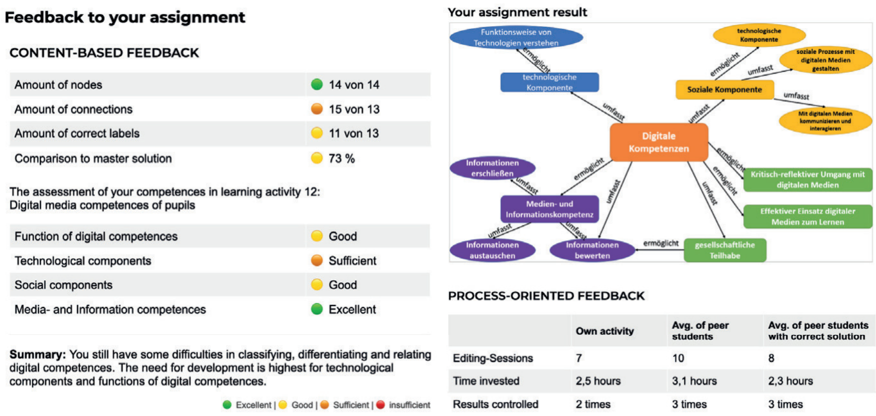
Concept maps are graphical representations of information and have been extensively used in a multitude of scientific disciplines for many years (Malone & Dekkers, 1984; Austin & Shore, 1995; Schaal, 2008; Novak & Cañas, 2006). When students create concept maps, significant learning is enabled because they can serve as an effective method for enhancing text learning and comprehension (Chang et al., 2002) by detecting significant aspects and placing them in context. They facilitate the creation of a mental model of a topic by illustrating its relationships with and distinctions from other topics. Concept maps provide more information than mind maps by incorporating directional arrows and labels on the edges and nodes. They assist students in synthesising a multitude of topics and making connections between separate concepts.

Figure 15: The Concept-DeLA Widget in the Moodle LMS.



So far, evaluation methods for digital concept mapping systems have been limited (Chang et al., 2005). The most important part of the evaluation has so far been how well the nodes are used. The relationships between them have been neglected. The Concept-DeLA (see figure 15) enables students to build, alter, and delete elements of a concept map of a particular subject entirely online. It is highly adaptable to any subject and can be used with a predetermined list of elements or, in "free mode", with an empty concept map without any required elements given in advance (Giorgashvili et al., submitted a).

Figure 16: Example of Concept-DeLA Feedback Dashboard from Giorgashvili et al. (submitted b).



With the advent of LA, the role of concept maps has become more relevant again (Cañas et al 2015; Carrillo et al., 2019). Within the Concept-DeLA, we revisited the assessment of digital concept maps by focusing on the one hand, on the process level of creating concept maps and on the other hand, the product level by comparing the outcomes of a concept map to a set of master solutions. We can provide highly informative feedback on the actual working phase of the students with the concept map as well as the overall correctness of the concept map. Similar to all DeLAs, the Concept-DeLA provides highly informative feedback on a dashboard, as shown in figure 16.

3.5.5 DeLA delivery system for Highly Informative Feedback

The OnTask System served as our inspiration for delivering HILA feedback to students (Pardo et al., 2018; Lim et al., 2021). OnTask is an open-source platform that allows educators to combine various data sources, such as the Student Information System (SIS), LMS, and Grade Book, for analytics purposes and ultimately send personalised messages to individual students.

Using OnTask, educators can create workflows that are customised to their teaching environment and students' needs. These workflows may include a variety of learning activities in the LMS, such as quizzes, assignments, surveys, and written material reports. For each workflow, a variety of feedback templates that are sent to students after they reach a particular milestone in a course or curriculum can be created. Each workflow activity can be customised with specific instructions, deadlines, and other data. The student's performance in the preceding learning activity can also be used to trigger the subsequent workflow activity. In this manner, students receive sequential automated feedback while interacting with the digital learning environment. Each student's OnTask feedback is personalised, taking into account their unique assignment results and learning outcomes. Based on a student's performance and conduct, the OnTask Personalisation System can employ data analysis and predictive models to provide individualised interventions and activities. Figure 17 depicts the OnTask architecture, where data is imported from LMS and SIS to the OnTask platform. These data sources are processed by Data Operators who request machine Data Analysis and Machine Learning models in order to acquire analytics results for the personalised feedback messages. The Data Operators provide a table view for each student's LA results after receiving this request. The instructors are then able to design customised feedback templates and workflows for particular learning activities based on the IF-THEN rules derived from the LA indicators in the student table. Finally, OnTask sends students personalised emails containing feedback.

Figure 17: The OnTask architecture from Pardo et al. (2018).

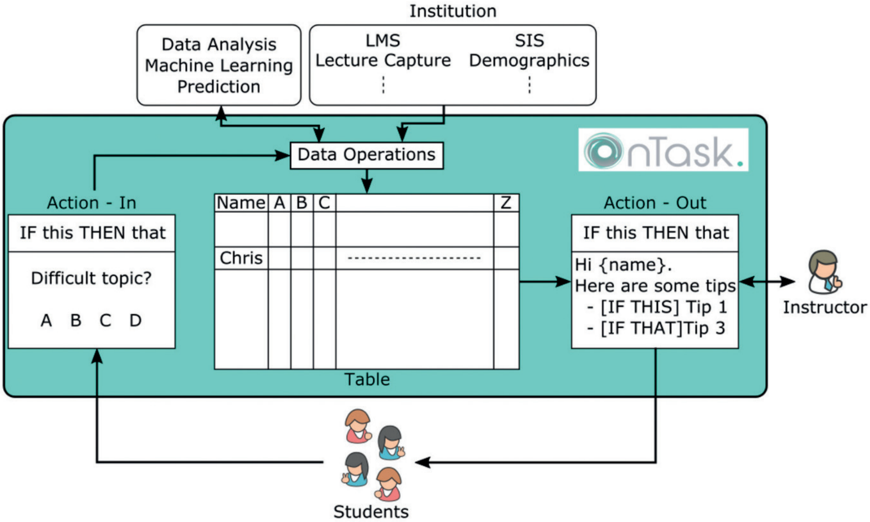
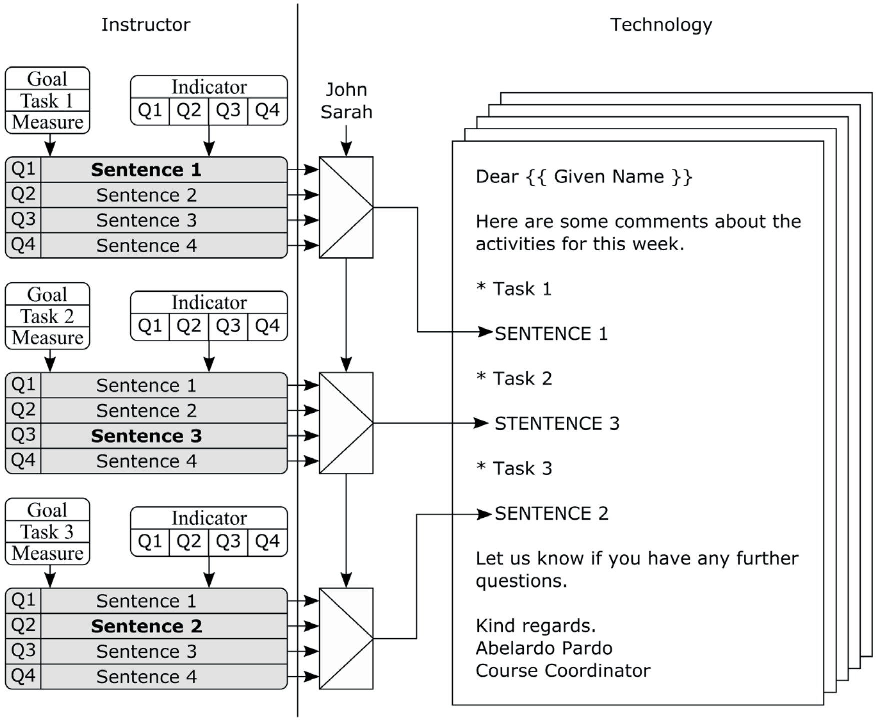


Figure 18 illustrates in greater detail the workflow for creating a personalised feedback message using the OnTask template. The feedback is generated in a bus-like system, represented in Figure 18 by an envelope. The envelope passes through numerous learning activities with specific objectives, tasks, and measurements. According to LA indicators, the students' performance is based on this information. Each indicator corresponds to a text excerpt. The relevant feedback text is activated and inserted into the feedback message depending on which indicator category a student belongs to. After all learning activities in a workflow have been requested, the student receives a personalised feedback message.

Figure 18: Example of a workflow of the OnTask Feedback System (Pardo et al., 2018).



