Learners' Preference for Reference Frame Before, During, and After Task Performance

Name: Karlijn Tuin

Student number: 5509890

First assessor: Liesbeth Kester

Second assessor: Eva Janssen

Date: 07-06-2021

Word count: 5800

Abstract

Learners often have difficulty with regulating their learning processes before, during, and after task performance. Learning environments need to offer support to help them with this. In immersive learning environments (ILEs), learning analytics (LA) are available that could provide such support by means of learning analytic dashboards (LADs). The visualization of LA in LADs can support SRL skills before, during, and after task performance. However, previous research suggest that this will only happen if learners appreciate the LA visualizations on their dashboards. Therefore, this study aims to identify preferences that learners have for LADs used before, during, and after task performance. Using a specific means to present LA, namely reference frames. Reference frames provide an anchor point for comparing LA that could help learners evaluate their learning. This study used three different types of reference frames: the social reference frame, the progress reference frame, and the achievement reference frame. This research reveals that, regardless of timing (i.e., before, during, or after task performance), learners have a preference for LADs with a progress reference frame. The results of this study might contribute to a more tailor made design of LADs in ILEs that could successfully support learners in regulating their learning.

Keywords: learning analytics, learning analytic dashboards, reference frames, self-regulated learning, immersive learning environment

Learners' Preference for Reference Frame Before, During, and After Task Performance

Digitalization and emerging novel technologies have an continuous impact on our lives and the way we work. Organizations feel the urge to go along with these developments and need to ensure that their employees can adapt to the constantly changing environment as well. Therefore, it is important that lifelong learning (i.e., continuous learning) can occur in the workplace (Tynjälä, 2008; Wang, 2017). In working environments, informal learning (e.g., autonomous, learning is a side task, learning is a means to an end) is more common than formal learning (e.g., controlled, structured, learning is the goal). It is often the responsibility of the employee to indicate whether he/she desires to develop themselves. However, it tends to be that most employees do not know how to learn or do not proactively initiate a learning process (Margaryan et al., 2013). To identify their personal learning needs and to initiate the learning processes, the employee needs to develop self-regulated learning (SRL) skills.

Panadero and Alonso-Tapia (2014) defined SRL as "the control that students have over their cognition, behavior, emotions, and motivation through the use of personal strategies to achieve the goals they have established" (p. 450). According to Zimmerman's (2000) cyclical SRL model, three phases can be distinguished: forethought, performance, and self-reflection.

The *forethought* phase can be seen as the phase which occurs before the task begins. The most important aspects of this phase are task analysis and self-motivation beliefs. The learner activates self-regulation by thinking about strategies that will help them to perform during the task (i.e., goal setting and strategic planning) and through self-motivational beliefs (i.e., self-efficacy, outcome expectations, task interest, task value, goal orientation). The expected level of performance and assessment criteria are essential variables during this phase.

In the *performance* phase, learners perform the actual task. At this phase, it is crucial that learners stay motivated and monitor their progress concerning their goals. Self-control (i.e., specific strategies, self-instruction, imagery, time management, help-seeking, interest incentives, self-consequences) and self-observation (i.e., self-monitoring, self-recording) are the two key processes during this phase.

The *self-reflection* phase, which takes place after the task is performed, is all about self-assessment of performance (i.e., self-evaluation, causal attribution) and recognizing success and failure (i.e., self-satisfaction/affect, adaptive/defensive decisions; Panadero, 2017; Panadero & Alonso-Tapia, 2014; Zimmerman and Moyland, 2009).

SRL skills are of great importance for workplace learning. However, little is known about how employees self-regulate their learning processes (Margaryan et al., 2013; Tynjälä, 2008). Research has found that advanced learning technologies such as simulations, serious games, virtual reality, intelligent tutoring systems, and so on could potentially foster self-regulated learning (Azevedo & Gašević, 2019)

Immersive Learning Environments, Learning Analytics, and Learning Analytic Dashboards

An emerging technology for workplace learning is the use of immersive learning environments (ILE), such as virtual reality training scenarios. These environments (i.e., realistic computer-generated worlds) enable learners to perform complicated, uncommon, and/or dangerous tasks (e.g., with people, objects, places, locations) in a real-life setting (Herrington et al., 2007). Subsequently, it enables learners to deepen their higher-order thinking skills (e.g., analyzing, reasoning, synthesizing, evaluating) to create new knowledge (Beckem, 2012). Research has shown that ILE contributes to the learner's engagement, motivation, inquiry and collaboration skills, and learning outcomes (Dede, 2009).

ILE yields data for learning analytics (LA), another emerging technology for workplace learning. LA enables the collection, analysis, and reporting of learner behavior (e.g., training effort, progress, grades). It aims to give meaningful information before, during, and after a task and stimulates reflecting on learning outcomes and informing decision-making as a result (Wang, 2017). Previous research acknowledges that LA could act as an instrument to stimulate students' SRL skills in an online learning setting (Kim et al., 2018). LA could be valuable for learners, teachers, and other stakeholders.

A manner to present the collected data is to use learning analytic dashboards (LADs). LADs are visual presentations –in the form of feedback– of the collected data and can potentially stimulate metacognitive skills as it provides insight into the learner's performance and behavior. Furthermore, it enables the learner to reflect and optimize their learning, assess their ability, and adapt their behavior (Matcha et al., 2020; Park & Jo, 2015; Jivet et al., 2020). For LADs to have impact, they need to be presented at the right moment. By doing so, the learner can use the information to support their learning processes (Gibbs & Simpson, 2005). Therefore, LADs can be presented during different task performances in the learning process: before the task performance, during the task performance, and after the task performance.

