Computers & Education 214 (2024) 105028

Contents lists available at ScienceDirect

Computers & Education

journal homepage: www.elsevier.com/locate/compedu

Learning analytics and the Universal Design for Learning (UDL): A clustering approach

Marvin Roski [a,*], Ratan Sebastian [b], Ralph Ewerth [b,c], Anett Hoppe [b,c], Andreas Nehring [a]

[a] Leibniz University Hannover, Am Kleinen Felde 30, Hannover, 30167, Lower Saxony, Germany
[b] TIB – Leibniz Information Centre for Science and Technology, Lange Laube 28, Hannover, 30159, Lower Saxony, Germany
[c] L3S Research Center, Appelstr. 9a, Hannover, 30167, Lower Saxony, Germany

ARTICLE INFO

Keywords:
Clustering
Web-based learning science
Education inclusive education

ABSTRACT

In the context of inclusive education, Universal Design for Learning (UDL) is a framework used worldwide to create learning opportunities accessible to all learners. While much research focused on the design and students' perceptions of UDL-based learning settings, studies on students' usage patterns in UDL-guided elements, particularly in digital environments, are still scarce. Therefore, we analyze and cluster the usage patterns of 9th and 10th graders in a web-based learning platform called I$_3$Learn.

The platform focuses on chemistry learning, and UDL principles guide its design. We collected the temporal usage patterns of UDL-guided elements of 384 learners in detailed log files. The collected data includes the time spent using video and/or text as a source of information, working on learning tasks with or without help and working on self-assessments. We used Exploratory Factor Analysis (EFA) to identify relevant factors in the observed usage behaviors. Based on the factor loadings, we extracted features for k-means clustering and named the resulting groups based on their usage patterns and learner characteristics. The EFA revealed four factors suggesting that learners remain consistent in selecting UDL-guided elements that require a decision (video or text, tasks with or without help). Based on these four factors, the cluster analysis identifies six different groups. We discuss these results as a starting point to provide individualized learning support through further artificial intelligence applications and inform educators about learner activity through a dashboard.

1. Introduction

By 2030, the global implementation of "Education for All" by the United Nations aims to provide all learners equal opportunities to receive a high-quality education, regardless of their backgrounds or circumstances (UNESCO- UNEVOC, 2021; UNESCO, 2005). In this context, the Universal Design for Learning (UDL) framework provides flexible and systematic recommendations for minimizing barriers and creating optimal learning processes for all. It takes the perspective that not the diverse learners create barriers but the

* Corresponding author. Leibniz University Hannover, Institute of Science Education, Am Kleinen Felde 30, Hannover, 30167, Lower Saxony, Germany.

E-mail addresses: roski@idn.uni-hannover.de (M. Roski), ratan.sebastian@tib.eu (R. Sebastian), ralph.ewerth@tib.eu (R. Ewerth), anett.hoppe@tib.eu (A. Hoppe), nehring@idn.uni-hannover.de (A. Nehring).

https://doi.org/10.1016/j.compedu.2024.105028
Received 14 September 2023; Received in revised form 26 February 2024; Accepted 29 February 2024
Available online 2 March 2024

content to be learned or the learning environment itself (CAST, 2018). The UDL framework is an internationally renowned, widely used framework that greatly impacts educational practices in many countries (Gronseth & Dalton, 2019). The question of how learners use UDL-guided elements and which learning outcomes relate to observed usage patterns remains open (Roski et al., 2021). We define UDL-guided elements as learning opportunities implemented based on the concrete suggestions in the UDL framework (CAST, 2018). Edyburn (2021) critically examines the methodologies employed in UDL research, highlighting the lack of consensus on documenting UDL intervention outcomes and the subsequent influence on the evidence of UDL. This underscores the urgent need for a more profound understanding of the complex interplay among learner characteristics, learning content, and technology (Edyburn, 2010). The meta-analysis by King-Sears et al., 2023 shows a medium positive effect on the academic performance of learners who engaged in learning activities based on the UDL framework. However, it is also apparent that more than half of the studies do not report to what extent UDL was integrated into their learning activities. This, in turn, reflects the problems noted by Edyburn (2021). The global Sars-CoV-2 pandemic and its impacts showed how fragile educational systems worldwide can be (United Nations, 2022). Billions of learners could not attend school and had to switch to more or less well-prepared web-based remote learning (Hodges et al., 2020; United Nations, 2020), leading to an increased demand for well-designed web environments supporting learning for a maximum of students.

This need has not only increased with the pandemic. In recent years, learning trends and strategies such as blended learning or informal learning opportunities, for example, Massive Open Online Courses (MOOCs), have become very popular (Alraimi et al., 2015; Fisher et al., 2021). Given the goals of an "Education for All" approach, web-based

learning platforms can also be used to promote educational equity, for example, by giving access to students in rural areas or by using technology to enable access to educational resources, such as text-to-speech functions (Edyburn, 2021; Parkes et al., 2015).

With the increasing demand for web-based learning, the question arises about how learners use such a platform. Especially with the easy-to-implement individualization options, learners are offered millions of different learning paths, even in a small-scale learning platform. Learning analytics can provide insights into usage behavior, allowing researchers to understand and individualize learning better (Ifenthaler & Drachsler, 2020). We contribute to the literature by analyzing the usage patterns of learners using UDL-guided elements in a web-based learning platform called "I$_3$Learn." The Center for Applied Special Technology (CAST) has developed the UDL framework and the associated guidelines and summarized them in a graphic organizer (CAST, 2024). The organizer visualizes the three principles, nine guidelines, and 31 checkpoints[1] (CAST, 2018). The following section thoroughly examines the principles, guidelines, and checkpoints. The analysis includes the time-based use of texts and videos (UDL checkpoint 1 "provide options for perception"), learning tasks with and without help (UDL checkpoint 8.2 "vary requirements and resources to optimize the challenge"), and self-assessment (UDL checkpoint 9.3 "develop self-assessment and reflection") in the learning platform I$_3$Learn (CAST, 2018). For this purpose, we have chosen an exploratory approach: First, factors underlying the usage pattern of UDL-guided elements are identified using exploratory factor analysis (EFA) (Beavers et al., 2013). Based on the factor loadings, we engineered features that use the unsupervised machine learning (ML) algorithm k-means to find groups based on the similarity of usage patterns (Hartigan & Wong, 1979). Finally, we compare the exploratively determined groups based on their similarities in terms of usage patterns and learner characteristics.

2. Theoretical background and literature review

2.1. The Universal design for learning

Universal Design for Learning (UDL) is a comprehensive framework that aims to shape the design of learning environments to be accessible and challenging for all learners. The overarching goal of UDL is to cultivate autonomous learners who are experts in taking responsibility for the learning process: They plan and design their learning process to get an optimal experience corresponding to their particular needs. The UDL framework focuses on changing the learning environment rather than changing the learners and emphasizes intentional design to minimize barriers and provide meaningful, challenging learning experiences for all students (CAST, 2018; Fisseler, 2023; Navaitienė & Stasiūnaitienė, 2021). Using UDL can support not only learners with disabilities or impairments but all learners in the classroom. Also, the degree of impairment is not a fixed system of categories but a spectrum (Edyburn, 2021; Florian & Spratt, 2013). The concept of UDL is widely used and has even found its way directly into national curricula in some countries, shaping learning environments in schools and universities (Edyburn, 2010; Capp, 2017; Fisseler, 2023). The UDL Guidelines serve as a practical tool that aligns with the UDL framework and provides support for educators and researchers to put UDL principles into practice. These guidelines are structured according to three principles:

Provide multiple means of engagement: Learners differ in their affect and, thus, their motivational readiness to participate in the learning process. In order to stimulate this for each learner, several guidelines refer to this principle addresses the learner's emotions affecting motivation and engagement in the learning process.

Provide multiple means of representation: Learners process information from sources differently and benefit from a wide range of sources and individualization options. Each learner can choose the type of information reception that suits her best.

Provide multiple means of action and expression: Diverse learners also prefer various options for action and expression, partly because physical limitations restrict the scope of expression. By offering several options, all learners can interact with their learning

[1] The graphic organizer can be found here: https://udlguidelines.cast.org.

environment during the learning process, regardless of preferences or limitations.

The three principles have resulted in guidelines with concrete checkpoints to enable inclusive learning. The suggestions are not obligatory but can be selected optionally (CAST, 2018). Concrete suggestions in the checkpoints include, for example, providing audio and visual information, a customizable interface, or a glossary explaining essential terms (CAST, 2018). Hollingshead et al. (2020) point out that experts' perceptions and interpretations of the

guidelines vary due to the complexity of UDL. This also becomes clear when examining how the UDL Guidelines have been implemented in other research projects. The systematic review of Bray et al. (2023) shows the technology-based integration of UDL. Here, greater attention is paid to implementing the "Provide multiple means of representation" principle by offering audiovisual alternatives, interactive learning options, text-to-speech, or translation functions. The "Provide multiple means of engagement" principle is realized by giving learners alternative ways to respond to learning tasks (text, audio, or drawing) through a dashboard or scaffolding options. The principle "Provide multiple means of action and expression" is implemented in the reported studies via multimedia presentation options or further scaffolding options so that learners can report on their success. In the study by Rao et al. (2015), learners responded positively to the provision of multiple representations of their learning materials in online learning. Besides texts and audio, this also includes interactive websites. In a digital learning environment by Roski et al. (2021), the principle of multiple representations was also utilized; in addition to a video, an interactive pop-up text and a comic were also used. For an online course, Rogers-Shaw et al. (2018) also provided multiple representations, but additionally alternative assessment options and fine-grained communication via email and forum to provide multiple means of engagement.