One effective way to improve students' learning outcomes is by providing them with timely and actionable feedback (Hattie & Timperley, 2007). Within the HILA research programme, it is not sufficient for us to send personalised feedback messages to the students, we need to consider two important aspects of the HILA research. First, the ability to *research how students interact with formative feedback* and, second, the *benefits of repeated formative feedback to students* in which performance in relation to individual competence and learning objectives is displayed and can be reflected.

Concerning the first point, researchers and educators can gain valuable insights into students' learning processes by investigating how they interact with feedback. The instructors can identify patterns and trends, understand what works and what does not, and make adjustments accordingly. This information can be used to improve the quality and effectiveness of feedback and ultimately, enhance students' learning outcomes. To ensure that students reflect on their learning, instructors can set reflection prompts for the learning objectives of the course based on feedback. Reflection prompts encourage students to consider their learning and its evolution over time.

Regarding the second point, it is vital for feedback to be repeated. Feedback that is provided only once or infrequently is unlikely to have a lasting impact on students' learning. Repeated feedback, on the other hand, enables students to track their progress over time and make continuous improvements. In addition, HILA feedback that displays performance in relation to competence learning objectives can be particularly valuable. This type of feedback enables students to see how their performance aligns with specific learning objectives, and identify areas where they need to improve. By reflecting on this feedback, students can set realistic goals, track their progress, and stay motivated to achieve their learning objectives.

Therefore, we developed a system similar to OnTask that transmits DeLA feedback to a LA dashboard (Giorgashvili et al., 2023). The LA dashboard enables students to receive regular feedback, compare their performance and reflections from previous DeLAs with the most recent one, and gain a better understanding of their learning progress. The feedback they receive is highly informative and relevant to their current learning goals and context, which helps them stay motivated and engaged in their learning process.

4 RESEARCHING THE EFFECTS OF HILA

As a field of study committed to understanding and increasing learning, LA must be able to provide empirical evidence for its effects on learning. This is particularly true today, where a growing number of LA methods are moving from the laboratory to educational practices and being implemented in educational institutions. Unfortunately, tightly controlled, randomised field trials are not often conducted in the LA field (Viberg et al., 2018; Weidlich et al., 2022). But that these experiments are feasible and also deliver promising results for the application of HILA has been demonstrated, for instance, by Meurers et al. (2019). The authors conducted a randomised field study with an Intelligent Tutoring System integrated into regular foreign language classes in Germany and demonstrated significant learning improvement from automated learner-guided feedback in comparison to standard feedback (true/false). Comparable studies in this area are still rare; instead observational data, from which it is difficult to discern causal links, is routinely employed.

There have recently been a number of critical articles on the existing research in the LA field. Weidlich et al. (2022) and Hicks et al. (2022) comment on the generation of findings from LA experiments while also demonstrating an approach to generating causal knowledge outside of highly controlled experiments. With the HILA research initiative, we aim to address this lack of research in the field by conducting more reliable and valid LA studies to determine the causal impact of LA on learning. We further aim to build a body of knowledge about the effects of LA on multiple target groups, since many LA methods can be used with different groups, from school-aged children to college students. When selecting to use a LA instrument in practice, the instrument should be selected with a specific objective and a particular group of respondents in mind. A first attempt to achieve this ambitious aim has been made by

Ahmad et al. (2022) and the Open Learning Analytics Indicator Repository (OpenLAIR[12]) as well as the development of the DeLAs (Biedermann et al., 2023a; Gombert et al., 2022; Karademir et al., submitted; Menzel et al., submitted; Giorgashvili et al., submitted a) that are highly adaptable to any domain but can be strongly controlled to enable experimental conditions.

4.1 Main research questions

As stated in Section 3, the goal of the HILA research programme is to overcome the shortcomings found in the existing LA research field and to provide students and teachers with highly informative feedback in authentic learning situations. We therefore investigate the following four primary research questions:

1. How can relevant data for the learner goals and outcomes of a course be extracted from digital learning environments?
2. How valid is the interpretation of indicators derived from digital traces?
3. What is the effect of different feedback types on assignment results, exam performance, and affective student variables?
4. How does feedback literacy influence students' interpretation and reaction to the received feedback?

The purpose of these questions is to provide a deeper understanding of how to build DeLAs that can provide relevant data to deliver highly informative feedback in order to improve student learning process and outcomes. We will also investigate the quality of several HILA feedback categories that carry personalised feedback for a learner. The primary types of HILA feedback are:

A. Norm-referenced feedback, which compares a student's performance to that of a certain group or cohort of students,
B. Criterion-referenced feedback, which primarily compares the student's performance to educational standards,
C. Combination of both norm-referenced and criterion-referenced feedback

As well as feedback on different levels of Hattie & Timperley (2007) feedback model:

D. Task-level, how well tasks are understood and performed?
E. Process-level, the main process needed to understand and perform tasks
F. Self-Regulation level, self-monitoring, directing, and regulations actions,
G. Self-level, personal evaluations and affect about the learner

and the three stages, according to Hattie & Timperley (2007) feedback model:

H. Feed-Up: What are you supposed to do?
I. Feed-Back: How did you do that?
J. Feed-Forward: What should you do next on the basis of the Feed-Back?

12 https://edutec.science/open-learning-analytics-indicator-repository-openlair/

The control group will receive feedback as usual, with non-personalised advice and broad recommendations for enhancing their learning results. In the following subsection, we describe the four research questions in more detail.

Research question 1: How can relevant data for the learner goals and outcomes of a course be extracted from digital learning environments?
Research question one focuses on capturing relevant and rich data from digital learning environments that are appropriate to deliver HILA feedback during the learning process and on student learning products. A bottom-up and top-down strategy, as stated in Section 3.1, is required to create data-enriched learning activities (see Section 3.4) in order to collect valuable data. This collected data can be merged with various machine learning techniques, and data geology and science methods are required to optimise the acquired data for the affective student characteristics, such as motivation, learning goals, and outcomes of a course or learning activity.

Research question 2: How valid is the interpretation of indicators derived from digital traces?
Regardless of how indicators are generated, their interpretation must be validated to verify that assumptions about, for instance, a learner's reading skills are justified (Biedermann et al., 2023a). The intersection of theory-driven psychometric indicators for motivation, engagement, or reading skills and data-driven LA indicators is uncharted territory. A strong correlation between both types of indicators (theory-based indicators and data-driven indicators) could validate an LA indicator and promote it as a promising alternative measurement that does not require any additional psychometric assessment. A weak correlation, on the other hand, challenges the validity of a LA indicator or the applicability of the psychometric method in that context. While varied studies seem to broadly find weak or non-existent links between psychometric scales and relevant indicators (Choi et al., 2023; Quick et al., 2020), this is still an emerging body of research. However, this may not apply in the same way when DeLA's, which are designed to deliver relevant data, are the main data source. For this reason, a central question of the HILA programme is to further investigate the validity of DeLA indicators to represent psychometric and educational constructs, like reading skill, domain competence, collaboration skills, and others.

Research question 3: What is the effect of different feedback types on assignment results, exam performance, and affective student variables?
The indicators derived from a digital learning environment for which validation was successful will be used for feedback purposes, such as on affective student variables like motivation, engagement, or academic achievement, to motivate unmotivated students or support the self-regulation of students who are unable to properly plan their learning trajectory. Among affective student factors, HILA aims to deliver highly informative feedback on the learning process and the end product created by the student throughout a learning activity. It is investigated if personalised HILA feedback based on validated indicators has a positive influence on the learning progress

of students as measured by assignment results, exam performance, and affective student factors. For this purpose, experimental pre-post designs will be conducted, with several treatment groups receiving variations of personalised highly informative feedback (norm-referenced, criteria-referenced, and norm- and criteria-referenced feedback as well as feedback on task-, process-, self-regulation-, and self-level) and a control group receiving non-personalized feedback as usual with general learning advice.

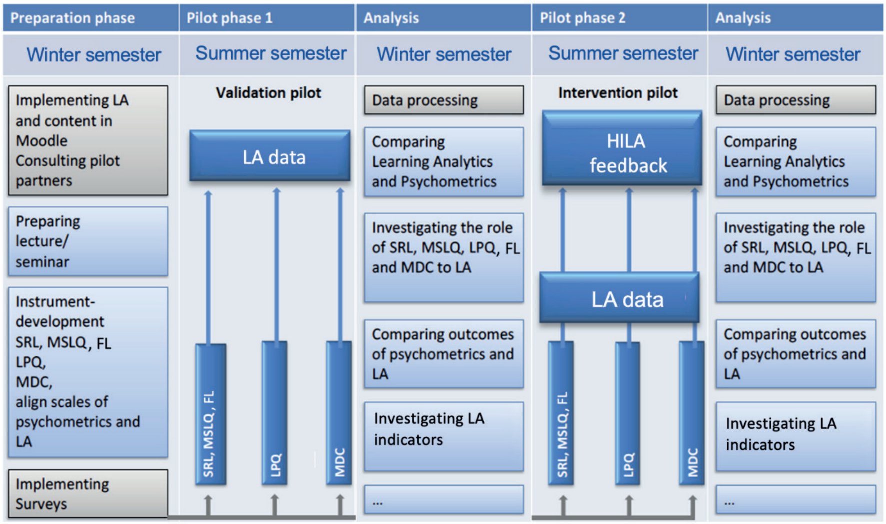
Research question 4: How does feedback literacy influence students' interpretation and reaction to the received feedback?
Feedback literacy is a frequently neglected factor that might decrease the effect of HILA feedback on learning behaviour and performance. According to Carless and Boud (2018) definition, feedback literacy comprises the information, abilities, and attitudes required to grasp and implement feedback to enhance one's learning or work techniques.