This research paper refer to "before task performance" when the LAD is shown before the learner starts a task. During this performance, aspects of the forethought phase of the SRL cycle could be targeted. LAD can present information that could attain data from similar or earlier attempts of the task by the learner. The learner might be more motivated to think about the right approach and set goals for performing the task, when confronted with their previous results.

This research paper refer to "during task performance" when the LAD is shown to the learner while still performing a task. Aspects of the performance phase of the SRL cycle could be targeted after the learner has completed the task. When the LAD presents information about the learner's progress so far, the learner might use this input to adjust or reconsider their learning strategy or seek help.

Lastly, the "after task performance", refers to when the LAD is shown to the learner after completing the task. During this performance, aspects of the self-reflection phase of the SRL cycle could be targeted. When presenting the LAD to the learner with the information referred to the completed task's performance, the learner might be stimulated to reflect on their performance and task approach.

According to Jivet et al. (2017), most LADs only support the self-reflection phase of the SRL cycle. These findings emphasize the need to encapsulate LADs into the instructional design in order to perform all the SRL phases. To promote sense-making and SRL skills, LADs need to be designed as pedagogical interventions (Matcha et al., 2020; Wise et al., 2014). A good understanding of how learning theories can be used is needed to design and use LADs. Previous research also discussed that learning analytic systems (e.g., ILE) need to be more user-centered. Yet, little is known about how learners make use of LADs. SRL support by LA visualizations on learners' dashboards will only be effective if learners also appreciate them. Therefore, this study focuses on the preferences learners have for LADs used before, during, and after task performance, using a specific means to present LA, namely reference frames.

Reference Frames

Reference frames can be defined as "the comparison point to which students orient when they examine their analytics" (Wise, 2014, p. 208). They are pedagogical interventions (e.g., personalized feedback) and encourage sense-making and the use of learning analytics.

Jivet et al. (2017) analyzed 26 LADs on how information was presented to the learners and distinguished three types of reference frames: (a) social, (b) progress, and (c) achievement. The *social* reference frame (SRF) enables learners to compare their performance with the results of their peers. The *progress* reference frame (PRF) enables learners to compare their present performance data with their past performance data. The *achievement* reference frame (ARF) enables learners to compare their results with goals proposed by themselves or defined by the teacher (Jivet et al., 2020). The achievement goal theory can be used to explain some of the potential merits of the different types of reference frames. This theory describes how beliefs and cognition motivate learners to academic success and achievement (Ames, 1992). In addition, the different types of reference frames might be suited to support different phases in SRL.

The SRF is most commonly used and can be applied in various ways. The comparison can be made with the whole class, teammates, previous graduates, top students, and peers with similar goals. Based on the achievement goal theory perspective, the SRF might benefit learners with a performance goal orientation. This orientation might be useful to define individuals' purposes for approaching and engaging in a task. Learners with a performance goal orientation focus on comparing themselves with others to see how successful they perform (Ames, 1992). By comparing the learner's outcome with their peers' outcomes, the SRF might stimulate the self-reflection phase of the SRL cycle.

The PRF allows learners to compare historical data of their performances with their current level of performance. As a result, the learner can see their progress over time. This type of frame might align with the mastery goal-oriented concept of achievement goal theory. Learners concern the importance of mastering the materials and tasks when they are mastery goal-oriented. They compare their learning outcomes with previous outcomes to see how they progressed (Ames, 1992). By comparing the learner's present performance data with their past performance data, the PRF might stimulate the performance phase of the SRL cycle.

The ARF allows learners to compare their performance level with desired goals from an internal perspective (i.e., set by themselves) or external perspective (i.e., set by their teacher). This frame might suit both performance goal-oriented learners and mastery goal-oriented learners because this frame makes it possible to focus on the learning outcomes and how they are mastering the materials and tasks. Since the ARF made it possible to compare the learner's current level of performance with the desired goal, this frame might stimulate the forethought phase of the SRL cycle.

According to Jivet et al. (2017), social reference frames are used more frequently than the achievement reference frame and the progress reference frame because motivation and engagement of learners increase by comparing to others. Furthermore, they argued that learning should be about improving knowledge and skills instead of competition between learners. User acceptance could play an important role in ensuring that the LAD stimulates SRL skills (Sedrakyan et al., 2020). The same type of feedback (e.g., reference frame) may not suit the learner during different task performances. The different needs of learners and their preference for reference frames needs should be taken into account. Considering the design of LADs should be functioning as a pedagogical aid to stimulate motivation and engagement of learners with various performance levels. Jivet et al. underline the need for further research on how different designs of LAD are perceived by learners and how they can improve their learning abilities.

Present Study

Previous research defined that LA mainly focuses on measuring SRL instead of supporting the different SRL phases. There is an urge to further investigate how LA support mechanisms (e.g., LADs) fosters SRL skills of learners in an online learning environment (Jivet et al., 2017; Viberg et al., 2020). Earlier studies also discussed that most of the LADs are designed to support teachers instead of learners, that LADs need to support learners SRL skills more, and that ILE needs to be more user-centered (Jivet et al., 2018; Margaryan et al., 2013; Matcha et al., 2020). For LADs to effectively support SRL skills, learners need to appreciate the visualizations. Therefore, the current study focuses on the preferences learners have for reference frames before, during, and after task performance. The following research question arises: What type of reference frame do learners prefer before, during, and after task performance? An exploratory study will be conducted in a workplace setting to answer this question. With these insights, we can contribute to the development of design guidelines for LA's presentation for ILE. The following sections describe the method used for the present study, explain the results, and conclude with a discussion about the research question.