The uncertainty and lack of clarity about the implementation of UDL yields the problem that the structures entailed by UDL require refinement in order to be applied in a future-proof fashion: "More work is needed to identify what, when, and how the UDL framework benefits educational policy and practice" (Hollingshead et al., 2020, p. 23). Learning analytics can contribute to understanding students' usages of UDL and to capturing practical implementations of UDL. This paper demonstrates one possible approach.

2.2. *Analysis of usage patterns in learning platforms*

Learning analytics seeks to gain better insight into learning by, among other things, applying machine learning-based approaches and thus analyzing and supporting both learning and teaching processes in greater depth (Ifenthaler & Drachsler, 2020; Long & Siemens, 2011): "Learning analytics is the measurement, collection, analysis, and reporting of data about learners and their context, for understanding and optimizing learning and the environments in which it occurs" (Long & Siemens, 2011, 34).

Schmitz et al. (2017) describe in their model the possibility of using learning analytics to observe the effect of the learning design on learning outcomes and to optimize it based on the insights gained. In addition, Vytasek et al. (2019) describe the possibilities of mining learning process data and applying learning analytics to understand student engagement during learning. Phillips et al. (2012) and Adji and Hamda (2019) use learning analytics to understand students' behavior and participation in online learning environments, revealing interactions that can only be tracked with learning analytics approaches. The digital tracking of interaction data and the use of learning analytics make it possible to look over all learners' shoulders simultaneously and record and analyze their learning behavior. These examples provide a brief insight into the possibilities of learning analytics and suggest that a transfer to implementing the UDL Guidelines and an in-depth analysis to promote understanding of these guidelines is feasible.

In learning analytics, researchers use machine learning to analyze data on learning behavior. (Furini et al., 2022). These include, for example, observing user behavior during learning with videos (Merkt et al., 2022), predicting dropout (Roski et al., 2023), modeling students current knowledge state (Pelánek, 2017), or grouping learners who show similar performance (Khosravi & Cooper, 2017).

Machine learning can generally be distinguished between supervised, unsupervised, and reinforcement learn-ing (Marsland, 2015). Using an unsupervised machine learning algorithm to cluster learner usage behavior is a more explorative approach than supervised machine learning. However, it has the advantage of not requiring labels of the ground truth of the data. Due to the scale of the data potentially generated during the usage of a learning platform, coding and thus an assignment of labels, as it is widely used in science education research (Cheung & Tai, 2021), cannot usually be realized. The data is often complex to interpret for humans, and its size exceeds the human capacity to find patterns.

Several approaches exist to utilize machine learning to enhance the comprehension of how learners interact on digital or web-based learning platforms. Hung and Zhang (2008) analyze the learning in a project-based business course in higher education and predict the outcomes. The unsupervised machine learning algorithm k-means was used to group learners based on their behaviors, such as time spent and learning performance. Dooley and Makasis (2020) draw on log data from a flipped classroom setting. Here, learner behavior is measured by the number of learning sessions, the number of topics covered per session, and whether learners refer to the online material before the classroom course. The k-means algorithm was also applied. Estacio and Raga (2017) use the widely known Moodle learning management

system to create various higher education courses. Log data about the activities was collected, but the timestamp of each activity, such as the use of the forum or a quiz, was not considered. The usage behavior of the different activities was visualized with the help of a vector space model. Jo et al. (2014) chose a different approach. Here, the focus is on the so-called creation of proxy variables. These variables are an abstract construct based on various time-based information from log files, such as the total login time or the login frequency. The data comes from a digital learning platform. The proxy variables' predictive power in learning performance was investigated using multiple linear regression analysis. To the best of our knowledge, previous analyses have some limitations.

2.3. The web-based learning plattform I_3Learn

We developed the web-based learning platform [I_3Learn using the content management system WordPress (Roski et al., 2024). The learning platform focuses on "ion bonding" and is adapted for 9th and 10th-grade inclusive chemistry teaching in Loxer Saxony, Germany. The content is based on the curriculum for integrated comprehensive school. This type of school has a particular focus on inclusive learning and teaching. Due to the SARS-CoV-2 pandemic, the learning platform was designed for remote learning and is thus suitable for independent learning from home. To provide access to the learning resource for all learners, the development of I_3Learn was guided by the UDL framework.

The learning platform is designed for a total learning time of 180 min. On a macro level, individual contents are arranged as chapters. The chapters can be divided into three categories: (1) chapters for repetition of previous and prerequisite content (in the context of ion bonding: the octet rule, the distinction of atoms and ions, and electron transfer), (2) chapters to achieve the minimum learning goals (phenomenon brittleness, ion bonding, ratio formula of ion bonding), and (3) chapters to achieve the extended learning goals for higher-performing learners (extended ratio formulas, nomenclature, salts in everyday life). Some chapters can be worked on in custom order and/or are optional for learners, such as repetition chapters. In addition, learners can follow individual learning paths within each chapter, which we call the micro level. Learners can choose between text and videos as a source of information or work on tasks with or without help. Additional UDL Guidelines and checkpoints implemented in I_3Learn are a thematic focus to an everyday phenomenon (brittleness of salt crystals), advanced organizer, self-assessment (learners rate from one to five stars how well they coped with a text/video/task), customizable interface (font, font size, high-contrast background, grayscale), read-aloud function, glossary, simple language, and device-independent use of the web-based learning platform (CAST, 2018). Table 1 provides an overview of the UDL-guided elements integrated into I_3Learn and the corresponding UDL Guidelines and checkpoints.

Accordingly, we generally follow a broad understanding of the concept of inclusion in our web-based learning platform. The goal of our project is to value diversity, and we aim to remove barriers to learning and to enable participation for all learners (Stinken-Rösner et al., 2020; Göransson & Nilholm, 2014; Ainscow et al., 2006; Roski et al., 2024), also in line with the "Education for All" approach of the United Nations (UNESCO, 2005; UNESCO-UNEVOC, 2021). We are not only interested in integrating vulnerable groups, such as students with a disability, into science learning according to a placement definition of inclusion as discussed by Göransson and Nilholm (2014) but also in creating broad access to the science curriculum for the benefit of all learners, including those with special needs, as highlighted by Roski et al. (2024). For this reason, the web-based learning platform was designed for use in high schools and integrated comprehensive school. This type of school stands out because it is inclusive and integrates learners from vulnerable groups into regular school classes.

3. Research purpose and questions

Even though there are already numerous approaches to finding the usage behavior of learners in a web-based learning platform, this paper highlights the usage of UDL-based elements and, thus, different learning paths in a web-based learning platform with a focus on K12 chemistry education. This bridges the gap in previous research regarding the importance of understanding the use of UDL-guided elements (Roski et al., 2021).

RQ1. How can the complexity of data on the usage behavior of UDL-guided elements be reduced purposefully?

RQ2. To what extent can learners be clustered into groups of similar behavior based on UDL-guided elements in a web-based learning platform?

Table 1
Overview of all UDL-guided elements in the learning platform I_3Learn. Bold elements were used in this study.

UDL-guided element	Description of implementation	UDL Guideline/checkpoint (CAST, 2018)
Multiple representations	**Text and videos are available for learning con-tent**	**1 Provide options for per- ception**
Customizable interface	Customizable in terms of font size, font type, and contrasts	1.1 Offer ways of customiz-ing the display of informa-tion
Read-aloud function	Texts can be read aloud; Reading speed is adjustable	1.2 Offer alternatives for au-ditory information
Glossary	All relevant terms relating to ion binding are explained again separately in the glossary	2.1 Clarify vocabulary and symbols
Simple language	All texts and audio information in the videos are written in simple language	2.1 Clarify vocabulary and symbols
Device- independent	The web-based learning platform can be accessed from any device with Internet connectivity	4 Provide options for physi-cal action
Chapter selection	The order of the learning chapters can be chosen freely; for the most part	7.1 Optimize individual choice and autonomy
Everyday phenomenon	A reference to everyday life was established with the help of an everyday phenomenon	7.2 Optimize relevance, value, and authenticity
Advanced organizer	Activation of prior knowledge and preparation of learners for new learning content	8.1 Heighten salience of goals and objectives
Tasks with or without help	Learners can choose scaffolding	8.2 Vary demands and re-sources to optimize chal-lenge
Self-assessment	Learners have the chance to reflect on the use of texts, videos, and tasks	9.3 Develop self-assessment and reflection

4. Method

The methodological approach can be summarized in four steps: (1) data collection, (2) data preparation, (3) exploratory factor analysis (EFA), and (4) k-means clustering (see Fig. 1).

4.1. Data collection

Data collection was conducted using the learning platform I_3Learn.

A total of 580 learners from 27 classes have logged in to the learning platform at least once during the first half of 2022. Almost 500,000 log files were generated, which form the basis for further analyses. The log files represent the time-accurate behavior of each learner on the I_3Learn learning platform. Further data about the learners is being collected through pre- and post-tests, supplementing the usage behavior documented in the log files. The adapted Bonding Representations Inventory (BRI) of Luxford and Bretz (2014) assesses the conceptual knowledge about ions and their bonding used in the pre and post-test. The original version of the BRI was first translated, suitable items for ion bonding were selected, and new items tailored to the learning content of I_3Learnwere added. The reliability was tested in a separate study with 407 learners. Cronbach's alpha is 0.87 (Meyer, 2021).

In addition, chemical self-concept and interest (Frey et al., 2009), socio-economic background (Torsheim et al., 2016), reading ability (Mayringer & Wimmer, 2014), cognitive ability (figure analogy; Heller and Perleth (2000), language spoken at home (as an indicator for migration background), gender, school type, and grade were collected. The Ministry of Science and Culture, Lower Saxony, Germany approved the entire data collection.

4.2. Data preparation

In order to address the research questions, the duration of usage of the UDL-guided elements in seconds is extracted from the log files. It involves using texts and/or videos, tasks with or without assistance, and self-assessments (CAST, 2018). Even though more UDL-guided elements were integrated as part of the development of I_3Learn (see Table 1), only these three are used in the subsequent analysis. This is because these UDL-guided elements are measurable and tangible learning opportunities that require a decision-making process on the part of the learner. Other UDL-based learning opportunities, such as the use of simple language, determine the design of the learning platform and do not require learners to make any decisions. The use of these and the remaining UDL-guided elements cannot be differentiated in the log files. From the entire data set, 23,072 log files were extracted.