Many elements will be investigated in order to acquire a deeper knowledge of how feedback literacy affects students' interpretation and use of feedback. They include the importance that students place on feedback, their interpretation of the feedback they get, the planned adjustments to their behaviour following feedback interpretation, and the influence of specific student characteristics such as feedback literacy.

4.2 Study designs

The study design for HILA experiments involves at least two empirical phases but could also be continued. The first phase (validation pilot), addressing research questions 1 and 2, aims at the extraction and validation of indicators based on process and product data from the learning environment (see figure 19). For this purpose, students' activities in the learning environment are logged when engaging in a learning activity. In addition, they also complete standardised assessment instruments outside the learning environment, tapping the targeted constructs of the pilot partners like learning engagement with the LPQ measure (Biggs, 1987), self-regulated learning with the MSLQ measure (Pintrich, et al., 1993), or feedback literacy (Woitt et al., submitted). The second phase (intervention pilot), which addresses research questions 3 and 4, uses the data from the first pilot study to compute machine learning models for the various DeLAs applied in a course. These machine learning models will then be applied to a new cohort of students completing the same course and provide HILA feedback for each DeLA a student completes. To investigate the effectiveness of the informative feedback, the experimental group receives HILA feedback based on machine learning outcomes, while the control group receives feedback-as-usual with general advice and less personalised information. As feedback-as-usual was the established approach it is a natural reference for comparison, i.e. an authentic control group in a real-world setting (ecological validity). In order to have a fair comparison between the treatment group and the control group, care has to be taken to align all feedback types in terms of length, density, and layout. At the end, the performance of both experimental conditions is compared on the basis of assignment results, exam performance, and affective student variables.

Figure 19: Example of the continuous research design cycle for HILA.



5 CHALLENGES OF HIGHLY INFORMATIVE LEARNING ANALYTICS

While working on the HILA research programme, various technical and educational challenges have been encountered in order to implement the research outcomes into university practices. These challenges can be clustered into technical and educational types. In the following subsection, an overview of the current challenges is given.

5.1 Technical Challenges

Within the research on HILA, a new set of technical challenges has emerged. These challenges relate to: 1) the need for systems and tools that can extract relevant learner attributes; 2) the annotation of textual data for machine learning; and 3) the delivery of LA-driven feedback to students. Overall, these three technical challenges are critical for the development of effective and scalable technology solutions for HILA.

5.1.1 Extraction of learner attributes in authentic contexts

To give students useful feedback on their learning goals and skills, the design and context of the learning environment need to be considered. This is because you need to gather relevant learner attributes from learning data. Learner attributes must be chosen to address specific course learning goals as well as possibly overarching competencies such as critical thinking or collaborative problem solving. For HILA research, we use the "evidence-centred design" method from the field of psychometrics. This method starts with a cognitive model of how learners with a certain level of competence process information. It starts with a description of the information processing level the learner is expected to go through when interacting with a learning activity. Based on this cognitive model of competence, we build learning activities that are rich in data (see Section 3.4 on DeLA) and have relevant LA indicators. The DeLA construction guarantees that meaningful data is collected and combined to maximise the information relevant for the learning goals of the course. This has been successfully done in the area of higher education in the HIKOF-DL[13] and IMPACT[14] project (Gombert et al., submitted) as well as for STEM education in schools in the AFLEK & ALICE project[15] (Kubsch et al., 2022). In the school context, the LA needed to inform about competences like being able to apply concepts of energy transition (see figure 13), as well as overarching competencies such as collaborative problem-solving with peers. In order to provide meaningful feedback on these competences it is required to extract relevant learner attributes for these competencies and thus intensively work in the field of practise.

5.1.2 Annotation of textual data for machine learning

Open-text fields are still an important part of a digital learning environment because they encourage critical and creative thinking more than other types of input fields. However, the automatic analysis of open-text fields is still costly and labour-intensive. The text corpora are rather large, consisting, for instance, of 600 - 800 abstracts submitted by students. The collected textual data needs to be labelled by human raters. The human raters therefore have to agree on a coding schema and also empirically test that they can achieve an acceptable inter-rater reliability (IRR) level. If the IRR is not above a certain threshold of at least 85%, the rating needs to be repeated, and further work on the coding schema is needed. Thus, labelling a text corpus is still a time-consuming and exhausting task, but it is also critical for training machine learning models. Burrows et al. (2015) say that a lot of natural language processing techniques have been used to automatically code answers to open-text assignments, but this problem is still not solved well enough. Therefore, there is still more research needed in order to gain machine learning solutions that reduce the effort of labelling textual data. Creating technologies to automate this process could speed it up, but this is not a trivial task. In addition to this technical challenge, the transfer to practical educational cases remains a significant challenge. A more general approach to semi-automatically labelling the local text corpus of a specific course is needed in order to address the mass of texts that are produced daily in an educational context.

13 https://hikof.uni-frankfurt.de/
14 https://impact.studiumdigitale.uni-frankfurt.de/
15 https://edutec.science/alice-and-aflek-project-meeting-in-kiel/

5.1.3 Delivering automated feedback to students

Developing a fully automated feedback system for DeLAs in digital learning scenarios is not a trivial task and requires careful planning, experimentation, and ongoing refinement of the feedback messages. Although we have achieved a maturity level that allows us to provide highly informative feedback to students, a higher level of automation of the systems is needed to scale our HILA system for educational practices and roll it out in an organisation. Creating textual templates that explain the results of an indicator to a learner requires the evaluation of the quality and tone of a feedback text for a LA dashboard. To fully automate the feedback delivery, we will need to connect the HILA feedback system with other institutional databases, the machine learning models and the LMS. Therefore, various software interfaces (REST APIs) need to be programmed and connected. Once the system is in place, it is important to evaluate it and make changes based on how students and teachers experience it.

5.2 Educational Challenges

The HILA research programme also faces four significant educational challenges: 1) developing a culture of HILA learning design; 2) improving the quality of automated feedback and feedback literacy of students; 3) evaluating bias in the feedback system; and 4) ensuring ecological validity in testing LA.

5.2.1 Developing an organisational culture of HILA learning design

Developing a culture of HILA requires actions on various levels. In essence, it is a matter of organisational change (Hage, 1999). First and foremost, it requires the support of the management of an organisation to support the adoption of HILA. The management can stimulate this change by establishing policies for data protection and privacy, providing significant time for training and support of the teaching staff and students, stimulating pilot studies to test the effectiveness of HILA, and increasing collaboration among different stakeholders (Hoover & Harder, 2015; Hansen et al., 2020; Tsai et al., 2018). Next to the support of the management, a HILA culture requires a commitment from the teaching staff and the students to embrace technology innovations and learn how to use them effectively (Wollny et al., 2023). The teaching staff needs to become aware of the potential of the feedback system and how to incorporate it into their daily teaching practices. They need to be trained to plan their teaching activities in more detail to get from the intended competencies and learning goals of a course to concrete learning activities that support the development of these competencies at the student's side (Ahmad et al., 2022; Schmitz et al., 2022).

5.2.2 Quality of automated feedback and literacy skills of students

Controlling for the quality of feedback given to students is essential for research on the effectiveness of automated feedback, as the quality of feedback can have a significant impact on students' learning outcomes (Nachtigall et al., 2020). Low quality feedback can be confusing, demotivating, and even counterproductive, whereas high quality feedback can be very encouraging, instructive, and supportive of students' development (Hattie & Timperley, 2007). Therefore, it is important to ensure that the feedback provided in research studies is accurate, relevant, and actionable, and that it aligns with students' needs and expectations.

Amid the quality of feedback supplied by researchers, there is a major role of learners in the utility of LA-based feedback (see Section 3.2), feedback literacy is an emerging research subject within the LA community (Tsai et al., 2022; Jivet et al., 2020). To investigate this topic, a precise measure of how well students can comprehend feedback is necessary. Multiple research groups are working on a reliable and validated psychometric measure for feedback literacy (Zhan, 2021; Song, 2022) in order to investigate the effects of different types of feedback on distinct student populations in authentic learning environments. To investigate how students perceive, engage with, and accept LA-based feedback, a thorough theoretical understanding of feedback literacy as a construct for measuring it is required. Pre-post experiments with an experimental group receiving personalised HILA feedback and a control group receiving non-personalised feedback are required for assessing whether and to what extent feedback literacy plays an important role in the effectiveness of HILA.