Method

Participants and Study Design

Participants were 70 employees of Merck KGaA, who were all trainees of Merck's chemical plant operator apprenticeship program. Sixty-eight participants completed the demographic questions. A majority of 64.3% are male, 22.9% are female, and 10.0% did not want to say their gender. 14.3% of the participants belong to the 18-19 age group, 27.1% are aged 20-21, 25.7% are aged 22-23, 11.4% are aged 24-25, 10.0% are aged 26-27, 4.4% are aged 28-29, and 4.4% are 30 years old and above. The participants were all German-speaking.

This study used particular reference frames to present learning analytics to examine students' preferences before, during, and after task performance. We conducted a within-subject quantitative, explorative study. Participants were recruited through purposive sampling. Ethical approval of the Ethics Review Board of the Faculty of Social & Behavioural Sciences was obtained. The researchers asked for active consent and informed the participants that they could withdraw at any moment without a given reason and disadvantage. Participation in this study was entirely voluntary; participants were neither compensated financially nor in credits for the apprenticeship program.

Instruments and Materials

The data collection was part of a more considerable study and consisted of questionnaires, adaptive comparative judgment (ACJ), and interviews. To increase the study's validity, the questionnaire and information in the ACJ were facilitated in German, the native language of most participants.

Demographic Data Questionnaire

With the use of a short Demographic Data Questionnaire, information about each participant's age and gender was collected. This questionnaire was distributed via the online survey software Microsoft Forms (Version: February 2021).

Reference Frames Mock-ups

The researchers created in total twelve print screens of reference frame mock-ups with fictional data. Four mock-ups per intended task performance: (a) before task performance, presented just before beginning the simulator task; (b) during task performance, presented when the participant pauses the simulator; and (c) after task performance, presented when the participant completed the task. The mock-ups were specially designed for a chemical process industry virtual reality training simulator and established through a collaboration between educational scientists and instructional designers from Utrecht University and content matter experts from the field of chemical engineering.

Each mock-up consisted of four elements with the aim to stimulate performance behaviors and self-regulated learning behaviors (see Figure 1). Element 1 provided information on which frame their results were compared with. Element 2 encouraged the participant to review the displayed and highlighted data, contextualized by the task performance. Element 3 illustrated the participant's training effort in the way of the time spent in the simulator. Element 4 provided feedback on the performance of the participant compared to a specific reference frame. Four reference frames are used: (a) social reference frame (SRF), which compared the participant's data with data of their peers (see Figure 2); (b) progress reference frame (PRF), which compared the participant's data with their own data from previous attempts; (c) external achievement reference frame (E-ARF), which compared the participant's data with the achievement goals set by the trainer and (d) internal achievement reference frame (I-ARF), which compared the participant's data with the achievement goals set by themselves. To ensure that preferences were made on the elements rather than the aesthetical design or other design elements, the look-and-feel in every task performance was as similar as possible.