Three hundred eighty-four learners used one of the UDL-guided elements at least once. Non-available data points are replaced by zero, as this corresponds to the actual usage time of the UDL-guided elements. Table 2 shows an overview of the participant group's characteristics.

The platform-generated log file data document the learners' click behavior, allowing us to derive the usage duration for each UDL-guided element. Each click is logged with a timestamp; the duration of use is the time interval between two learner actions. However, possible interruptions in the learning process cannot be traced. For this reason, a cut-off value of 30 min was defined for using videos and texts. Usage periods of more than this time were filtered. The cut-off value corresponds to 6 to 13.5 times the video duration. Furthermore, a minimum usage time of 30 s was selected for videos and texts. A minimum viewing time between 10% and 22% of the videos aligns with this criterion. We assume that a lower usage time is insufficient to take in the information. For texts, a similar usage time as for videos was assumed. The script for the learning videos is based on the texts to ensure the same information content and to keep texts and videos comparable. The spoken text in the video was recorded slowly and corresponded to a leisurely reading pace.

An upper cut-off value of 5 min was chosen for self-assessment and tasks. This corresponds to sufficient time to determine the ratio formula tasks, which can be considered the most time-consuming in the learning platform. The estimation is based on the judgment of experts with at least a Master of Education in chemistry education. A value above this threshold is replaced with the mean value of the usage time. There is no minimal threshold defined for self-assessments and tasks. Both can be completed in a few seconds, for instance, if a ratio formula is known. Table 3 shows all features extracted from the dataset. Due to the architecture and the peculiarities of the tasks in I_3Learn, there are some particularities in the features that require clarification. Notably, the "Ratio Formula 1″ task is not included in the "tasks" feature category because it does not allow learners to complete it without help; it is exclusively available with help in I_3Learn to familiarize the learner with this topic. Given the absence of a choice, we did not categorize it as a UDL-guided element. However, we have classified the associated self-assessment as such. A similar situation arises with other tasks, such as the "ion grid" task, where learners arrange cations and anions. In this case, we have also classified the associated self-assessment as a UDL-guided element.

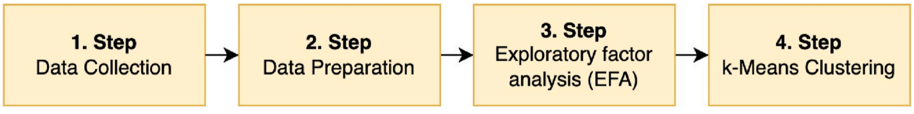


Fig. 1. Methodological steps for clustering learner behavior.

Table 2
Sample characteristics of the data collected with the web-based learning platform I_3Learn.

Sample characteristic		N	Percentage
Gender	Male	148	41.5%
	Female	183	52.7%
	Diverse [a]	26	7.3%
School	Integrated comprehensive school	102	27.4%
	Cooperative comprehensive school	54	14.5%
	High School	209	56.2%
	Other	9	1.9%
Grade	9th	233	62.6%
	10th	139	37.4%
Language at home	German	333	89.5%
	Other language	39	10.5%

[a] In Germany, individuals who do not identify as either male or female have this as a legal option.

Table 3
Overview of the features included in the exploratory factor analysis.

UDL-guided element	Feature	Feature details
Multiple representations	Text	Text about ions
		Text about noble gas configuration
		Text about electron transfer
		Text about ion bonding
		Text about ratio formula
	Video	Video about ions
		Video about noble gas configuration
		Video about electron transfer
		Video about ion bonding
		Video about ratio formula
Tasks	With help	What is sodium chloride?
		Ratio formula 2
		Ratio formula 3
		Advanced ratio formula 1
		Advanced ratio formula 2
		Advanced ratio formula 3
	Without help	What is sodium chloride?
		Ratio formula 2
		Ratio formula 3
		Advanced ratio formula 1
		Advanced ratio formula 2
		Advanced ratio formula 3
Self-assessment		*For each feature mentioned above*
		Task ions
		Task noble gas configuration
		Task electron transfer
		Task ion bonding
		Task ion grid
		Task ion grid characteristics
		Task nomenclature
		Task with help ratio formula 1

4.3. Exploratory factor analysis

We chose exploratory factor analysis (EFA) to study the underlying structure of the student usage data, reduce the complexity for the following cluster analysis, and enable more effective and interpretable clustering. The underlying structure is linked to observed features through factors and factor loadings. A factor is a latent variable representing the correlations between different measured variables, such as the usage time in seconds shown in this work. The factor loading describes how much a measured variable correlates with the factor. The Bartletts sphericity test checks whether the data correlate entirely (Kaiser & Rice, 1974; Klopp, 2010; Weygandt, 2021). The result should be statistically significant. Furthermore, the Kaiser-Mayer-Olkin coefficient (KMO value) is determined, which checks for the partial correlation of the variables/features entered. The KMO value can be between 0 and 1, whereby a value above 0.6 is considered acceptable (Bartlett, 1937; Klopp, 2010; Weygandt, 2021). If the data is suitable for an EFA, it is calculated without a rotation to obtain the eigenvalue of each variable/feature. Using a Scree plot, where the factors are plotted on the x-axis and the eigenvalues on the y-axis, the number of factors can be determined by the resulting “elbow” in the graph. After the subsequent selection of the factors, the rotation method “Varimax” is chosen to interpret the factor loadings better. A cut-off value of 0.4 or higher for incorporated factor loadings is selected, ensuring that only items with a high relation to a factor are included. The choice of this factor loading leads to interpretable factor loadings, more representative factors, and a reduction of noise (Klopp, 2010; Werner, 2014; Weygandt, 2021). The EFA was calculated with the Python module *factor_analyzer* (Biggs & Madnani, 2022).

4.4. K-means clustering

To explore similar usage behaviors of UDL-guided elements, we use the widely-used k-means clustering approach (MacQueen, 1967). K-means is an unsupervised machine-learning algorithm that assigns data points to cluster centers (centroids) based on similarity. The number of centroids, determined by k, must be set in advance. During initiation, k data points are randomly selected as centroids, and the other data points are assigned to these centers. The assignment revolves around selecting a specific distance measure, with the Euclidean distance relative to the cluster center chosen for this purpose. Afterward, the centroids are updated by calculating the mean value of the respective clusters; the mean value serves as a new centroid, to which the data points are assigned again. This is repeated until the assignment of the data points does not change anymore (Jain et al., 1999; Jin & Han, 2011; MacQueen, 1967; Manning et al., 2008). To determine the number of clusters, we consider the silhouette score, the Bayesian Information Criterion (BIC), and the Akaike Information Criterion (AIC) as metrics (Everitt et al., 2011; Shahapure & Nicholas, 2020). The silhouette score is a metric in the range of −1 to 1. The higher the value, the closer together the data points of a cluster are. A reasonable structure with a silhouette score above 0.5 (Kaufman & Rousseeuw, 1990) can be assumed. The BIC or AIC is suitable for comparing different models regarding relative information loss. Consequently, in the comparison, the lowest value for both information criteria indicates the model to be selected. The BIC penalizes complexity to a greater extent (for more detail, see Everitt et al. (2011)).

We chose k-means as the clustering algorithm because it is a simple, efficient, robust, and interpretable approach. There is no overlap for the assignment to a cluster, and the individual clusters can be described easily, which supports answering the research questions (Wu, 2012). Furthermore, we resort to k-means clustering because it is an established and widely used research method in educational research as well as in learning analytics to cluster learner data (Cerezo et al., 2016; Liao & Wu, 2023; Lust et al., 2011). The features for clustering were formed from the four factors of the

EFA. Based on each factor, a weighted sum score was calculated as a feature for k-means (Distefano et al., 2009). The weighted sum score S is calculated as $S = \sum_i w_i \times v_i$, where w_i is the factor loading and v_i is usage time in seconds

of the ith UDL-guided element within one of the four factors. As a result, four features for k-means have been created

for each learner, analogous to the four factors. Since the assignment to the centroids is based on distance, the features for k-means clustering are normalized separately for each feature (Mohamad & Usman, 2013). We used the Python library *scikit-learn* for data scaling and k-means clustering (normalization: *sklearn.preprocessing.StandardScaler*; k-means clustering: *sklearn.cluster.KMeans*; Pedregosa et al. (2011)). The default parameters were selected for k-means clustering.

5. Results

5.1. Exploratory factor analysis

The results of both the Bartlett test (chi-square = 5389.13, $p < 0.001$) and the Kaiser-Mayer-Olkin coefficient (KMO value = 0.677) indicate the general feasibility of EFA with the data set (Bartlett, 1937; Kaiser & Rice, 1974; Klopp, 2010). The elbow argues for a five-factor solution (see Fig. 2).

However, the factor loading results showed that the individual factor loadings are poorly interpretable. Only a few factor loadings are above 0.4, so individual factors are partly based on a few variables/features. In addition, individual variables/features load on multiple factors. Consequently, we reduce to a four-factor solution promoting the interpretability of the individual factors. More variables/features can be assigned to the individual factors, while the variables/features only load on one of the four factors (see Table 4).

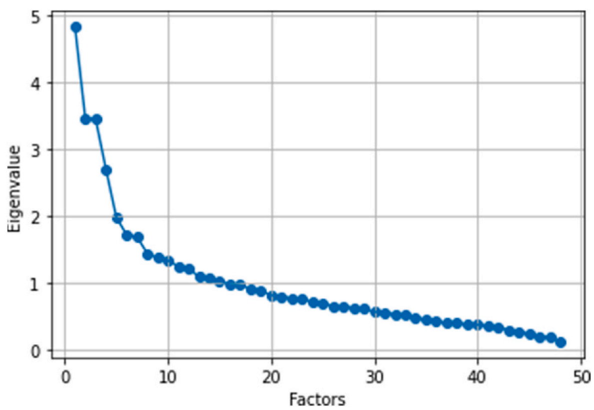


Fig. 2. Scree plot of the exploratory factor analysis.