5.2.3 Ethics and biases within HILA

There are a number of issues associated with the ethics and biases of HILA systems. An example of these issues is, for instance, the dominance of white or male students in a study programme, as well as the labelling of data for machine learning by a homogeneous group as opposed to a heterogeneous group of coders as it has been researched on platforms like Mechanical Turk (Difallah et al., 2018). In order to learn and make decisions, the HILA machine learning models rely on data. If there is a bias in the data, then also the HILA system is biased, which is also known under the term 'algorithmic bias'. In the field of education, this may indicate that the HILA systems are biased against particular student groups, resulting in unequal treatment or outcomes (Dieterle et al., 2022). Next to bias in data, a machine learning model can be opaque, making it difficult to comprehend how decisions are made. This lack of transparency makes it difficult to identify and correct systemic prejudice. With respect to privacy, a HILA system may gather and analyse sensitive information about students, including their academic achievement, conduct, and affective learning characteristics. This raises worries about privacy and the possible misuse of this information. Within the HILA research programme, we have established a code of behaviour for the implementation of HILA in an effort to mitigate these risks (Hansen et al., 2020). This code of conduct also specifies HILA's defined accountability roles. Thus, it is essential to approach the use of HILA with a critical mind to monitor potential risks like ethics, bias and privacy.

5.2.4 Ecological valid testing of HILA

The development and use of LA systems outside of educational practice is not very promising, because the research into effective feedback formats can only be empirically advanced in authentic educational contexts with adequate consideration of relevant parameters and their interactions. It is therefore essential for HILA to support learners and teachers by developing digital learning environments that offer real added value for actual teaching in order to obtain ecologically valid data about HILA and deepen our research findings.

Being a discipline devoted to understanding and enhancing learning, LA must be able to give empirical evidence for cause and effect in learning. Despite their facility to infer causality, randomised controlled field trials are not consequently done among the LA community (Sonderlund et al., 2018; Viberg et al., 2018), while the AIED community is for instance more used to randomised controlled trials. Until recently, many characterisations of LA research methods were limited to observation of student data generated from real educational systems, with inferences gleaned primarily from statistical modelling, visualisations, and dashboards based on these existing data resources. Instead of comparative research, observational studies, from which it is significantly more difficult to discern causal relationships, are regularly applied in this field (Weidlich et al., 2022). In effect, this means that the ecological validity and practical usability of the conclusions for informing educational decisions, i.e. the implementation of HILA feedback, are often compromised. Large-scale randomised field studies are recognised as the gold standard in educational science for addressing this deficiency (Styles & Torgerson, 2018), but they then need interventions that can be scaled to hundreds of students in diverse, realistic environments, posing a significant logistical and methodological challenge. The HILA researchers are committed to striving for this gold standard and providing empirical evidence on the effects of HILA for an empirical body of knowledge on LA. However, where practical constraints do not allow for fully randomised field trials, HILA will make use of state-of-the-art methods to infer causality, even in research designs that are not strictly experimental. A promising, yet largely underutilised approach to this is the use of Directed Acyclic Graphs (Weidlich et al., 2022) to reason about and communicate causal assumptions.

6 CONCLUSIONS

In conclusion, this book provides a comprehensive overview of the TLA research foundation and current challenges of the HILA research programme. Building upon the foundational elements of TLA, such as Stakeholders, Objectives, Data, Instruments, External constraints, and Internal limitations, the necessary advancements to HILA are explained. The book explores related research areas, including psychometrics, feedback theory, and learning design, to generate highly informative feedback for students and teachers. A range of DeLA technologies, such as those for reading, writing, collaboration, and concept modelling, are demonstrated as sources to collect relevant data and provide HILA feedback. Additionally, potential study designs and the main research questions used to investigate the effects of HILA in educational settings are presented. Finally, an in-depth analysis of the technical and educational challenges associated with implementing HILA is provided, including the extraction of learner attributes, annotation of textual data, delivering automated feedback, developing an organisational culture for HILA, ensuring the quality of feedback and feedback literacy skills, addressing ethics and biases, and conducting ecologically valid testing for the field of LA.

In the concluding paragraphs, it is pertinent to emphasise certain ideas of significance for the HILA research programme. The overall objective of HILA is to improve the quality of feedback that students and teachers receive from LA. The programme has three main focus areas that are aimed at achieving this goal.

The first focus area is on the micro-level of LA. The HILA programme aims to give students and teachers meaningful feedback on their learning experiences. This focus on the micro-level is important because it provides a more detailed understanding of the learning process and allows for more effective and efficient feedback. By using machine learning, the HILA programme can make certain teacher tasks like assessment and feedback less time-consuming and more efficient, giving teachers more time to create a rich and interesting learning environment for their students. In turn, teachers can use HILA to provide immediate and personalised feedback on the student's outcomes, resulting in motivated and excited students and teachers who can provide effective feedback and create an enjoyable learning environment. To increase the effectiveness of LA-based feedback, the HILA programme shifts its attention from *providing feedback* to *receiving feedback*. This necessitates taking into account both the characteristics and perspectives of students as well as having rich and relevant data that can be utilised by LA and teachers to provide HILA feedback.

The second focus area of the HILA program is the acquisition of rich and relevant data about learning through the DeLAs. These DeLAs aim to provide comprehensive coverage of learning activities and can be applied to almost any study programme. Their context and discipline independence allow to serve various study fields and provide HILA feedback to a diverse range of stakeholders ranging from higher education to schools. By collecting data that is meaningful and aligned with the actual learning tasks, DeLAs can offer valuable feedback to its users.

The third focus area of the HILA programme is on building a body of knowledge for LA. This involves critically considering already applied LA approaches and studying them in different settings and with different stakeholder groups to receive an evidence-based inventory of the effects of LA. HILA therefore, wants to explore the intersection or the middle space between highly controlled assessment in the field of psychometrics and broadly applicable DeLAs that collect relevant data for the learning process and the learning products. The DeLAs are a powerful mean to achieve this intersection as DeLAs offer fairly stable and reliable conditions that make them well suited for experimental settings to test and evaluate the effects of LA interventions in various settings and with different stakeholder groups. They therefore ultimately contribute to building a body of knowledge and provide insights into effective and less effective approaches for different stakeholders.

The HILA approach has faced criticism regarding its implementation in institutions, as it necessitates significant effort and resources. However, based on the experiences gained so far, it is acknowledged that the initial stages of HILA implementation are labor-intensive, but as the process progresses, there is potential for synergies and increased effectiveness and efficiency in rolling out HILA solutions. The DeLAs serve as an exemplar for this, as they require substantial effort during their development but are relatively effortless in transferring them to new contexts and additional learning scenarios once they are created. DeLAs have been successfully transferred from the higher education sector to the school sector, and this process will continue. As the collection of DeLAs grows, it will be possible to address a high percentage of common teaching and learning situations in almost all faculties and thus provide HILA feedback to vast numbers of students while reducing the time investment of teachers.

On a personal node, as a computer and educational scientist, I have actively worked over the past to build LA that carry a value for educational stakeholders and can prove this value by empirically evidence. In my opinion, this is only achievable through long-term, interdisciplinary collaboration that combines computer science research with educational reality. The transfer of LA research and the collaboration with the practioners are not only desirable, but also the only way to gain research insights into teaching and learning in ecologically valid contexts. Machine learning techniques applied in HILA require authentic data to foster individual learning processes. In order to obtain ecologically valid an authentic learning data, it is necessary to develop digital platforms, such as the HILA infrastructure, that provide real value to both students and teachers, considering the current educational context and individual learning goals. The HILA research, therefore, stands complementary to internally valid but reductionist lab research. A variety of LA modelling levels (e.g., combination of different DeLAs) are systemically needed to accurately address individual learning processes. Moreover, the variety of these LA models also needs to be orchestrated and researched in their combination.

This is not a trivial task and therefore requires strong institutional partners with complimentary profiles. I am privileged to follow the HILA research agenda with the strong background in computer science at the DIPF Leibniz Institute and the computer science faculty of the Goethe University of Frankfurt, as well as the educational faculty of the Open Universiteit and the research group on Online Learning and Instruction. This outstanding partnership has the potential to achieve the objectives of the HILA research agenda. Together, we can provide valuable insights and tangible outcomes for students, teachers, researchers, practitioners, and policymakers interested in HILA feedback. Therefore, let me close the speech with a quote by Andy Hargreaves that succinctly sums up the HILA research efforts in one sentence:

"*Measure what you value; don't value what you can easily measure.*"

Ik heb gezegd.

7 REFERENCES

Ahmad, A., Schneider, J., Griffiths, D., Biedermann, D., Schiffner, D., Greller, W., & Drachsler, H. (2022). Connecting the dots – A literature review on learning analytics indicators from a learning design perspective. *Journal of Computer Assisted Learning*. https://doi.org/10.1111/jcal.12716

Ahmad, A., Schneider, J., & Drachsler, H. (2020). OpenLAIR an Open Learning Analytics Indicator Repository Dashboard. In *European Conference on Technology Enhanced Learning* (pp. 467-471). Springer, Cham.