Figure 1

Example of Reference Frame Mock-up with the Four Elements Highlighted

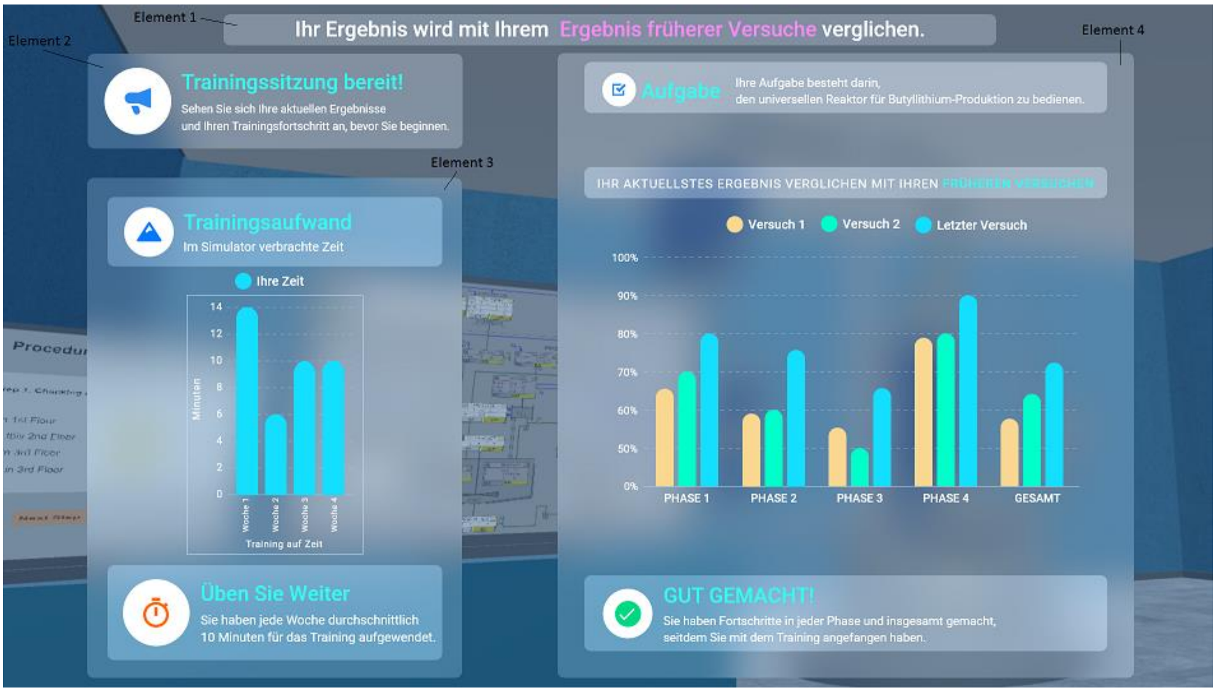
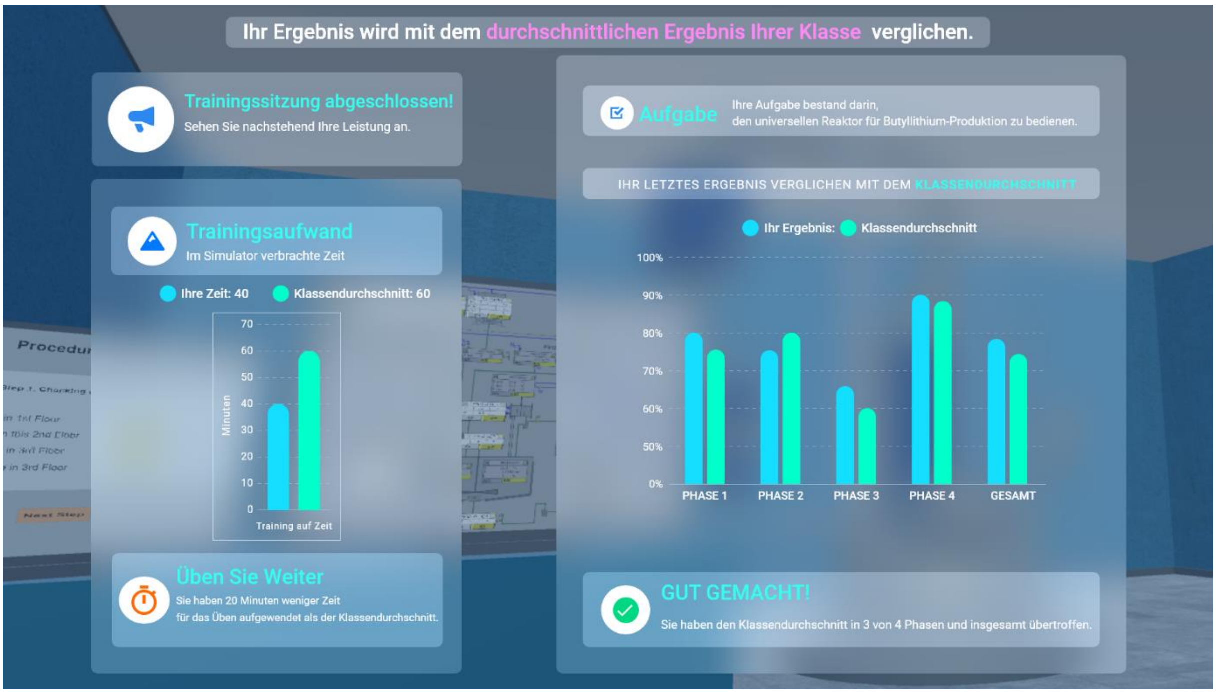


Figure 2

Example of the Social Reference Frame Mock-up in the After Task Performance



RM Compare

This study used the online system software RM compare to measure the learner's preference for reference frame within different task performances (https://www.compare.rm.com). This software system allows learners to make adaptive comparative judgments (ACJ). ACJ is a modification of comparative judgment (CJ) introduced by Thurstone (1927). Thurstone referred to CJ as *the Law of comparative judgment*. Two presentations are being compared and judged which one is preferred. It is seen as a more straightforward and instinctive technique since decisions are based on comparisons instead of judging one by one (Laming, 2003). Pollitt (2004) modified the CJ technique to a web-based adaptive comparative judgment through the continuing evolution of technology. After each object that a participant judged, an algorithm revised the information about the object and automatically generated a new pair for judgments. When every object is judged once, the algorithm generated pairs that are expected to be close in quality (e.g., both objects which won their comparison). The iterations of the pairs chosen by the algorithm continue for multiple rounds. It is finished when enough data is collected to produce a highly reliable and accurate rank order (Pollitt, 2012; Bartholomew & Yoshikawa, 2018). The advantage of adaptiveness is efficiency. Since not every pair needs to be seen by every judge, it takes less time to create a reliable rank order.

In the present study, participants judged the four particular designs for the different task performances: before, during, and after. For every comparison in the before task performance, participants were asked, "Which design do you think will help you best with planning, goal setting, and motivation?". For every comparison in the during task performance, participants were asked, “Which design do you think will help you best with keeping track of how well you are doing at the task and following a plan you made?”. For every comparison in the after task performance, participants were asked, “Which design do you think is best for helping you reflect upon your performance and judging or assessing your own performance?”. The researchers had set up the system that each participant had to make seven comparisons. This was to ensure that the participant compared every possible pair of reference frames. Finally, after all the comparisons, RM Compare provided the rank order and the associated data.

Procedure

One week prior to the data collection, the participants received the information letter and informed consent. If the participant was willing to participate and undersigned the informed consent, the at Merck situated researcher sent an email with an invitation to join one of the six webinars with the research team. In this webinar, the team walked through the research procedure. These webinars were held in German by a native German-speaking researcher. To maintain confidentiality, the responses were pseudonymized, which concealed the identity of the answers to the questions and judgments.