Table 4
Overview of factor loadings due to the exploratory factor analysis.

Variable/Feature	1	Factor 2	3	4
Text about ratio formula	0.43			
Task with help - Advanced ratio formula 1	0.53			
Task with help - Advanced ratio formula 2	0.52			
Task with help - Advanced ratio formula 3	0.48			
Task with help - Ratio formula 2	0.59			
Task with help - Ratio formula 3	0.68			
Self-assessment - Ratio formula 2 with help	0.55			
Self-assessment - Text ratio formula	0.44			
Video about noble gas configuration		0.50		
Video about electron transfer		0.60		
Self-assessment - Text ions		0.57		
Self-assessment - Video ratio formula		0.44		
Self-assessment - Video electron transfer		0.76		
Self-assessment - Video noble gas configuration		0.64		
Task without help - Advanced ratio formula 2			0.57	
Task without help - Ratio formula 2			0.53	
Task without help - Ratio formula 3			0.56	
Self-assessment - Ratio formula 2			0.43	
Self-assessment - Advanced ratio formula 2			0.48	
Self-assessment - Task nomenclature			0.44	
Text about ions				0.46
Text about noble gas configuration				0.67
Text about electron transfer				0.57
Self-assessment - Task electron transfer				0.43
Self-assessment - Task noble gas configuration				0.44
Self-assessment - Text electron transfer				0.52
Self-assessment - Text noble gas configuration				0.52
Self-assessment - Video ions				0.45

Furthermore, an interpretation of the four factors is also possible regarding content. Factor 1 focuses on text on ratio formulas (ratio of ions in an ion grid), related tasks that rely exclusively on help, and two matching self-assessments. Factor 2 focuses almost exclusively on the use of videos and self-assessments on videos. The self-assessment regarding the use of the "ions" text stands out. Only text or video on one main topic can be found in this factor. The third factor also covers ratio formulas but focuses on unassisted tasks. More advanced learners can work on it since it is a voluntary chapter at the end of the learning platform. In addition, the self-assessment on the nomenclature of salts, which is based on the architecture of I_3Learn, is a chapter that also assumes that more advanced learners tend to work on unassisted tasks. In the case of Factor 4, the content can be assumed to be the opposite of the second factor. Here, the focus is on using texts and the appropriate self-assessments. Also, in line with the second factor, the self-assessment on using the content "ions" catches the eye here, but this time the video instead of the text.

5.2. K-means clustering

When comparing the silhouette score and BIC/AIC at different numbers of clusters, the silhouette score shows a maximum at $k = 6$, and the BIC/AIC shows a minimum at the same number of clusters (see Table 5). Thus, both metrics indicate the use of six clusters. The silhouette score for the six clusters is 0.506. Furthermore, the clusters identified can be defined in terms of the usage behavior of UDL-guided elements and the differences of learners within the clusters. The default settings of the *scikit-learn* software library were used for k-means clustering, involving ten initializations and a maximum of 300 iterations. The best initialization is automatically selected based on the lowest inertia (Pedregosa et al., 2011).

Table 5
Silhouette Score, BIC, and AIC depending on the number of clusters. The metric with the best result was marked.

Number of clusters	Silhouette Score	BIC	AIC
2	0.445	765.73	651.16
3	0.466	−481.47	−655.29
4	0.478	−442.93	−676.02
5	0.501	−454.12	−746.46
6	**0.506**	**−3454.13**	**−3805.73**
7	0.438	−2066.45	−2477.31
8	0.441	−1766.34	−2236.47
9	0.461	−1683.51	−2212.90

6. Interpretation and discussion of the results

In this study, we used two experimental methods to gain insight into the usage patterns of learners using UDL-guided elements. First, the results of the EFA are interpreted and discussed with regard to the first research question. This is followed by the description of the six clusters identified by k-means clustering. This serves to answer the second research question, which seeks to identify learner groups of similar usage patterns in I_3Learn.

Within each cluster, learners with similar usage patterns are grouped. The clusters can subsequently be described concerning the learner characteristics gathered during the pre- and post-test.

6.1. Interpretations of the factors

The four factors show a certain degree of consistency in content and reflect correlations in learners' usage behavior among the UDL-guided elements, which are aggregated into a joint factor. Factors 1 and 3 mainly involve tasks. What is striking here is the clear separation between the factors. Factor 1 mainly contains tasks with help, Factor 3 mainly without help. This implies that learners remain consistent in decision-making behavior throughout using I_3Learn: They tend to choose the tasks with or without help, regardless of where the tasks are located on their learning journey. The situation is similar for Factors 2 and 4. Factor 2 mainly indicates the use of videos as a source of information. Factor 4 mainly refers to text (see Table 4). Here, too, there are indications that learners make a decision on this option and do not change it later. This situation persists despite the UDL-based content being dispersed across multiple locations within the learning platform. Furthermore, understanding why learners remain engaged with UDL-guided elements when deciding their learning paths warrants thorough exploration. This also paves the way for an ML-based recommendation system, offering learners personalized suggestions to select the most suitable learning path.

6.2. Cluster differences

An in-depth analysis of the clusters shows differences in the characteristics of the learners, such as the chemistry self-concept, and their learning patterns, such as the use of videos. At first glance, the clusters differ in the number of learners assigned to them (see Table 6). The largest cluster is Cluster 1, with more than every second learner. The other clusters are smaller in comparison. Cluster 2, in particular, is small, with only seven learners assigned.

Fig. 3 shows the average usage time of the UDL-guided elements for each of the six clusters and reveals considerable differences between the clusters. Despite the limited size of Cluster 2, which is comprised of only seven learners, notable disparities in usage duration are evident, particularly when compared to Cluster 1. Obviously, learners in Cluster 2 invest particularly more time in UDL-guided elements.

The usage patterns depicted in Fig. 3 are subsequently translated into a z-standardized representation for each UDL-guided element (Fig. 4) to clearly show the differences between the clusters. The associated error bars denote the standard deviation. The visualized results show substantial differences in the usage behavior of the UDL-guided elements in I_3Learn. For instance, Cluster 2 is distict from the other clusters by the extended duration of text usage (Fig. 4a), while Clusters 0 and 3 also exhibit prolonged durations of text usage. Additionally, Cluster 2 is prominent for its extended video usage duration (Fig. 4b), and Cluster 3 similarly stands out for an extended usage of tasks with assistance (Fig. 4d). Overall, the clusters generated by k-means seem to represent meaningful groups of learners with distinct characteristics. Thus, we conclude that it makes sense to investigate the cluster differences further.

In addition to variations of the usage behavior regarding the distinct UDL-guided elements, learners can be categorized according to their demographics and the scores obtained from different test instruments. Fig. 5 shows the distribution among the six clusters regarding gender, the language used at home, and grade level. Fig. 5b shows that the language used at home, an indicator of the migrant background of the learners, is at a similar level in all six clusters. This cannot be interpreted as a comprehensive bias analysis, but it seems that migration background is not a factor that dominates or correlates with a particular cluster.

Fig. 6 offers a more detailed perspective on the cluster differences: It visualizes the learners' data collected through the chemical knowledge test in the pre- and post-test. The values from Fig. 6 have been z-standardized and the accompanying error bars represent the standard deviation. Further differences between the clusters can be identified from the learning characteristics. It is evident, for example, that Cluster 2 and 5 have apparent differences in their self-concept and interest in chemistry (Fig. 6a and b). In addition, Cluster 0, for example, stands out from the others regarding reading ability (Fig. 6d), and Cluster 1 shows below-average post-test scores (Fig. 6h).

Table 6
Distribution of learners per cluster.

Cluster	Number of learners	Relative number of learners
0	21	5.5%
1	226	58.9%
2	7	1.8%
3	68	17.7%
4	43	11.2%
5	19	5.0%

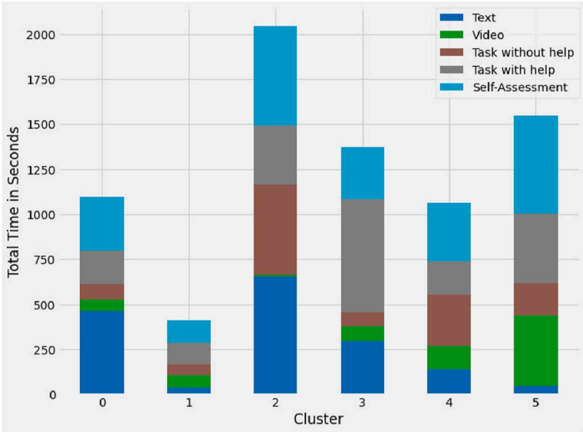


Fig. 3. Mean usage time in seconds by learners in each cluster concerning the UDL-based elements.

Figs. 3–6 each provide a new perspective on the clusters and, above all, on the differences between them. In the following, the information from these figures is used to characterize the six clusters with respect to all observed aspects.

6.2.1. Cluster 0 - engaged (primarily female) reading learners

Cluster 0 learners show the characteristic of being mainly female 9th-grade learners. Compared to other clusters, they show a higher mean value in the reading score with a comparably wider standard deviation. For socio-economic background and knowledge gain, high values are evident in contrast to other clusters. Regarding usage patterns, the time spent reading the texts is particularly striking.