Alvarez, C. P., Martinez-Maldonado, R., & Buckingham Shum, S. (2020). LA-DECK: A card-based learning analytics co-design tool. *Proceedings of the 10th International Conference on Learning Analytics and Knowledge (LAK 2020)*, 23–27 March 2020, online (pp. 63–72). ACM. https://doi.org/10.1145/3375462.3375476

Arnold K. E., & Pistilli M. D. (2012). Course signals at Purdue: using learning analytics to increase student success. In: *Proceedings of the 2nd International Conference on Learning Analytics and Knowledge (LAK '12)*. Association for Computing Machinery, New York, NY, USA, 267–270. https://doi.org/10.1145/2330601.2330666

Austin, L. B., & Shore, B. M. (1995). Using concept mapping for assessment in physics. *Physics Education, 30*(1), 41.

Bakharia, A., & Dawson, S. (2011). SNAPP: a bird's-eye view of temporal participant interaction. In *Proceedings of the 1st international conference on learning analytics and knowledge* (pp. 168-173).

Banihashem, S. K., Noroozi, O., Van Ginkel, S., Macfadyen, L. P., & Biemans, H. J. (2022). A systematic review of the role of learning analytics in enhancing feedback practices in higher education. *Educational Research Review*, 100489.

Berg, A., Scheffel, M., Drachsler, H., Ternier, S., & Specht, M. (2016). Dutch Cooking with xAPI Recipes: The Good, the Bad and the Consistent. *In Proceedings of the International Conference on Advanced Learning Technologies, ICALT '16*, 234–236. DOI: 10.1109/ICALT.2016.48.

Biedermann, D., Kalbfell, L., Schneider, J., & Drachsler, H. (2019). Stakeholder attitudes towards digitalization in higher education institutions. *DeLFI 2019 – Die 17. Fachtagung Bildungstechnologien der Fachgruppe Bildungstechnologien der gi e.V. Humboldt Universität zu Berlin*, 16.-19. September 2019.

Biedermann, D., Schneider, Ciordas-Hertel, G-P., Hahnel, C., Eichmann, B., Goldhammer, F. & Drachsler, H. (2023). Detecting the Disengaged Reader - Using Scrolling Data to Predict Disengagement during Reading. In *LAK23: 13th International Learning Analytics and Knowledge Conference (LAK23).* Association for Computing Machinery, New York, NY, USA.

Biedermann, D., Ciordas-Hertel, G-P., Winter, M., Mordel, J., & Drachsler, H. (2023b). Contextualized Logging of On-Task and Off-Task Behavior during Learning. *Journal of Learning Analytics*.

Biggs, J. B. (1987). *Learning Process Questionnaire Manual. Student Approaches to Learning and Studying.* Australian Council for Educational Research Ltd., Radford House, Frederick St., Hawthorn 3122, Australia.

Biggs, J., and Tang, C. (2011). EBOOK: *Teaching for Quality Learning at University*. McGraw hill education (UK).

Buckingham Shum, S. (2012). *Learning Analytics Policy Brief.* UNESCO Institute of Technologies in Education. Available at: http://iite.unesco.org/files/policy_briefs/pdf/en/learning_analytics.pdf

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centered learning analytics. *Journal of Learning Analytics, 6*(2), 1-9.

Burrows, S., Gurevych, I. & Stein, B. The Eras and Trends of Automatic Short Answer Grading. *Int J Artif Intell Educ 25*, 60–117 (2015). https://doi.org/10.1007/s40593-014-0026-8

Carless, D., & Boud, D. (2018). The development of student feedback literacy: enabling uptake of feedback. *Assessment & Evaluation in Higher Education*, *43*(8), 1315-1325.

Carrillo, R., Prié, Y., & Lavoué, É. (2019). Observing learner engagement on mind mapping activities using Learning Analytics. In *Transforming Learning with Meaningful Technologies: 14th European Conference on Technology Enhanced Learning, EC-TEL 2019, Delft, The Netherlands, September 16–19, 2019, Proceedings 14* (pp. 668-672). Springer International Publishing

Catrysse, L., Gijbels, D., & Donche, V. (2018). It is not only about the depth of processing: What if eye am not interested in the text? *Learning and Instruction, 58*, 284-294

Cañas, A. J., & Novak, J. D. (2014). Concept mapping using CmapTools to enhance meaningful learning. *Knowledge cartography: Software tools and mapping techniques*, 23-45.

Cañas, A. J., Novak, J. D., & Reiska, P. (2015). How good is my concept map? Am I a good Cmapper?. *Knowledge Management & E-Learning, 7*(1), 6.

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2016). Creating effective learning analytics dashboards: Lessons learnt. In *European Conference on Technology Enhanced Learning* (pp. 42-56). Springer, Cham.

Chang, K. E., Sung, Y. T., & Chen, I. D. (2002). The effect of concept mapping to enhance text comprehension and summarization. *The Journal of Experimental Education, 71*(1), 5-23.

Chang, K. E., Sung, Y. T., Chang, R. B., & Lin, S. C. (2005). A new assessment for computer-based concept mapping. *Journal of Educational Technology & Society, 8*(3), 138-148.

Choi, H., Winne, P. H., Brooks, C., Li, W., & Shedden, K. (2023). Logs or Self-Reports? Misalignment Between Behavioral Trace Data and Surveys When Modeling Learner Achievement Goal Orientation. In *Proceedings of the Thirteens International Learning Analytics & Knowledge Conference (LAK*23). ACM, New York, USA, 11–21. https://doi.org/10.1145/3576050.3576052

Cocea, M., & Weibelzahl, S. (2011). Disengagement Detection in Online Learning: Validation Studies and Perspectives. *IEEE Transactions on Learning Technologies 4*, *2* (April 2011), 114–124. https://doi.org/10.1109/TLT.2010.14.

Ciordas-Hertel, G., P., Schneider, J., Ternier, S., & Drachsler, H. (2019). Adopting Trust in Learning Analytics Infrastructure: A Structured Literature Review. *Journal of Universal Computer Science, Vol. 25, No. 13*, pp. 1668-1686.

Ciordas-Hertel, G. P., Schneider, J., Ternier, S., & Drachsler, H. (2020a). Adopting trust in learning analytics infrastructure: A structured literature review. *Journal of Universal Computer Science, 25*(13), 1668–1686. https://doi.org/10.3217/jucs-025-13-1668

Ciordas-Hertel, G. P., Schneider, J., & Drachsler, H. (2020b). Which strategies are used in the design of technical LA infrastructure?: A qualitative interview study. *IEEE Global Engineering Education Conference (EDUCON), Porto, Portugal, 2020*, 96–105. https://doi.org/10.1109/EDUCON45650.2020.9125363

Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G. J. (2017). Follow the successful crowd: raising MOOC completion rates through social comparison at scale. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17). ACM, New York, NY, USA*, 454-463. DOI: 10.1145/3027385.3027411.

De Bruyckere, P., Kirschner, P. A., & Hulshof, C. D. (2015). Chapter 4—Myths about Technology in Education. In P. D. Bruyckere, P. A. Kirschner, & C. D. Hulshof (Eds.), *Urban Myths about Learning and Education* (pp. 127–164). Academic Press. https://doi.org/10.1016/B978-0-12-801537-7.00005-6

De Bruyckere, P., Kirschner, P.A., & Hulshof, C. (2019). *More Urban Myths About Learning and Education: Challenging Eduquacks, Extraordinary Claims, and Alternative Facts (1st ed.)*. Routledge. https://doi.org/10.4324/9781351132435

Devlin, J., Chang, M.-W., Lee, K., & Toutanova, K. (2019). BERT: Pre-training of Deep Bidirectional Transformers for Language Understanding. *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, 4171–4186. https://doi.org/10.18653/v1/N19-1423

Dieterle, E., Dede, C., & Walker, M. (2022). The cyclical ethical effects of using artificial intelligence in education. *AI & society*, 1-11.