The data collection was part of a more considerable study. Once the participant was informed about the procedure, the data collection started and took about 20 minutes. There were no time limits for participating in this study. The data collection consisted of three steps. In the first step, the participant filled in the questionnaire on their demographics: age and gender. This was directly followed by the Goal Orientation questionnaire distributed in Microsoft Forms. In the second step, the adaptive pairwise comparisons were made in the ACJ software system RM Compare. A paired comparison (i.e., two different reference frames) was shown to the participant, and the participant was asked which one he or she liked better. After the comparison, a new paired comparison was shown. This went on to seven paired comparisons for every task performance. Finally, the Self-regulated Online Learning – Questionnaire-Revised to measure self-regulated learning was distributed and filled in using Microsoft Forms. The data collection concluded with semi-structured interviews. A sample participant, proficient English-speaking, of the larger group was asked to participate in this.

Analytic Approach

To determine the preferred reference frame for each task performance, this study used the ACJ online software system RM Compare. The data collection yielded in total 1416 pairwise comparisons. Each participant compared seven pairs of reference frames during each task performance except for three participants. These three participants did not make all the seven comparisons but only one each. However, since we only use the comparisons, the three participants were included in the total number of comparisons.

Before analyzing the results, assumptions for ACJ were checked. Every reference frame within each task performance could be distinguished from each other reference frame. Each judgment was independent, and the judges were equal in their ability to distinguish between the reference frames. No violations of these assumptions were found. RM Compare automatically computed the level of reliability after each task performance. The reliability of the ACJ ranking can be seen as a judgment consistency coefficient. It is equivalent to Cronbach's alpha and indicated the consistency level between all the judges and their judgments.

Subsequent to this, we checked whether the reference frames were chosen as a win over another perchance. Therefore, we used all comparisons made per task performance. Since there is only a categorical, nominal, dependent variable (i.e., reference frames), a chi-square test for goodness of fit (with $\alpha = .05$) was conducted. Before conducting the analyses, the assumptions were checked. All participants were independent, and all the expected frequencies were greater than five. The chi-square test was conducted with the frequencies of how many times a reference frame had 'won' the comparison over another reference frame.

Results

Before Performance Task

Table 1 shows the results after 472 adaptive comparative judgments for the before task performance. The ranking showed a high level of judge consistency coefficient of 0.84.

Table 1

Descriptive Statistics for the Reference Frames in the Before Task Performance

Reference frame	Win frequency		Rank
	n	%	
Social reference frame	88	37.1	4
Progress reference frame	151	64.0	1
External achievement reference frame	122	51.7	2
Internal achievement reference frame	111	47.2	3

Note. Win frequencies are shown for comparisons with the social reference frame (*n* = 237), progress reference frame (*n* = 236),) external achievement reference frame (*n* = 236), and internal achievement reference frame (*n* = 235).

A chi-square test for goodness of fit (with $\alpha = .05$) was used to assess whether the four reference frames were significantly more or less preferred for the before task performance. There was a significant difference between the frequencies, $\chi2$ (3, N = 472) = 17.41, p = .001, indicating that the participants do have a preference.

Subsequent to this, a chi-square test for goodness of fit (with $\alpha = .05$) between pairs of reference frames was used to assess whether one frame over another was significantly more preferred. The frequencies between the SRF and PRF (i.e., rank 4 vs. rank 1) differed significantly from each other, $\chi2$ (1, n = 239) = 16.61, $p < .001$, indicating that the PRF was more preferred over the SRF. Furthermore, the frequencies between the SRF and E-ARF (i.e., rank 4 vs. rank 2) differed significantly from each other, $\chi2$ (1, n = 210) = 5.51, p = .019, suggesting that the E-ARF was more preferred over the SRF. Although the I-ARF had more wins than the SRF (i.e., rank 4 vs. rank 3), the frequencies were statistically equal from each other, $\chi2$ (1, n = 199) = 2.66, p = .103. Suggesting that this result should be handled with care. The PRF has more wins than the E-ARF (i.e., rank 1 vs. rank 2). Yet there was no significant difference between the frequencies, $\chi2$ (1, n = 273) = 3.08, p = .079. Suggesting that this result should be handled with care. There was a statistically significant difference between the PRF and I-ARF frequencies (i.e., rank 1 vs. rank 3), $\chi2$ (1, n = 262) = 6.11, p = .013, indicating that the PRF was more preferred over the I-ARF. Despite the fact that the E-ARF has more wins than the I-ARF (i.e., rank 2 vs. rank 3), the frequencies were statistically equal from each other, $\chi2$ (1, n = 233) = 0.52, p = .471. Suggesting that this result should be handled with care.

During Task Performance

Table 2 shows the results after 472 adaptive comparative judgments for the during task performance. The ranking showed a moderate level of judge consistency coefficient of 0.65.

Table 2

Descriptive Statistics for the Reference Frames in the During Task Performance

Reference frame	Win frequency		Rank
	n	%	
Social reference frame	100	42.4	4
Progress reference frame	137	58.1	1
External achievement reference frame	133	56.4	2
Internal achievement reference frame	102	43.2	3

Note. Every reference frame was in the same amount of comparisons (n = 236).

A chi-square test for goodness of fit (with α = .05) was used to assess whether the four reference frames were significantly more or less preferred within during task performance. The frequencies differed significantly from each other, $\chi 2$ (3, N = 472) = 9.88, p = .020, indicating that the participants do have a preference.