6.2.2. Cluster 1 - learners investing minimally in their learning

Cluster 1 learners are notable for constituting the largest cohort. Not surprisingly, the z-standardized characteristics are usually around 0. Nonetheless, the low variability of the characteristics stands apart, as indicated by the small error bars. The usage patterns show that the learners have invested the minimum learning time with the interaction of I_3Learn. Compared to the other clusters, the learners of Cluster 1 tend to use less text. They show almost no gain in conceptual knowledge.

6.2.3. Cluster 2 - diverse learners investing a lot and gaining a lot

Cluster 2 is notable for its number of female learners, mirroring the pattern observed in Cluster 0. The characteristics of this small group possess great variability in almost all facets. Nevertheless, self-concept and interest exhibit a noticeable difference in mean values compared to other clusters, independent of cognitive and reading ability. The socio-economic background is lower than for other learners. Learners show the greatest knowledge gain. In terms of usage patterns, the extensive use of the learning platform in terms of time is particularly noteworthy. Learners prefer texts to videos and tasks without help compared to other clusters. In addition, learners make extensive use of the self-assessment option.

6.2.4. Cluster 3 - learners with balanced characteristics using text-based learning and task assistance

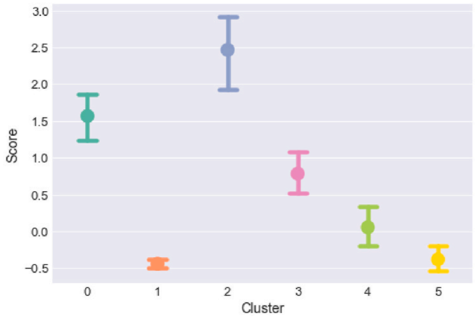
Cluster 3 stands out little in terms of learner characteristics. The mean value for the z-standardized generally ranges around 0 or deviates only slightly. Regarding the usage patterns in I_3Learn, there are indications of increased use of text. Furthermore, learners in cluster 3 are distinctive for dedicating much time to tasks with assistance when directly compared to the other clusters.

6.2.5. Cluster 4 - learners with (comparatively) strong self-concept, interest, and learning autonomy taking the challenge without using help

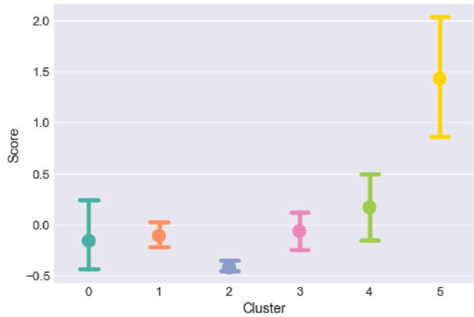
Cluster 4 can be described in terms of learner characteristics as a cluster with tendencies toward larger mean z-scores in terms of self-concept, interest, cognitive ability, and socio-economic background. These may not hold an exceptionally central position, but they consistently maintain the second-highest values compared to the other clusters. The differences become more apparent when individual clusters are distinguished separately. Compared to Cluster 3, learners in Cluster 4 show a greater level of interest or a better post-test score than Cluster 1. Concerning usage patterns, it is evident that more time is spent on tasks without additional assistance and that the offer of self-assessment is also used to a greater extent than in the largest cluster, at least in terms of time. Spending more time on tasks without additional help does not initially indicate whether they have been mastered, but it does show that the learners in this cluster accept the challenges. They spend more time than other learners on these tasks, even though they always have the option to switch to the task with help. The z-score for tasks with help is significantly lower than the other clusters.

6.2.6. Cluster 5 - diverse learners prefer videos and reflect on their learning process

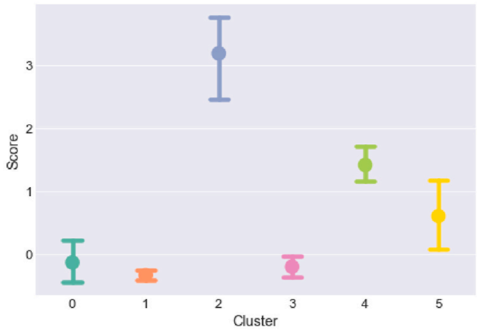
Cluster 5, like Cluster 2, constitutes a smaller cluster comprising merely five learners. Amidst the substantial variance in learner characteristics, a conspicuous pattern emerges, indicating lower mean values for self-concept and interest in chemistry. The usage



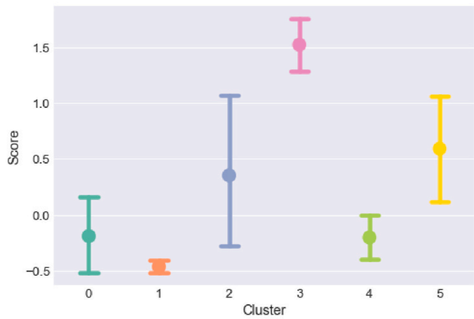
(a) Cluster differences within text usage.



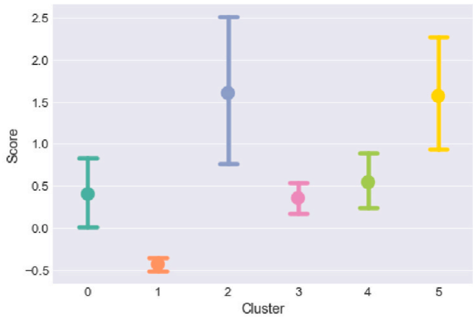
(b) Cluster differences within video usage.



(c) Cluster differences within tasks without help usage.



(d) Cluster differences within tasks with help usage.



(e) Cluster differences within self-assessment usage.

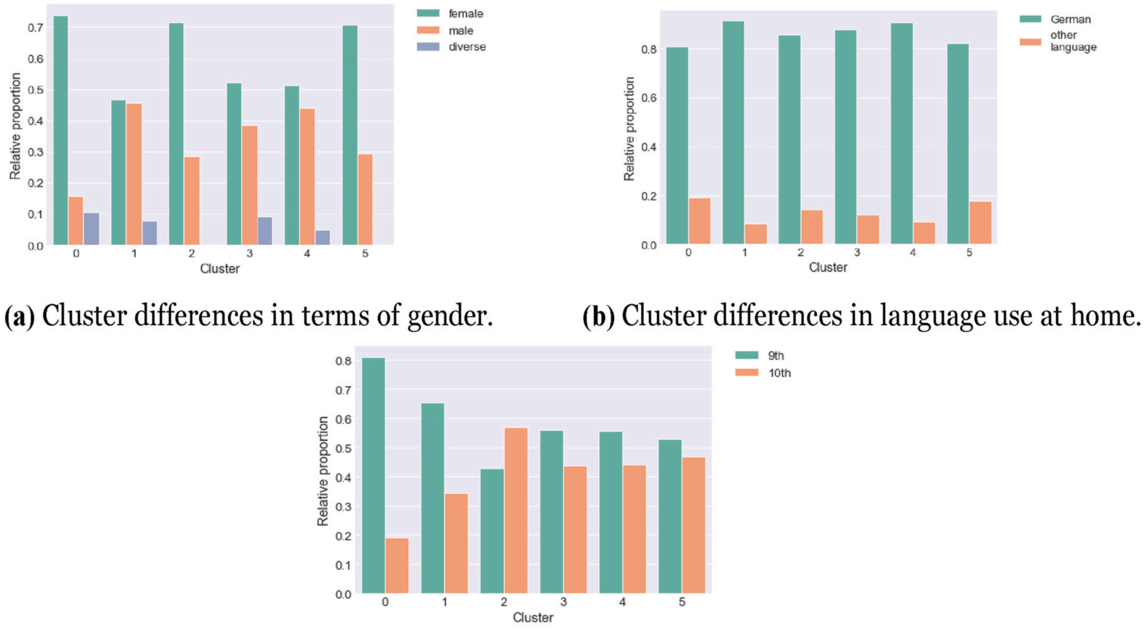
Fig. 4. Comparison of clusters based of (a) text usage, (b) video usage, (c) task without help usage, (d) task with help usage, and (e) self-assessment usage. Point-plot shows the z-standardized mean value. Error bars show the standard deviation.

patterns consistently reveal a preference for videos. Texts are used to a small extent by the learners in Cluster 1, mainly characterized by the common use of the entire learning platform. Furthermore, the usage is also characterized by learners spending much time on self-assessment. Regarding utilizing information sources (text/video) and self-assessment, Cluster 2 and Cluster 5 emerge as notable cases. Both have remarkably high durations of use of information sources, with learners from Cluster 2 relying on texts and learners from Cluster 5 relying on videos. Furthermore, both clusters stand out in the extensive use of self-assessments. Only in the case of recourse to tasks with or without help is there no clear tendency for Cluster 5 in terms of the mean value, whereas Cluster 2 tends strongly toward long processing times for tasks without help.

7. Conclusions

In this research paper, we analyzed the usage patterns related to UDL-guided elements of 384 learners using the web-based platform I_3Learn for inclusive chemistry education. The first research question focuses on underlying factors in the usage patterns of videos, texts, tasks with or without help, and self-assessment in seconds. The application of EFA for this purpose calculated four factors in content consistency, which indicate that learners do not show changes in their learning tendencies. For instance, when utilizing learning videos, learners generally use them within the rest of the learning platform rather than switching to alternatives like text.

The second research question relates to using k-means clustering for identifying learner clusters with similar usage patterns. Six



(a) Cluster differences in terms of gender.

(b) Cluster differences in language use at home.

(c) Cluster differences in terms of grade.

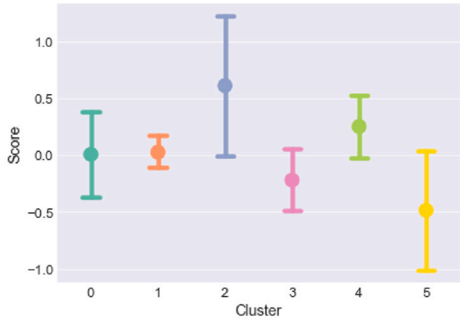
Fig. 5. The cluster differences among the six clusters in terms of (a) gender, (b) language use at home, and (c) grade.

clusters were identified and described based on the usage of I_3Learn and characteristics. The clusters show different usage behaviors of the UDL-guided elements and give a deeper insight into the learning process of the learners.