Difallah, D., Filatova, E., & Panos, I. (2018). Demographics and Dynamics of Mechanical Turk Workers. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining (WSDM '18). Association for Computing Machinery, New York, USA*, 135–143. https://doi.org/10.1145/3159652.3159661

Di Mitri, D., Schneider, J., Specht, M., & Drachsler, H. (2018). From signals to knowledge: A conceptual model for multimodal learning analytics. *Journal of Computer Assisted Learning. 34*: 338– 349. https://doi.org/10.1111/jcal.12288

Di Mitri, D., Schneider, J., & Drachsler, H. (2021). Keep Me in the Loop: Real-Time Feedback with Multimodal Data. *International Journal of Artificial Intelligence in Education.* https://doi.org/10.1007/s40593-021-00281-z

Di Mitri, D., Schneider, J., Trebing, K., Sopka, S., Specht, M., and Drachsler, H. (2020). Real-Time Multimodal Feedback with the CPR Tutor. In: Bittencourt I., Cukurova M., Muldner K., Luckin R., Millán E. (eds) *Artificial Intelligence in Education*. AIED 2020. Lecture Notes in Computer Science, vol 12163. Springer, Cham. Doi: https://doi.org/10.1007/978-3-030-52237-7_12

Dowell, N. M. M., Nixon, T. M., & Graesser, A. C. (2019). Group communication analysis: A computational linguistics approach for detecting sociocognitive roles in multiparty interactions. *Behavior Research Methods*, 51(3), 1007–1041. https://doi.org/10.3758/s13428-018-1102-z

Drachsler, H., Jansen, J., & Kirschner, P. A. (2021). Adoption of learning technologies in times of pandemic crisis. *Journal of Computer Assisted Learning, 37*: 1509-1512. https://doi.org/10.1111/jcal.12626

Drachsler, H., Stoyanov, S., & Specht, M. (2014). The impact of learning analytics on the Dutch education system. In *Proceedings of the Fourth International Conference on Learning Analytics and Knowledge, LAK '14.* pages 158-162, New York, NY, USA. ACM. DOI: 10.1145/2567574.2567617.

Drachsler, H. & Greller, W. (2016). Privacy and Analytics – it's a DELICATE issue. A Checklist to establish trusted Learning Analytics. In *Proceedings of the Sixth International Conference on Learning Analytics and Knowledge, LAK '16*, pages 89–98, New York, NY, USA. ACM. DOI: 10.1145/2883851.2883893.

Drachsler, H., & Goldhammer, F. (2020). Learning Analytics and eAssessment—Towards Computational Psychometrics by Combining Psychometrics with Learning Analytics. In. Burgos, D. (Ed.) (2020). *Radical Solutions and Learning Analytics*. pp. 67-80. Singapore: Springer. DOI: https://doi.org/10.1007/978-981-15-4526-9

Fazeli, S., Drachsler, H., Bitter-Rijpkema, B., Brouns, F., Van der Vegt, W., & Sloep, P. B. (2017). User-centric Evaluation of Recommender Systems in Social Learning Platforms: Accuracy is Just the Tip of the Iceberg. *IEEE Transactions on Learning Technologies.* DOI: 10.1109/TLT.2017.2732349.

Ferguson, R., & Clow, D. (2017). Where is the evidence? A call to action for learning analytics. In: *LAK '17 Proceedings of the Seventh International Learning Analytics & Knowledge Conference, ACM International Conference Proceeding Series, ACM, New York, USA*, pp. 56–65. DOI: https://doi.org/10.1145/3027385.3027396

Gasevic, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends 59*(1), 64–71.

Giorgashvili, T., Jivet, I. & Drachsler, H. (2023) Closing the Feedback Loop - A Moodle Plugin for (Semi-)automated Feedback. European Association for Research on Learning and Instruction (EARLI 2023), Thessaloniki, Greece.

Giorgashvili, T., Jivet, I., Biedermann, D., Ciordas-Hertel, G. P., Mendzheritskaya, J., Winter, M., Bengs, D., Eichmann, B., Hahnel, C., Goldhammer, F., & Drachsler, H. (submitted a). For your eyes only: How students reflect on personalized learning analytics feedback in large lectures: a case study. In *Proceedings of the 18th European Conference on Technology Enhanced Learning (EC-TEL 2023)* (pp. 70-81).

Giorgashvili, T., Jivet, I., Mendzheritskaya, J., Winter, M., Catrysse, L., Jarodzka, H., & Drachsler, H. (submitted b). A webcam-based eye-tracking study on a concept mapping dashboard for higher education students. *Journal of Computer Assisted Learning.*

Goedecke, P. J., Dong, D., Shi, G., Feng, S., Risko, E., Olney, A. M., D'Mello, S., & Graesser, A. C. (2015). *Breaking off Engagement: Readers' Disengagement as a Function of Reader and Text Characteristics.* International Educational Data Mining Society. https://eric.ed.gov/?id=ED560881

Goodyear P., & Dimitriadis Y. (2013). In medias res reframing design for learning. *Research in Learning Technology*, *21*. https://doi.org/10.3402/rlt.v21i0.19909

Gombert, S., Di Mitri, D., Karademir, O., Kubsch, M., Kolbe, H., Tautz, S., Grimm, A., Bohm, I., Neumann, K., & Drachsler, H. (2022). Coding educational concepts in students essays with explainable NLP models. *Journal of Computer Assisted Learning.* https://doi.org/10.1111/jcal.12767

Gombert, S., Fink, A., Giorgashvii, T., Jivet, I., Di Mitri, D., Yau, J., Frey, A., & Drachsler, H. (submitted). From the automated assessment of student essay content to highly informative feedback: a case study. *Journal of Artificial Intelligence in Education*, AIED.

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

Hage, J. (1999). Organizational Innovation and Organizational Change. Annual Review of Sociology - ANNU REV SOCIOL. 25. 10.1146/annurev.soc.25.1.597.

Hahnel, C., Kroehne, U., Goldhammer, F., Schoor, C., Mahlow, N., & Artelt, C. (2019). Validating process variables of sourcing in an assessment of multiple document comprehension. *Br J Educ Psychol, 89*: 524-537. https://doi.org/10.1111/bjep.12278

Harks, B., Rakoczy, K., Hattie, J., Besser, M., & Klieme, E. (2014). The effects of feedback on achievement, interest and self-evaluation: the role of feedback's perceived usefulness, *Educational Psychology, 34*:3, 269-290. https://doi.org/10.1080/01443410.2013.785384

Hansen, J., Rensing, C., Hermann, O., & Drachsler, H. (2020). *Verhaltenskodex für Trusted Learning Analytics: Entwurf für die Hessischen Hochschulen*. Frankfurt am Main: Innovationsforum Trusted Learning Analytics Goethe-Universität Frankfurt am Main. Available at: https://www.researchgate.net/publication/340183667_Verhaltenskodex_fur_Trusted_Learning_Analytics.

Hattie, J. (2009). *Visible learning: A synthesis of 800+ meta-analyses on achievement*. London: Routledge.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

Hicks, B., Kitto, K., Payne, L., & Buckingham Shum, S. (2022). Thinking with causal models: A visual formalism for collaboratively crafting assumptions. *Proceedings of the 12th International Conference on Learning Analytics and Knowledge (LAK '22)*, *21–25 March 2022, Online* (pp. 250–259). ACM Press. https://doi.org/10.1145/3506860.3506899

Hoover, E., & Harder, M. K. (2015). What lies beneath the surface? The hidden complexities of organizational change for sustainability in higher education. *Journal of cleaner production*, 106, 175-188. https://www.sciencedirect.com/science/article/abs/pii/S0959652614001073

Jivet, I., Scheffel, M., Schmitz, M., Robbers, S., Specht, M., & Drachsler, H. (2020). From students with love: An empirical study on learner goals, self-regulated learning and sense-making of learning analytics in higher education. *Internet and Higher Education, 47*, 100758. https://doi.org/10.1016/j.iheduc.2020.100758

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness is not enough: pitfalls of learning analytics dashboards in the educational practice. In *European Conference on Technology Enhanced Learning* (pp. 82-96). Springer, Cham.

Jivet, I., Wong, J., Scheffel, M., Valle Torre, M., Specht, M., & Drachsler, H. (2021). Quantum of Choice: How learners' feedback monitoring decisions, goals and self-regulated learning skills are related. *In LAK21: 11th International Learning Analytics and Knowledge Conference (LAK21). ACM. USA*, 416–427.

Kalz, M., Van Bruggen, J., Giesbers, B., Waterink, W., Eshuis, J., & Koper, R. (2014). A study about placement support using semantic similarity. *Journal of Educational Technology & Society, 17*(3), 54-64.

Karademir, O., Ahmad, A., Schneider, J., Di Mitri, D., Jivet, I., & Drachsler, H. (2021). Designing the Learning Analytics Cockpit-A Dashboard that Enables Interventions. In *International Conference on Methodologies and Intelligent Systems for Technology Enhanced Learning* (pp. 95-104). Springer, Cham.

Karademir, O., Ahmad, A., Schneider, J., Di Mitri, D., Jivet, I., & Drachsler, H. (submitted). Co-designing a Learning Analytics Dashboard with Teachers - A Dashboard that Enables Interventions. *Journal of Computer Assisted Learning*.

Kreijns, K., Kirschner, P. A., & Jochems, W. (2003). Identifying the pitfalls for social interaction in computer-supported collaborative learning environments: a review of the research. *Computers in human behavior, 19*(3), 335-353.

Kim, J., Jo, I. H., & Park, Y. (2016). Effects of learning analytics dashboard: analyzing the relations among dashboard utilization, satisfaction, and learning achievement. *Asia Pacific Education Review, 17*(1), 13-24.