In addition, a chi-square test for goodness of fit (with α = .05) between pairs of reference frames was used to assess whether one frame over another was significantly more preferred. A significant difference in frequencies between the SRF and PRF (i.e., rank 4 vs. rank 1), $\chi 2$ (1, n = 237) = 5.78, p = .016, and between the SRF and E-ARF (i.e., rank 4 vs. rank 2), $\chi 2$ (1, n = 233) = 4.67, p = .031, were found. Indicating that the PRF and E-ARF were more preferred over the SRF. Although the I-ARF had more wins than the SRF (i.e., rank 3 vs. rank 4), statistical equality was found in frequencies, $\chi 2$ (1, n = 202) = 0.02, p = .888. Indicating that this result should be handled with care. Despite the fact that the PRF had more wins than the E-ARF (i.e., rank 1 vs. rank 2), the frequencies were statistically equal, $\chi 2$ (1, n = 270) = 0.06, p = .808. Indicating that this result should be handled with care. There was a significant difference in frequencies between the PRF and I-ARF (i.e., rank 1 vs. rank 3), $\chi 2$ (1, n = 239) = 5.13, p = .024, indicating that the PRF was more preferred over the I-ARF. The frequencies between the E-ARF and I-ARF (i.e., rank 2 vs. rank 3) were statistically different from each other, $\chi 2$ (1, n = 235) = 4.09, p = .043, indicating that the E-ARF was more preferred over the I-ARF.

After Task Performance

Table 3 shows the results after 472 adaptive comparative judgments for the after task performance. The ranking showed a high level of judge consistency coefficient of 0.84.

Table 3

Descriptive Statistics for the Reference Frames in the After Task Performance

Reference frame	Win frequency		Rank
	n	%	
Social reference frame	94	39.8	4
Progress reference frame	164	69.2	1
External achievement reference frame	120	50.8	2
Internal achievement reference frame	94	40.0	3

Note. Win frequencies are shown for comparisons with the social reference frame (n = 236), progress reference frame (n = 237), external achievement reference frame (n = 236), and internal achievement reference frame (n = 235).

A chi-square test for goodness of fit (with α = .05) was used to assess whether the four reference frames were significantly more or less preferred for the after task performance. The frequencies were statistically different from each other, $\chi 2$ (3, N = 472) = 27.73, p <.001, indicating that the participants do have a preference.

Subsequent to this, a chi-square test for goodness of fit (with α = .05) between pairs of reference frames was used to assess whether one frame over another is significantly more preferred. The frequencies between the SRF and PRF (i.e., rank 4 vs. rank 1) were statistically different from each other, $\chi 2$ (1, n = 258) = 18.99, p < .001, indicating that the PRF was more preferred over the SRF. Numerically the E-ARF had more wins than the SRF (i.e., rank 2 vs. rank 4). Nevertheless, the frequencies were statistically equal, $\chi 2$ (1, n = 214) = 3.16, p =

.076. Suggesting that this result should be handled with care. The win distribution between I-ARF and the SRF (i.e., rank 2 vs. rank 4) was equal. However, the win percentage for I-ARF was higher. The frequencies were statistically equal from each other, $\chi2$ (1, n = 188) = 0.00, p = 1. Suggesting that this result should be handled with care. There was a statistically significant difference between the PRF and E-ARF frequencies (i.e., rank 1 vs. rank 2), $\chi2$ (1, n = 284) = 6.82, p = .009, indicating that the PRF was more preferred over the E-ARF. Moreover, the frequencies between the PRF and I-ARF (i.e., rank 1 vs. rank 3) were statistically different from each other, $\chi2$ (1, n = 258) = 18.88, p < .001. Suggesting that the PRF was more preferred over the I-ARF. The E-ARF and I-ARF (i.e., rank 2 vs. rank 3), differ in the amount of wins, but the frequencies were statistically equal, $\chi2$ (1, n = 214) = 3.16, p = .076. Indicating that this result should be handled with care.

Discussion

This study aimed to explore the preferred reference frame for different task performances from a learners' perspective. We focused on four reference frames: (a) social reference frame, (b) progress reference frame, (c) external achievement reference frame, and (d) internal achievement reference frame. Furthermore, we focused on three task performances: (a) before task performance, (b) during task performance, and (c) after task performance. This study used adaptive comparative judgments performed in the computer software system RM Compare. This software system provided us with a rank order of the four reference frames for the three task performances. In the following paragraphs, the findings, limitations, and implications are discussed.

Our research question asked what type of reference frame learners preferred for the different task performances. This study reveals that there is a preference for reference frames for the different task performances. The progress reference frame was most preferred over the other three reference frames for every performance task based on the win frequencies.

Furthermore, the frequency distribution was analyzed between all possible pairs of reference frames for each performance task. The results indicate that we cannot say with certainty that the progress reference frame is most preferred for the three performance tasks. Only in the after performance task, the progress reference frame was significantly different when compared with the other three reference frames.

A possible explanation for these results could be found in of the achievement goal theory. The trainees in this study might be more mastery goal-oriented. Trainees who are mastery goal-oriented are self-driven and have higher levels of perseverance. They tend to learn continuously and improve their skills (Ames, 1992; Dweck & Leggett, 1988). The preferred reference frame depends on what the learner wants to know about their performance. If the trainees are more mastery goal-oriented, they might be more interested in their progression relative to previous attempts. This could be an possible explanation for our results.