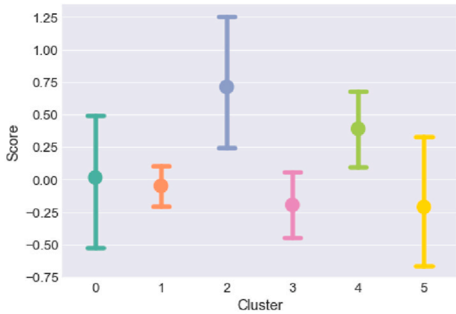
This explorative approach lays the foundation for individualized learning and support for all learners. The ability to analyze and identify specific usage patterns can be used as a starting point in future work. Machine learning-based recommender systems can react to learners' performance and behaviors. For instance, they could suggest more advanced tasks without assistance, tailoring such recommendations to support individualized learning journeys. Educators can subsequently use this information to adapt the learning and teaching process. With the help of a dashboard, educators and learners can gain insight into the learning process. Clustering gives teachers an overview of the usage of web-based learning platforms. For example, they can see immediately if one group tends to under-utilize learning opportunities. The results of the first research question imply a necessity for such a described personalized learning support: In I_3Learn, learners tend to decide on, for instance, the preferred type of learning resource (e.g., text vs. video) at the beginning of usage and then hold on to that decision for the rest of their learning time. The results do not provide insight into the learning outcomes of individual learners based on their usage patterns, nor do they imply the best learning path through the learning platform for each learner. However, they indicate that learners do not change their preferred usage pattern during an individual work phase in I_3Learn. This can be explained by the fact that learners know themselves best and know how they can learn best. Still, it can also mean for individual learners that they do not choose a learning path that is the best option for them because it is contrary to their previous behavior pattern. These learners can be identified in future work, and an alternative learning path can be suggested, or they could be informed that their learning progress is proceeding well enough that they are, for example, ready for tasks without additional assistance. However, this is also possible through dashboards for learners, the obtained clusters and their descriptions can be communicated. Learners get feedback for their learning process and can improve self-regulated learning. The context of I_3Learn and the results of our study reflect research needs described in Maier and Klotz (2022): The review findings indicate a lack of personalized feedback in digital learning environments in K-12 education, and only a few studies relied on ML-based methods (such as k-means clustering).

Learning analytics can consolidate the evidence base of the UDL framework. On the one hand, implementing and interpreting the UDL framework can be problematic. On the other hand, reports on implementations of UDL-guided elements often do not provide sufficient details. However, advancing our consideration, it should not be presumed that learners will embrace and utilize UDL-guided elements merely based on availability. The findings of this study show that all UDL-guided elements are used voluntarily among the learners. Since the described learning opportunities (e.g., text and video) are repetitive and learners tend to stick to their choices, it can be inferred that the different offerings are appreciated, and learners stick to their choices. Adding UDL-guided elements to a learning opportunity proves to be beneficial as a result. Therefore, it can be assumed that implementing several UDL-guided elements in a web-based platform diversifies the learning opportunities and supports self-regulated learning. At this point, learning analytics can also strengthen the UDL framework's evidence base by linking the active use of UDL-guided elements to their beneficial use for specific groups of learners.

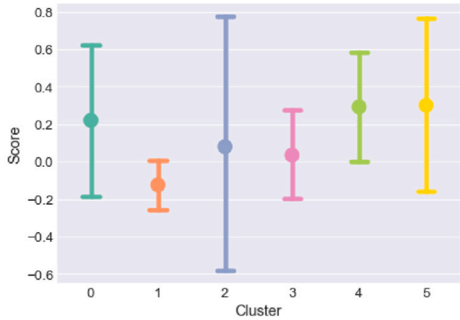
Research in the field of inclusive education showed over and over again how categorizing students as being disadvantaged (e.g.,



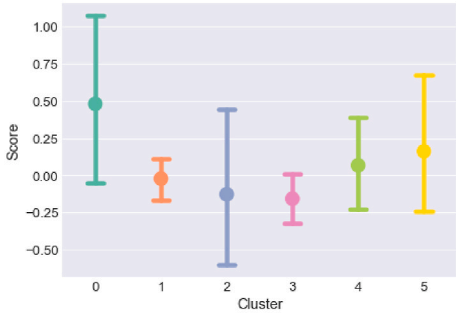
(a) Cluster differences within self-concept in chemistry.



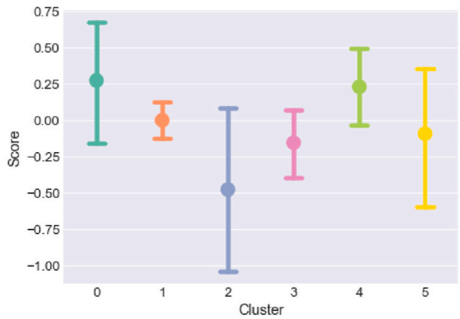
(b) Cluster differences within interest in chemistry.



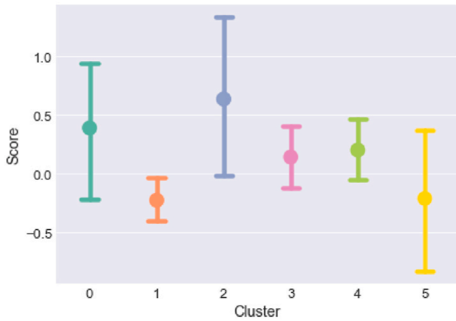
(c) Cluster differences within cognitive ability.



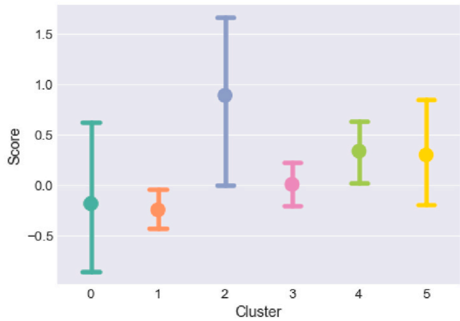
(d) Cluster differences within reading ability.



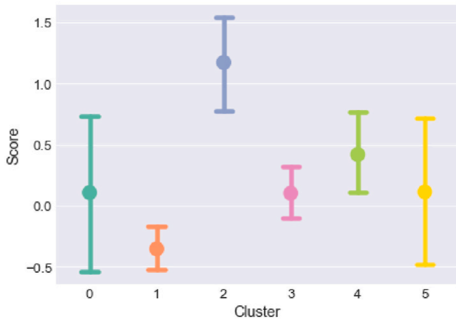
(e) Cluster differences within socio-economic background.



(f) Cluster differences within conceptual knowledge gain.



(g) Cluster differences within pre-test score.



(h) Cluster differences within post-test score.

Fig. 6. Comparison of learner characteristics between clusters in terms of (a) self-concept in chemistry, (b) interest in chemistry, (c) cognitive ability, (d) reading ability, (e) socio-economic background, (f) conceptual knowledge gain,(g) pre-test score, and (h) post-test-score. Point-plot shows the z-standardized mean value.

with or without special educational needs, gender, migration background) can entail moments of exclusion and thus be detrimental to inclusion (Florian & Black-Hawkins, 2011; Spörer et al., 2021; Szumski et al., 2017). Using learning analytics, we can transfer user behavior into human-readable clusters. This constitutes a form of categorization. However, it is only based on ways students deal with UDL-based features. This approach deconstructs typical categories used in schools, some based on sociological or legal categories, and reconstructs categories only through data on the learning processes. It makes applying the UDL framework more visible and inclusive learning

more feasible, as learners are assigned to actionable and understandable categories that are not based on discriminatory characteristics of the learners.

8. Study limitations and future directions

It should be noted that this approach is exploratory. According to the current state, it cannot be assumed that the clusters found have general validity and can be found universally. It cannot be assumed that similar clusters can be found in other studies. For this, further analyses with similar learning platforms have to be conducted. Nevertheless, the approach gives us an insight into learners' usage behavior of UDL-based learning offerings, which is still open in current research (Roski et al., 2021). In addition, further UDL-based offers are implemented in I_3Learn, which the collected log files cannot track. A conclusion that a specific usage behavior of the UDL-guided elements leads to more knowledge gain cannot be pursued due to interference. Further, it should be noted that there is evidence that more UDL-guided elements do not necessarily have better effects on knowledge gain (Roski et al., 2021). Furthermore, we only considered the learners' time use in this exploratory approach. This gives an insight into the learning process. However, additional information, such as the correct or incorrect answers to the learning tasks or the ratings in the self-assessment, could provide even better insight in the future. The sole use of time must include information on whether the learners have interacted with the content. We have defined cut-off values to protect the dataset from refined look-through behavior or long pause times. Nevertheless, it is theoretically possible that learners invest an average amount of time in a learning task but do not answer it. However, this information is available in the complete data set and could be linked to future work.

Based on the results of this work, two directions can be taken. Firstly, the results can be used to inform teachers and learners about the learning process. As in the case of Dickler et al. (2021, p. 6126), educators may receive direct indications of a learner's problems and concrete support prompts. Alternatively, like Holstein et al. (2018), educators can use a mixed-reality smart glass to display a real-time dashboard that indicates, for example, "misusing" or "struggling" learners. The effect of an integrated feedback mechanism or dashboard-like system on learning success and self-regulated learning can be tested with a follow-up study. In addition, the effect of dropping out of the learning platform can be explored (Roski et al., 2023). On the other hand, the methodological approach and other learning analytics approaches can be used to investigate and improve the effectiveness of the UDL framework. This work has identified different clusters that show different usage behaviors. UDL-guided elements are assumed to be used differently by learners, which is the purpose behind the guidelines (CAST, 2018). However, learning analytics can identify and clarify these usage clusters. In further work, effective integrations or combinations of UDL Guidelines for certain learner groups may be identified and predicted at an early stage with the help of machine learning in order to make learners aware of these learning opportunities.