Kirschner, P. A. (2017). Stop propagating the learning styles myth. Computers & Education, 106, 166-171.

Kirschner, P., Kreijns, K., Phielix, C., & Fransen, J. (2015). Awareness of cognitive and social behaviour in a CSCL environment. *Journal of Computer Assisted Learning, 31*(1):59–77.

Krumm, A. E., Waddington, R. J., Lonn, S., & Teasley, S. D. (2014). A Learning Management System-Based Early Warning System for Academic Advising in Undergraduate Engineering. In: Larusson J., White B. (eds) *Learning Analytics.* Springer, New York, NY. DOI: 10.1007/978-1-4614-3305-7_6

Koper, R., & Olivier, B. (2004). Representing the learning design of units of learning. *Journal of Educational Technology & Society, 7*(3), 97-111.

Kollom, K., Tammets, K., Scheffel, M., Tsai, Y.S., Jivet, I., Muñoz-Merino, P.J., Moreno-Marcos, P.M., Whitelock-Wainwright, A., Ruiz Calleja, A., Gasevic, D., Delgado Kloos C., Drachsler, H., & Ley, T. (2020). A four-country cross-case analysis of academic staff expectations about learning analytics in higher education. *The Internet and Higher Education*, 2020, 100788, ISSN 1096-7516 https://doi.org/10.1016/j.iheduc.2020.100788.

Kovanović, V., Gašević, D., Dawson, S., Joksimović, S., Baker, R. S., & Hatala, M. (2015). Penetrating the black box of time-on-task estimation. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 184-193).

Kube, D., Weidlich, J., Jivet, I., Kreijns, K., & Drachsler, H. (2022). "Gendered differences versus doing gender": a systematic review on the role of gender in CSCL. Unterrichtswissenschaft, 1-28. https://doi.org/10.1007/s42010-022-00153-y

Kube, D., John, N., Gombert, S., Weidlich, J., Kreijns, K., & Drachsler, H. (submitted). Escaping Binary Gender Roles: Gender Diversity Dynamics in a CSCL-Escape Game. *Journal of Computer Assisted Learning*

Kubsch M., Czinczel B., Lossjew J., Wyrwich T., Bednorz D., Bernholt S., Fiedler D., Strauß S., Cress U., Drachsler H., Neumann K., & Rummel N. (2022). Toward learning progression analytics — Developing learning environments for the automated analysis of learning using evidence-centered design. *Front. Educ. 7*:981910. https://doi.org/10.3389/feduc.2022.981910

Law, N., & Liang, L. (2020). A multilevel framework and method for learning analytics integrated learning design. *Journal of Learning Analytics, 7*(3), 98-117. https://doi.org/10.18608/jla.2020.73.8

Lipnevich A.A., & Panadero, E. (2021) A Review of Feedback Models and Theories: Descriptions, Definitions, and Conclusions. *Front. Educ.* 6:720195. https://doi.org/10.3389/feduc.2021.720195

Lipnevich, A. A., & Lopera-Oquendo, C. (2022). Receptivity to instructional feedback: A validation study in the secondary school context in Singapore. *European Journal of Psychological Assessment*. Advance online publication. https://doi.org/10.1027/1015-5759/a000733

Lipnevich, A. A., & Smith, J. K. (2022). Student – Feedback Interaction Model: Revised. *Studies in Educational Evaluation*, *75*, 101208. https://doi.org/10.1016/j.stueduc.2022.101208

Lim, L. A., Dawson, S., Gašević, D., Joksimović, S., Pardo, A., Fudge, A., & Gentili, S. (2021). Students' perceptions of, and emotional responses to, personalised learning analytics-based feedback: an exploratory study of four courses. *Assessment & Evaluation in Higher Education, 46*(3), 339-359. https://doi.org/10.1080/02602938.2020.1782831

Lui, A. M., & Andrade, H. L. (2022, March). Inside the Next Black Box: Examining Students' Responses to Teacher Feedback in a Formative Assessment Context. In *Frontiers in Education* (Vol. 7, p. 751549). Frontiers Media SA. https://doi.org/10.3389/feduc.2022.751549

Macfadyen, L. P., Lockyer, L., & Rienties, B. (2020). Learning design and learning analytics: Snapshot 2020. *Journal of Learning Analytics, 7*(3), 6–12. https://doi.org/10.18608/jla.2020.73.2

Malone, J., & Dekkers, J. (1984). The concept map as an aid to instruction in science and mathematics. *School science and mathematics, 84*(3), 220-231.

Mangaroska, K., Sharma, K., Gasevic, D., & Giannakos, M. (2020). Multimodal learning analytics to inform learning design: Lessons learned from computing education. *Journal of Learning Analytics, 7*(3), 79–97. https://doi.org/10.18608/jla.2020.73.7

Menzel, L., Gombert, S., Di Mitri, D., & Drachsler, H. (2022). Superpowers in the Classroom: Hyperchalk is an Online Whiteboard for Learning Analytics Data Collection. In Educating for a New Future: Making Sense of Technology-Enhanced Learning Adoption. *17th European Conference on Technology Enhanced Learning* (ECTEL 2022). Lecture Notes in Computer Science, vol 13450. Springer International Publishing. https://doi.org/10.1007/978-3-031-16290-9_37

Menzel, L., Gombert, S., Weidlich, J., Fink, A., Frey, A. & Drachsler, H. (submitted). Why you should give your students automatic process feedback on their collaboration: Evidence from a randomized experiment. *18th European Conference on Technology Enhanced Learning*.

Meurers, D., De Kuthy, K., Nuxoll, F., Rudzewitz, B., & Ziai, R. (2019). Scaling Up Intervention Studies to Investigate Real-Life Foreign Language Learning in School. *Annual Review of Applied Linguistics, 39*, 161-188. doi:10.1017/S0267190519000126

Mills, C., & D'Mello, S. (2015). Toward a Real-Time (Day) Dreamcatcher: Sensor-Free Detection of Mind Wandering during Online Reading. *International Educational Data Mining Society*. https://eric.ed.gov/?id=ED560533.

Mislevy, R. J., Almond, R. G., & Lukas, J. F. (2003). A brief introduction to evidence-centered design. *ETS Research Report Series, 2003*(1), i-29. doi:10.1002/j.2333-8504.2003.tb01908.x

Mislevy, R. (2019). On integrating psychometrics and learning analytics in complex assessments. In H. Jiao, R. W. Lissitz, & A. van Wie (Eds.), *Data analytics and psychometrics* (pp. 1–52). Charlotte, NC, USA: Information Age Publishing.

Motz, B., Quick, J., Brooks, C., Bergner, Y., Gray, G., Lang, C., Li, W., & Marmolejo-Ramos, F. (2022, April 3). A LAK of Direction: Misalignment Between the Goals of Learning Analytics and its Research Scholarship. *EdArXiv. April 3.* doi:10.35542/osf.io/bzahw.

Nachtigall, V., Serova, K., & Rummel, N. (2020). *When failure fails to be productive: Probing the effectiveness of productive failure for learning beyond stem domains.* Instructional Science. Advance online publication. https://doi.org/10.1007/s11251-020-09525-2

Nelson, T. O., & Narens, L. (1994). Why investigate Metacognition. In J. Metcalfe & A. P. Shimamura (Eds.), Metacognition. *Knowing about Knowing.* (pp. 1-25). Cambridge, Mass: Cambridge University Press.

Novak, J. D., & Cañas, A. J. (2006). The theory underlying concept maps and how to construct them. *Florida Institute for Human and Machine Cognition, 1*(1), 1-31.

OECD. (2013). *OECD Skills Outlook 2013: First Results from the Survey of Adult Skills.* Paris: OECD Publishing.

OECD. (2016). *PISA 2015 Results (Volume I): Excellence and Equity in Education.* Paris: OECD Publishing.

Pardo, A., Bartimote, K., Buckingham Shum, S., Dawson, S., Gao, J., Gašević, D., Leichtweis, S., Liu, D., Martínez-Maldonado, R., Mirriahi, N., Moskal, A. C. M., Schulte, J., Siemens, G., & Vigentini, L. (2018). OnTask: Delivering Data-Informed, Personalized Learning Support Actions. *Journal of Learning Analytics, 5*(3), 235–249. https://doi.org/10.18608/jla.2018.53.15

Panadero, E., & Lipnevich, A.A. (2022). A review of feedback models and typologies: Towards an integrative model of feedback elements. *Educational Research Review, 35*, 100416. https://www.sciencedirect.com/science/article/pii/S1747938X21000397

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1993). Reliability and predictive validity of the Motivated Strategies for Learning Questionnaire (MSLQ). *Educational and psychological measurement, 53*(3), 801-813.