Previous research showed that comparing to peers who perform better can lead to negative feelings (e.g., stress and intimidation) and lower academic self-concept (Dijkstra et al., 2008; Wise et al., 2014). However, social comparison could also stimulate motivation, increase learner performance, and lead to feelings of superiority (Corrin & de Barba, 2015; Major et al., 1991). Despite these mixed results, the social reference frame is the most common reference frame for comparisons. Earlier research done by Jivet et al. (2017) revealed that the social reference frame was rated by students as the lowest item among 26 other items. Findings from the current study confirmed that the social reference frame is least preferred based on the win frequencies. However, we cannot say this with certainty. Only a significant difference was found in the during performance task between de social reference frame and the other tree frames. These findings may contribute in refuting to per default use the social reference frame and to consider the learner's need for reference.

Limitations and Future Research

It should be noted that our study knows several potential limitations. A first potential limitation is the limited generalizability of the study. Even though we do get an insight into learner preferences for reference frames, it only represents a homogeneous group in terms of the workplace. Future research could elaborate on this and replicate the study with a more heterogeneous sample of participants and a larger sample size.

A second limitation is the use of fictional data. The reference frame mock-ups were specially designed by a team of educational scientists, instructional designers, and content matter experts from the field of chemical engineering. However, it was designed with fictional data. It can be discussed whether the same results would be found with the learner's personal learning analytics. It is possible that learners have more feeling with the reference frames when it comes to their own learning analytics and that the preference might differ when compared to the present study. Future research could examine the preference for reference frames during virtual reality training followed by the participants.

Lastly, the way this research used the ACJ software system could be a potential limitation. ACJ software systems are designed for a great sample sizes. The advantage of ACJ is that not every possible comparison needs to be seen by every judge. After only 40-50% of the total comparisons, the required level of accuracy is reached (Pollitt, 2011). However, in this research, we set the number of comparisons at seven, and we only had four objects. Although we have reliable results, this set number of comparisons could have influenced the reliability and the other results. Looking into the comparisons made per participant, it occured that pairs were compared more than once. Sometimes with different outcomes. Future research could replicate this study but using the ACJ software system in an appropriate way.

Implications

Previous research has discussed the lack of insight into the perspective of and how learners perceive LADs (Buckingham Shum et al., 2019; Ochoa & Wise, 2020). To our knowledge, this study was the first to explore learner preferences for learning analytics regarding reference frames. Jivet (2017) already identified the three types of reference frames used in LADs. This study shows that learners have a preference for the progress reference frame.

In terms of practical use, this study can help inform policymakers with designing guidelines for LA's presentation for ILE. LADs are mainly developed and designed through the expertise of educational designers and LA researchers. However, in order for LADs to be fully effective, it is also essential to take the learner's perspective into account. This study gives an insight into the learner's perspective. This research might make the educational designers and LA researchers aware of the fact that learners could have a different preference than initially thought.

Conclusion

Where previous research focused on teacher preferences for LAD design, the current study used a user-centered approach and focused on the learner and reference frames. This study shows that learners do have a preference for a reference frame. The progress reference frame had the most win frequencies for all three task performances. However, one should be cautious when interpreting these conclusions. Indeed, the progress reference frame was only statistically different from the other three reference frames in the after performance task. With this research, we were able to gain more insight into learner preferences for presenting learning analytics. With these insights, we could contribute to the design of LADs in ILEs that learners appreciate and might potentially support them in regulating their learning.

Reference List

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271. https://doi.org/10.1037/0022-0663.84.3.261

Azevedo, R., & Gašević, D. (2019). Analyzing Multimodal Multichannel Data about Self-Regulated Learning with Advanced Learning Technologies: Issues and Challenges. *Computers in Human Behavior, 96*, 207–210. https://doi.org/10.1016/j.chb.2019.03.025

Bartholomew, S., Yoshikawa, E., & Connolly, P. (2018, June 18-21). *Exploring the potential for identifying student competencies in design education through adaptive comparative judgment.* In 36th International Pupils' Attitudes Towards Technology Conference (p. 187). https://www.skolfi.se/wp-content/uploads/2018/04/PATT36-Proceedings.pdf

Bartholomew, S. R., Ruesch, E. Y., Hartell, E., & Strimel, G. J. (2019). Identifying design values across countries through adaptive comparative judgment. *International Journal of Technology and Design Education*, *30*(2), 321–347. https://doi.org/10.1007/s10798-019-09506-8

Beckem, J. M. (2012). Bringing life to learning: Immersive experiential learning simulations for online and blended courses. *Journal of Asynchronous Learning Networks, 16*(5), 61-70.

Bramley, T. (2015). *Investigating the reliability of adaptive comparative judgment.* Cambridge Assessment.

Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-Centred Learning Analytics. *Journal of Learning Analytics*, *6*(2). https://doi.org/10.18608/jla.2019.62.1

Corrin, L., & De Barba, P. (2015, March). How do students interpret feedback delivered via dashboards?. In *Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 430-431).

Dede, C. (2009). Immersive interfaces for engagement and learning. *Science, 323*(5910), 66–69.

Dijkstra, P., Kuyper, H., van der Werf, G., Buunk, A. P., & van der Zee, Y. G. (2008). Social Comparison in the Classroom: A Review. *Review of Educational Research, 78*(4), 828–879. https://doi.org/10.3102/0034654308321210

Gibbs, G., & Simpson, C. (2005). Conditions under which assessment supports students' learning. *Learning and teaching in higher education,* (1), 3-31.