In addition, while conceptualizing the web-based learning platform, another fundamental challenge became apparent: How can we operationalize the UDL framework? As can be seen in Table 1, various checkpoints were taken up in order to integrate UDL into the platform. However, it is impossible to directly translate all of the comprehensive principles of the UDL framework into educational decisions solely through the use of checkpoints. Using videos and texts to offer learners different representations is not directly expressed as a checkpoint and allows for various forms of operationalizing UDL. This makes it challenging for stakeholders to implement and interpret the UDL framework. In future work, the UDL framework should be revised to simplify its implementation by reducing interpretability and adding precisely formulated checkpoints. This can also help strengthen the UDL framework's evidence base if implementations can be accurately compared and reported. In addition, future work should track the use of the UDL-guided elements in detail, for example, using log files and evaluating them with the help of learning analytics. This is necessary to track the active use of the UDL-guided elements and draw conclusions to strengthen the evidence of individual UDL checkpoints.

CRediT authorship contribution statement

Marvin Roski: Writing – original draft, Visualization, Software, Resources, Project administration, Methodology, Investigation, Formal analysis, Conceptualization. **Ratan Sebastian:** Writing – review & editing, Validation, Methodology. **Ralph Ewerth:** Writing – review & editing, Supervision. **Anett Hoppe:** Writing – review & editing, Supervision. **Andreas Nehring:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Data availability

Data will be made available on request.

Acknowledgments

This paper results from the PhD program LernMINT, founded by the Ministry of Science and Culture, Lower Saxony, Germany.

References

Adji, S. S., & Hamda, S. (2019). Utilization of learning analytic on measuring student participation in online learning. *Journal of Physics: Conference Series 1280*. https://api.semanticscholar.org/CorpusID:213125473.

Ainscow, M., Booth, T., & Dyson, A. (2006). *Improving schools, developing inclusion*. Routledge. https://doi.org/10.4324/9780203967157

Alraimi, K. M., Zo, H., & Ciganek, A. P. (2015). Understanding the moocs continuance: The role of openness and reputation. *Computers in Education, 80*, 28–38. https://doi.org/10.1016/j.compedu.2014.08.006

Bartlett, M. S. (1937). Properties of sufficiency and statistical tests. *Proceedings of the Royal Society of London. Series A - Mathematical and Physical Sciences, 160*, 268–282. https://doi.org/10.1098/rspa.1937.0109

Beavers, A. S., Lounsbury, J. W., Richards, J. K., Huck, S. W., & Skolits, G. J. (2013). Practical considerations for using exploratory factor analysis in educational research. *Practical Assessment, Research and Evaluation, 18*. https://doi.org/10.7275/qv2q-rk76

Biggs, J., & Madnani, N. (2022). Factoranalyzer. URL: https://factor-analyzer.readthedocs.io/en/latest/factor_analyzer.html.

Bray, A., Devitt, A., Banks, J., Sanchez Fuentes, S., Sandoval, M., Riviou, K., Byrne, D., Flood, M., Reale, J., & Terrenzio, S. (2023). What next for universal design for learning? A systematic literature review of technology in udl implementations at second level. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13328

Capp, M. J. (2017). The effectiveness of universal design for learning: A meta-analysis of literature between 2013 and 2016. *International Journal of Inclusive Education, 21*, 791–807. https://doi.org/10.1080/13603116.2017.1325074

Cast. (2018). *Universal Design for Learning Guidelines version 2.2*. http://udlguidelines.cast.org.

CAST. (2024). *AboutCAST*. https://www.cast.org/about/about-cast.

Cerezo, R., Sánchez-Santillán, M., Paule-Ruiz, M. P., & Núñez, J. C. (2016). Students' lms interaction patterns and their relationship with achievement: A case study in higher education. *Computers in Education, 96*, 42–54. https://doi.org/10.1016/j.compedu.2016.02.006

Cheung, K. K. C., & Tai, K. W. H. (2021). The use of intercoder reliability in qualitative interview data analysis in sci- ence education. arXiv *Research in Science & Technological Education*. https://doi.org/10.1080/02635143.2021.1993179, 10.1080/02635143.2021.1993179.

Dickler, R., Gobert, J., Adair, A., & Olsen, J. (2021). *Using a teacher dashboard to support students remotely on science inquiry*. EasyChair Preprint.

Distefano, C., Zhu, M., & Mîndrilã, D. (2009). Understanding and using factor scores: Considerations for the applied researcher. *Practical Assessment, Research and Evaluation, 14*, 20. https://doi.org/10.7275/da8t-4g52

Dooley, L., & Makasis, N. (2020). Understanding student behavior in a flipped classroom: Interpreting learning analytics data in the veterinary pre- clinical sciences. *Education Sciences, 10*, 1–14. https://doi.org/10.3390/educsci10100260

Edyburn, D. (2010). Would you recognize universal design for learning if you saw it? Ten propositions for new directions for the second decade of udl. *Learning Disability Quarterly, 33*, 33–41. https://doi.org/10.1177/073194871003300103

Edyburn, D. L. (2021). Universal usability and universal design for learning. *Intervention in School and Clinic, 56*, 310–315, 10.1177/1053451220963082.

Estacio, R. R., & Raga, R. C. (2017). Analyzing students online learning behavior in blended courses using moodle. *Asian Association of Open Universities Journal, 12*, 52–68. https://doi.org/10.1108/AAOUJ-01-2017-0016

Everitt, B. S., Landau, S., Leese, M., & Stahl, D. (2011). *Cluster analysis*. Wiley.

Fisher, R., Perényi, Á., & Birdthistle, N. (2021). The positive relationship between flipped and blended learning and student engagement, performance and satisfaction. *Active Learning in Higher Education, 22*, 97–113. https://doi.org/10.1177/1469787418801702

Fisseler, B. (2023). Universal design for learning in der lehrer*innenbildung. ein weg zu mehr digitaler inklusion und teilhabe? In J. R. Schluchter, & J. Betz (Eds.), *Schulische Medienbildung und Digitalisierung im Kontext von Behinderung und Benachteiligung* (pp. 313–333).

Florian, L., & Black-Hawkins, K. (2011). Exploring inclusive pedagogy. *British Educational Research Journal, 37*, 813–828. https://bera-journals.onlinelibrary.wiley.com/doi/abs/10.1080/01411926.2010.501096.

Florian, L., & Spratt, J. (2013). Enacting inclusion: A framework for interrogating inclusive practice. *European Journal of Special Needs Education, 28*, 119–135. https://doi.org/10.1080/08856257.2013.778111

Frey, A., Taskinen, P., Schütte, K., Prenzel, M., Artelt, C., Baumert, J., Blum, W., Hammann, M., Klieme, E., & Pekrun, R. (2009). *PISA-2006- Skalenhandbuch Dokumentation der Erhebungsinstrumente*. Waxmann.

Furini, M., Gaggi, O., Mirri, S., Montangero, M., Pelle, E., Poggi, F., & Prandi, C. (2022). Digital twins and artificial intelligence. *Communications of the ACM, 65*, 98–104. https://doi.org/10.1145/3478281

Göransson, K., & Nilholm, C. (2014). Conceptual diversities and empirical shortcomings - a critical analysis of research on inclusive education. *European Journal of Special Needs Education, 29*, 265–280. https://doi.org/10.1080/08856257.2014.933545

Gronseth, S. L., & Dalton, E. M. (2019). *Universal access through inclusive instructional design: International perspectives on UDL. Routledge*.

Hartigan, J. A., & Wong, M. A. (1979). Algorithm as 136: A k-means clustering algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics), 28*, 100–108. URL: http://www.jstor.org/stable/2346830.

Heller, K. A., & Perleth, C. (2000). *Kognitiver Fähigkeitstest für 4. bis 12*. Revision: Klassen.

Hodges, C., Moore, S., Lockee, B., Trust, T., & Bond, A. (2020). The difference between emergency remote teaching and online learning. https://er.educause.edu/articles/2020/3/the-difference-between-emergency-remote-teaching-and-online-learning.

Hollingshead, A., Lowrey, K. A., & Howery, K. (2020). Universal design for learning: When policy changes before evidence. *Educational Policy, 1*. https://doi.org/10.1177/0895904820951120. –27doi.

Holstein, K., Hong, G., Tegene, M., McLaren, B. M., & Aleven, V. (2018). The classroom as a dashboard: Co-Designing wearable cognitive augmentation for k-12 teachers. In *Proceedings of the 8th international Conference on learning analytics and knowledge* (pp. 79–88). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/3170358.3170377. URL:.

Hung, J. L., & Zhang, K. (2008). Revealing online learning behaviors and activity patterns and making predictions with data mining techniques in online teaching. *Journal of Online Learning and Teaching, 4*.

Ifenthaler, D., & Drachsler, H. (2020). Learning analytics. In H. Niegemann, & A. Weinberger (Eds.), *Lernen mit Bildungstechnologien: Prax- isorientiertes Handbuch zum intelligenten Umgang mit digitalen Medien* (pp. 1–20). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-662-54373-3_42-1.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*, 264–323. https://doi.org/10.1145/331499.331504

Jin, X., & Han, J. (2011). K-means clustering. In C. Sammut, & G. I. Webb (Eds.), *Encyclopedia of machine learning* (pp. 563–564). Boston, MA: Springer US. https://doi.org/10.1007/978-0-387-30164-8_425.

Jo, I. H., Kim, D., & Yoon, M. (2014). Analyzing the log patterns of adult learners in lms using learning analytics. In *Proceedings of the Fourth international Conference on learning analytics and knowledge* (pp. 183–187). New York, NY, USA: Association for Computing Machinery. https://doi.org/10.1145/2567574.2567616.