Quick, J., Motz, B., Israel, J., & Kaetzel, J. (2020). What college students say, and what they do: aligning self-regulated learning theory with behavioral logs. *In Proceedings of the Tenth International Conference on Learning Analytics & Knowledge (LAK '20).* Association for Computing Machinery, New York, NY, USA, 534–543. https://doi.org/10.1145/3375462.3375516

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

Rakoczy, K., Pinger, P., Hochweber, J., Klieme, E., Schütze, B., & Besser, M. (2019). Formative assessment in mathematics: Mediated by feedback's perceived usefulness and students' self-efficacy. *Learning and Instruction, 60*, 154-165. https://doi.org/10.1016/j.learninstruc.2018.01.004

Rakoczy, K., Harks, B., Klieme, E., Blum, W., & Hochweber, J. (2013). Written feedback in mathematics: Mediated by students' perception, moderated by goal orientation. *Learning and Instruction, 27*, 63-73. https://doi.org/10.1016/j.learninstruc.2013.03.002

Schaal, S. (2008). Concept mapping in science education assessment: An approach to computer-supported achievement tests in an interdisciplinary hypermedia learning environment. In *Proceedings of the 3rd International Conference on Concept Mapping*.

Saqr, M., Jovanovic, J., Viberg, O., & Gašević, D. (2022). Is there order in the mess? A single paper meta-analysis approach to identification of predictors of success in learning analytics. *Studies in Higher Education, 47*(12), 2370-2391. https://doi.org/10.1080/03075079.2022.2061450

Saqr, M., & López-Pernas, S. (2022). How CSCL roles emerge, persist, transition, and evolve over time: A four-year longitudinal study. *Computers & Education*, 189, 104581. https://doi.org/10.1016/j.compedu.2022.104581

Scheffel, M., Drachsler, H., De Kraker, J., Kreijns, K., Slootmaker, A., & Specht, M. (2017). Widget, widget on the wall, am I performing well at all? *IEEE Transactions on Learning Technologies, 10*(1):42–52. DOI: 10.1109/TLT.2016.2622268.

Scheffel, M., Tsai, Y.-S., Gašević, D., & Drachsler, H. (2019). Policy matters: Expert recommendations for learning analytics policy. In M. Scheffel, J. Broisin, V. Pammer-Schindler, A. Ioannou, & J. Schneider (Eds.), *Proceedings of the 14th European Conference on Technology Enhanced Learning (EC-TEL 2019)*, LNCS, Springer, Cham.

Schmitz, M., Scheffel, M., Bemelmans, R., & Drachsler, H. (2022). FoLA2 — A Method for Co-creating Learning Analytics–Supported Learning Design. *Journal of Learning Analytics, 9*(2), 265-281. https://doi.org/10.18608/jla.2022.7643

Schneider, J., Börner, D., Van Rosmalen, P., and Specht, M. (2017). Do you know what your nonverbal behavior communicates? Studying a self-reflection module for the presentation trainer. *In Immersive Learning Research Network: Third International Conference*, iLRN 2017, Coimbra, Portugal, June 26–29, 2017. Proceedings 3 (pp. 93-106). Springer International Publishing.

Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research, 78*(1), 153–189. https://doi.org/10.3102/0034654307313795

Shute, V. J. (2011). Stealth assessment in computer-based games to support learning. *Computer Games Instruction, 55*, 503–524.

Singer, N. (2014). InBloom Student Data Repository to Close. Article published online in *The New York Times*, April 21, 2014. Available at: https://bits.blogs.nytimes.com/2014/04/21/inbloom-student-data-repository-to-close/?_r=0

Smith, P. L., & Ragan, T. J. (2004). *Instructional design.* John Wiley & Sons.

Sonderlund, A., Hughes, E., and Smith, J. (2019). The efficacy of learning analytics interventions in higher education: A systematic review. *British Journal of Educational Technology*, 50(5), 2594-2618

Song, B. K. (2022). Bifactor modelling of the psychological constructs of learner feedback literacy: conceptions of feedback, feedback trust and self-efficacy. *Assessment & Evaluation in Higher Education,* 1-14.

Sousa, E.B.G., Alexandre, B., Ferreira Mello, R., Pontual Falcão, T., Vesin, B., & Gašević, D. (2021). Applications of Learning Analytics in High Schools: A Systematic Literature Review. *Front. Artif. Intell. 4*:737891. doi: 10.3389/frai.2021.737891

Styles B., & Torgerson C. (2018). Randomised controlled trials (RCTs) in education research –methodological debates, questions, challenges, *Educational Research, 60*:3, 255-264, DOI: 10.1080/00131881.2018.1500194

Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers & Education*, *89*, 53-74.

Tsai, Y.-S., Gašević, D., Whitelock-Wainwright, A., Muñoz-Merino, P. J., Moreno-Marcos, P. M., Fernández, A. R., Kloos, C. D., Scheffel, M., Jivet, I., Drachsler, H., Tammets, K., Calleja, A. R., & Kollom, K. (2018) *SHEILA: Supporting Higher Education to Intergrade Learning Analytics Research Report*. Last accessed at https://sheilaproject.eu/wp-content/uploads/2018/11/SHEILA-research-report.pdf on 27.03.2023.

Tsai, Y. S., & Martinez-Maldonado, R. (2022). Human-Centered Approaches to Data-Informed Feedback. In *Handbook of Learning Analytics* (pp. 213-222). Society for Learning Analytics Research.

Vezzoli, Y., Mavrikis, M., & Vasalou, A. (2020). Inspiration cards workshops with primary teachers in the early co-design stages of learning analytics. In *Proceedings of the tenth International Conference on Learning Analytics & Knowledge* (pp. 73-82).

Viberg, O., Hatakka, M., Baelter, O., and Mavroudi, A. (2018). The current landscape of learning analytics in higher education. *Computers in Human Behavior*, 89, 98-110.

Weidlich, J., Gasevic, D., & Drachsler, H. (2022). Causal Inference and Bias in Learning Analytics: A Primer on Pitfalls using Directed Acyclic Graphs. *Journal of Learning Analytics 9*(3), 1-17. https://doi.org/10.18608/jla.2022.7577

Weinberger, A., Stegmann, K., & Fischer, F. (2010). Learning to argue online: Scripted groups surpass individuals (unscripted groups do not). *Computers in Human Behavior*, 26(4), 506–515. https://doi.org/10.1016/j.chb.2009.08.007

Whitelock-Wainwright, A., Gašević, D., & Tejeiro, R. (2017). What do students want? towards an instrument for students' evaluation of quality of learning analytics services. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference (LAK '17). Association for Computing Machinery, New York, NY, USA,* 368–372. DOI: https://doi.org/10.1145/3027385.3027419

Whitelock-Wainwright, A., Gašević, D., Tsai, Y. S., Drachsler H., Scheffel, M., Muñoz Merino, P. J., Tammets, K., & Delgado Kloos, C. (2020). Assessing the validity of a learning analytics expectation instrument: A multinational study. *Journal of Computer Assisted Learning. 2020*; 1– 32. https://doi.org/10.1111/jcal.12401.

Wild, F., Stahl, C., Stermsek, G., and Neumann, G. (2005). Parameters driving effectiveness of automated essay scoring with LSA. *In: Proceedings of the 9th Conference on Computer Assisted Assessment (CAA)*, 5-6 Jul 2005, Loughborough, Loughborough University.

Wilson, T. D. (1999). Models in information behaviour research. *Journal of Documentation*, 55(3), 249–270. doi:10.1108/EUM0000000007145

Wisniewski, B., Zierer, K., & Hattie, J. (2020). The power of feedback revisited: A meta-analysis of educational feedback research. *Frontiers in Psychology, 10*. 1-14. doi: 10.3389/fpsyg.2019.03087

Winstone, N., & Carless, D. (2019). *Designing effective feedback processes in higher education: A learning-focused approach.* Routledge.

Woitt, S., Weidlich, J., Jivet, I., Göksün, D., Drachsler, H., and Kalz, M. (submitted). Students' Feedback Literacy in Higher Education: An Initial Scale Validation Study. *Assessment and Evaluation in Higher Education*.

Wolf M., and Barzillai, M. (2009). The importance of deep reading. *Educational Leadership 66*, 6 (March 2009).

Wollny, S., Di Mitri, D., Jivet, I., Muñoz-Merino, P., Scheffel, M., Schneider, J., Tsai, Y.-S., Whitelock-Wainwright, A., Gasevic, D., & Drachsler, H. (2023). Students' expectations of Learning Analytics across Europe. *Journal of Computer Assisted Learning,*1–14. https://doi.org/10.1111/jcal.12802

Wollny, S., Schneider, J., Di Mitri, D., Weidlich, J., Rittberger, M., & Drachsler, H. (2021). Are we there yet? – A Systematic Literature Review on Chatbots in Education. *Frontiers in Artificial Intelligence, 4*. https://doi.org/10.3389/frai.2021.654924

Zhan, Y. (2021). Developing and validating a student feedback literacy scale. *Assessment & Evaluation in Higher Education,* 1-14.

Zieky, M. J. (2014). An introduction to the use of evidence-centered design in test development. *Psicología Educativa, 20*(2), 79–87. https://doi.org/10.1016/j.pse.2014.11.003.