Herrington, J., Reeves, T. C., & Oliver, R. (2007). Immersive learning technologies: Realism and online authentic learning. *Journal of Computing in Higher Education, 19*(1), 80–99. https://doi.org/10.1007/bf03033421

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice. *Data-Driven Approaches in Digital Education*, 82–96. https://doi.org/10.1007/978-3-319-66610-5_7

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. *The Internet and Higher Education, 47*, 100758. https://doi.org/10.1016/j.iheduc.2020.100758

Jivet, I., Scheffel, M., Specht, M., & Drachsler, H. (2018). License to evaluate. *Proceedings of the 8th International Conference on Learning Analytics and Knowledge*, 31–40. https://doi.org/10.1145/3170358.3170421

Kim, D., Yoon, M., Jo, I.-H., & Branch, R. M. (2018). Learning analytics to support self-regulated learning in asynchronous online courses: A case study at a women's university in South Korea. *Computers & Education*, *127*, 233–251. https://doi.org/10.1016/j.compedu.2018.08.023

Kimbell, R., Wheeler, T., Stables, K., Sheppard, T., Martin, F., Davies, D., Pollitt, A., & Whitehouse, G. (2009). *E-scape portfolio assessment: Phase 3 report.* Goldsmiths' College. https://www.teachertoolkit.co.uk/wp-content/uploads/2014/08/e-scape_phase3_report.pdf

Laming, D. (2003). *Human judgment: The eye of the beholder*. Cengage Learning EMEA.

Major, B., Testa, M., & Blysma, W. H. (1991). *Responses to upward and downward social comparisons: The impact of esteem-relevance and perceived control.* In J. Suls & T. A. Wills (Eds.), *Social comparison: Contemporary theory and research* (p. 237–260). Lawrence Erlbaum Associates, Inc.

Margaryan, A., Littlejohn, A., & Milligan, C. (2013). Self-regulated learning in the workplace: strategies and factors in the attainment of learning goals. *International Journal of Training and Development*, *17*(4), 245–259. https://doi.org/10.1111/ijtd.12013

Matcha, W., Uzir, N. A., Gasevic, D., & Pardo, A. (2020). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *13*(2), 226–245. https://doi.org/10.1109/tlt.2019.2916802

Ochoa, X., & Wise, A. F. (2020). Supporting the shift to digital with student-centered learning analytics. *Educational Technology Research and Development*, *69*(1), 357–361. https://doi.org/10.1007/s11423-020-09882-2

Park, Y., & Jo, I. H. (2015). Development of the learning analytics dashboard to support students' learning performance. *Journal of Universal Computer Science*, *21*(1), 110.

Panadero, E. (2017). A Review of Self-regulated Learning: Six Models and Four Directions for Research. *Frontiers in Psychology*, *8*, 1–28. https://doi.org/10.3389/fpsyg.2017.00422

Panadero, E., & Alonso-Tapia, J. (2014). How do students self-regulate? Review of Zimmerman's cyclical model of self-regulated learning. *Anales de psicologia*, *30*(2), 450-462. http://doi.org/10.6018/analesps.30.2.167221

Pollitt, A. (2004, June). *Let's stop marking exams* [paper presentation]. International Association of Educational Assessment Conference, Philadelphia, PA, United States.

Pollitt, A. (2011). Comparative judgement for assessment. *International Journal of Technology and Design Education*, *22*(2), 157–170. https://doi.org/10.1007/s10798-011-9189-x

Pollitt, A. (2012). The method of Adaptive Comparative Judgement. *Assessment in Education: Principles, Policy & Practice*, *19*(3), 281–300. https://doi.org/10.1080/0969594x.2012.665354

Sedrakyan, G., Malmberg, J., Verbert, K., Järvelä, S., & Kirschner, P. A. (2020). Linking learning behavior analytics and learning science concepts: Designing a learning analytics dashboard for feedback to support learning regulation. *Computers in Human Behavior, 107*, 105512. https://doi.org/10.1016/j.chb.2018.05.004

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological Review*, *34*(4), 273–286. https://doi.org/10.1037/h0070288

Tynjälä, P. (2008). Perspectives into learning at the workplace. *Educational Research Review*, *3*(2), 130–154. https://doi.org/10.1016/j.edurev.2007.12.001

Viberg, O., Khalil, M., & Baars, M. (2020). Self-regulated learning and learning analytics in online learning environments. *Proceedings of the Tenth International Conference on Learning Analytics & Knowledge*, 524–533. https://doi.org/10.1145/3375462.3375483

Wang, M. (2017). *E-Learning in the Workplace*. Springer Publishing. https://doi.org/10.1007/978-3-319-64532-2

Wise, A. F. (2014). Designing pedagogical interventions to support student use of learning analytics. *Proceedings of the Fourth International Conference on Learning Analytics And Knowledge - LAK '14*, 203–211. https://doi.org/10.1145/2567574.2567588

Wise, A. F., Zhao, Y., & Hausknecht, S. N. (2014). Learning Analytics for Online Discussions: Embedded and Extracted Approaches. *Journal of Learning Analytics*, *1*(2), 48-71.

Zimmerman, B. J. (2000). Attaining Self-Regulation. *Handbook of Self-Regulation*, 13–39. https://doi.org/10.1016/b978-012109890-2/50031-7

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *The educational psychology series. Handbook of metacognition in education* (p. 299–315). Routledge/Taylor & Francis Group.