Kaiser, H. F., & Rice, J. (1974). *Little jiffy, mark iv. Educational and Psychological Measurement 34*, 111–117. https://doi.org/10.1177/001316447403400115

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data*. John Wiley and Sons, Inc. https://doi.org/10.1002/9780470316801

Khosravi, H., & Cooper, K. (2017). Using learning analytics to investigate patterns of performance and engagement in large classes. *SIGCSE '17: Proceedings of the 2017 ACM SIGCSE Technical Symposium on Computer Science Education*, 309–314. https://doi.org/10.1145/3017680.3017711

King-Sears, M. E., Stefanidis, A., Evmenova, A. S., Rao, K., Mergen, R. L., Owen, L. S., & Strimel, M. M. (2023). Achievement of learners receiving udl instruction: A meta-analysis. *Teaching and Teacher Education, 122*, Article 103956. URL: https://www.sciencedirect.com/science/article/pii/S0742051X22003316.

Klopp, E. (2010). *Explorative Faktorenanalyse*. https://psycharchives.org/en/item/8bd2d5e7-0941-4198-ab23-e94e19e47223.

Liao, C. H., & Wu, J. Y. (2023). Learning analytics on video-viewing engagement in a flipped statistics course: Relating external video-viewing patterns to internal motivational dynamics and performance. *Computers in Education, 197*. https://doi.org/10.1016/j.compedu.2023.104754

Long, P., & Siemens, G. (2011). Penetrating the fog: Analytics in learning and education. *Educause Review, 46*.

Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., & Clarebout, G. (2011). Tool-use in a blended undergraduate course: In search of user profiles. *Computers and Education, 57*, 2135–2144. https://doi.org/10.1016/j.compedu.2011.05.010

Luxford, C. J., & Bretz, S. L. (2014). Development of the bonding representations inventory to identify student misconceptions about covalent and ionic bonding representations. *Journal of Chemical Education, 91*, 312–320.

MacQueen, J. (1967). Classification and analysis of multivariate observations. In *5th Berkeley Symp. Math. Statist. Probability* (pp. 281–297). University of California Los Angeles LA USA.

Maier, U., & Klotz, C. (2022). Personalized feedback in digital learning environments: Classification framework and literature review. *Computers in Education: Artificial Intelligence 3*, Article 100080. URL: https://www.sciencedirect.com/science/article/pii/S2666920X22000352 10.1016/j.caeai.2022.100080.

Manning, C. D., Raghavan, P., & Schütze, H. (2008). *Introduction to information Retrieval*. Cambridge University Press, 10.1017/CBO9780511809071.

Marsland, S. (2015). *Machine learning* (Second edi ed.). CRC Press.

Mayringer, H., & Wimmer, H. (2014). *Salzburger Lese-Screening für die Schulstufen 2–9* (SLS 2–9).

Merkt, M., Hoppe, A., Bruns, G., Ewerth, R., & Huff, M. (2022). Pushing the button: Why do learners pause online videos?. In *, 176*. *Computers & Education*. Elsevier. https://doi.org/10.1016/j.compedu.2021.104355.

Meyer, S. (2021). *Entwicklung und Evaluation eines Multiple-Choice-Tests im three- tier format zur Erhebung von Daten über Lernendenvorstellungen im Themenbereich der Bildung und Bindung von Ionen (Master's thesis)*. Leibniz University Hannover.

Mohamad, I. B., & Usman, D. (2013). Standardization and its effects on k-means clustering algorithm. *Research Journal of Applied Sciences, Engineering and Technology, 6*, 3299–3303. https://doi.org/10.19026/rjaset.6.3638

Navaitienė, J., & Stasiūnaitienė, E. (2021). The goal of the universal design for learning: Development of all to expert learners. In A. Galkienė, & O. Monkevičienė (Eds.), *Improving inclusive education through universal design for learning* (pp. 23–57). Cham: Springer International Publishing. https://doi.org/10.1007/978-3-030-80658-3_2.

Parkes, M., Gregory, S., Fletcher, P., Adlington, R., & Gromik, N. (2015). Bringing people together while learning apart: Creating online learning environments to support the needs of rural and remote students. *Australian and International Journal of Rural Education, 25*, 66–78. https://doi.org/10.3316/aeipt.215238

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, E. (2011). Scikit-learn: Machine learning in python. *Journal of Machine Learning Research, 12*, 2825–2830.

Pelánek, R. (2017). Bayesian knowledge tracing, logistic models, and beyond: An overview of learner modeling techniques. *User Modeling and User-Adapted Interaction, 27*, 313–350, 10.1007/s11257-017-9193-2. exponential moving averag.

Phillips, R., Maor, D., Preston, G., & Cumming-Potvin, W. (2012). Exploring learning analytics as indicators of study behaviour. https://api.semanticscholar.org/CorpusID:18364213.

Rao, K., Edelen-Smith, P., & Wailehua, C. U. (2015). Universal design for online courses: Applying principles to pedagogy. *Open Learning: The Journal of Open, Distance and e-Learning, 30*, 35–52. https://doi.org/10.1080/02680513.2014.991300

Rogers-Shaw, C., Carr-Chellman, D. J., & Choi, J. (2018). Universal design for learning: Guidelines for accessible online instruction. *Adult Learning, 29*, 20–31. URL: https://api.semanticscholar.org/CorpusID:196050316.

Roski, M., Ewerth, R., Hoppe, A., & Nehring, A. (2024). Exploring Data Mining in Chemistry Education: Building a Web-Based Learning Platform for Learning Analytics. *Journal of Chemical Education*. https://doi.org/10.1021/acs.jchemed.3c00794

Roski, M., Hoppe, A., & Nehring, A. (2024). *Supporting Inclusive Science Learning through Machine Learning: The AIISE Framework. In Uses of Artificial Intelligence in STEM Education*. Oxford University Press.

Roski, M., Sebastian, R., Ewerth, R., Hoppe, A., & Nehring, A. (2023). In N. Wang, G. Rebolledo-Mendez, N. Matsuda, O. C. Santos, & V. Dimitrova (Eds.), *Artificial Intelligence in Education. AIED 2023. Lecture Notes in Computer Science, 13916*. Cham: Springer. https://doi.org/10.1007/978-3-031-36272-9_42.

Roski, M., Walkowiak, M., & Nehring, A. (2021). Universal Design for Learning: The More, the Better? *Educ. Sci., 11*, 164. https://doi.org/10.3390/educsci11040164.

Schmitz, M., van Limbeek, E., Greller, W., Sloep, P., & Drachsler, H. (2017). Opportunities and challenges in using learning analytics in learning design. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven approaches in digital education* (pp. 209–223). Cham: Springer International Publishing.

Shahapure, K. R., & Nicholas, C. (2020). Cluster quality analysis using silhouette score. In *2020 IEEE 7th international Conference on data science and advanced analytics* (pp. 747–748). DSAA). https://doi.org/10.1109/DSAA49011.2020.00096.

Spörer, N., Henke, T., & Bosse, S. (2021). Is there a dark side of co-teaching? A study on the social participation of primary school students and their interactions with teachers and classmates. *Learning and Instruction, 71*, Article 101393. URL: https://www.sciencedirect.com/science/article/pii/S0959475219309120.

Stinken-Rösner, L., Rott, L., Hundertmark, S., Baumann, T., Menthe, J., Hoffmann, T., … Abels, S. (2020). Thinking Inclusive Science Education from two Perspectives. *Inclusive Pedagogy and Science Education. RISTAL, 3*, 30–45.

Szumski, G., Smogorzewska, J., & Karwowski, M. (2017). Academic achievement of students without special educational needs in inclusive classrooms: A meta-analysis. *Educational Research Review, 21*, 33–54. URL: https://www.sciencedirect.com/science/article/pii/S1747938X17300131.

Torsheim, T., Cavallo, F., Levin, K. A., Schnohr, C., Mazur, J., Niclasen, B., & Currie, C. (2016). Psychometric validation of the revised family affluence scale: A latent variable approach. *Child Indicators Research, 9*, 771–784. https://doi.org/10.1007/s12187-015-9339-x

UNESCO. (2005). Guidelines for inclusion: Ensuring access to education for all. URL: http://www.ibe.unesco.org/sites/default/files/Guidelines_for_Inclusion_UNESCO_2006.pdf.

UNESCO-UNEVOC. (2021). *Medium-term Strategy for 2021-2023*. URL: https://unesdoc.unesco.org/ark:/48223/pf0000375156.locale=en.

United Nations. (2020). *Quality education: Why it matters*. The Sustainable Development Goals. URL: https://www.un.org/sustainabledevelopment/wp-content/uploads/2017/02/4_Why-It-Matters-2020.pdf.

United Nations. (2022). *The sustainable development goals report 2022*. URL: https://desapublications.un.org/publications/sustainable-development-goals-report-2022.

Vytasek, J. M., Patzak, A., & Winne, P. H. (2019). Analytics for student engagement. In *Machine learning Paradigms*. URL: https://api.semanticscholar.org/CorpusID:109938478.

Werner, C. (2014). *Explorative Faktorenanalyse: Einführung und Analyse mit R*.

Weygandt, B. (2021). Durchführung und ergebnisse der explorativen faktorenanalyse. In *Mathematische Weltbilder weiter denken: Empirische Untersuchung des Mathematikbildes von Lehramtsstudierenden am Übergang Schule–Hochschule sowie dessen Veränderungen durch eine hochschuldidaktische Mathematikvorlesung* (pp. 253–289). Wiesbaden: Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-34662-1_14.

Wu, J. (2012). Cluster analysis and k-means clustering: An introduction. In *Advances in K-means clustering: A data mining Thinking* (pp. 1–16). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-642-29807-3_1.