Self-regulated Learning as a Predictor for Learning Analytics Dashboard Reference Frame Preferences

Author:
Gerco Blom 6588352

Assessors:
1st: Prof. dr. Liesbeth Kester
2nd: Dr. Eva Janssen

Utrecht University

Date: 12-05-22

Universiteit Utrecht

Word count: 6597

Abstract

In Immersive Learning Environments (ILEs) the feedback tools are often passive displays which do not necessarily help learners to progress in their learning process. Learners could really benefit from receiving the appropriate feedback through Learning Analytics Dashboards (LADs) that supports them and enables them to take immediate action. This can, however, only happen if the feedback that is given is suitable for the learner based on their needs and skill level. Therefore this study aims to investigate how the Self-Regulated Learning (SRL) skills predict the preference for a certain type of LAD reference frame. Three different types of reference frames were used in this study: the Progress RF, the Social RF, and the Achievement RF. This research found that learners with higher SRL skills have a preference for the Progress RF. The findings can contribute to the educational sciences and the designing of LADs in ILEs in order to foster learning for learners of all levels.

Keywords: Immersive Learning Environment, Learning Analytics, Learning Analytic Dashboard, Self-regulated learning, Reference frame preferences

Introduction

Technology is changing rapidly in this digital era, and it is changing the way instructors teach and learners learn. The focus becomes more and more on user engagement so educators are trying to find new ways to incorporate interactive content in their lessons using e.g. online learning platforms or virtual reality (Allison et al., 2010). Immersive learning, learning through participation in a comprehensive and realistic digital environment, is becoming more advanced and suitable to have students engage in, learn, and transfer from classroom to real-world settings (Dede, 2009). For theorists and practitioners this requires knowledge about what instructional design is needed in order to enhance learning in an online environment. And even though the e-learning designs may be up to date and suitable to foster learning, the hardest challenge is to keep learners engaged throughout their learning process (Holley & Oliver, 2010).

Learning Analytics Dashboards (LADs) can help the learners to stay engaged, and the feedback provided by these dashboards can play an important role in the learners' ability to regulate their own learning process (Charleer et al., 2016). Unfortunately, the interface design of a LAD often provides insufficient meaningful data to foster learning in an Immersive Learning Environment (ILE) (Kitto et al., 2015). This is because the focus has often been too much on the design of the LAD, while the users of the LAD, the learners, were unable to decode, interpret, and make sense of the feedback shown (Verbert et al., 2013). Jivet et al. (2020) recognised this problem in their study, and they suggest that the focus should thus shift more towards the learners' needs instead of the fancy design. It should focus on how LADs can provide feedback that caters the needs of the learner to help them achieving their learning goals. Therefore, research is needed on what learners find relevant on a LAD. However, in their meta- analysis of LADs, Jivet et al. (2017) found that in reality, many current LAD designs lacked a proper scientific backing because oftentimes designers had chosen feedback reference frames they deemed fit, rather than using one that meets the learners' needs. Jivet et al. (2020) found that self-regulated learning (SRL) can be used as a useful theoretical foundation for LAD design, as SRL can be indicative for what kind of feedback learners require during their learning process (e.g. like goal setting, reflection, tracking progress, etc.). Because every learner regulates their learning process differently, it is necessary to consider the learners' SRL skills, and how these skills predict what type of feedback they require on a LAD (Jivet et al., 2020).

Thus, the goal of this study is to contribute to the field by providing insight in how SRL skills predict preference for a LAD design. The findings of this study will potentially be used to improve immersive learning environment learning analytics design.

Immersive learning

We are on the brink of entering a new era in which immersive learning plays a large role. Immersive learning is "the subjective impression that one is participating in a comprehensive, realistic experience" (Dede, 2009). According to Dede et al. (1999) this means that the learner is truly immersed in the learning environment through advanced design strategies that combine actional, symbolic, and sensory factors causing the disbelief that the learner is ‘inside’ a digitally enhanced setting. Furthermore, they say that the advantages of learning in an immersive setting are that one is able to learn through multiple perspectives, which helps the learner understanding complex phenomena. Also, immersive learning environments can foster situated learning for which authentic contexts, activities, and assessments are needed. Situated learning focuses on accomplishment of knowing in action and in practice, and should be integral to everyday practice in workplace or other settings (Handley et al., 2006). Creating such a complex real-world setting in the classroom is rather challenging but immersive interfaces are practical for situated learning by using digital simulations of authentic problem solving in a social interactive setting, as well as on different levels adaptive to the learner (Dede, 2009). These naturalistic interactions with the environment narrow the gap in learning transfer, however, proper feedback to the learner is key to offer most effective learning experience (De Freitas, 2010).

Designing Immersive Learning Environments, Workplace learning, and Learning Analytics

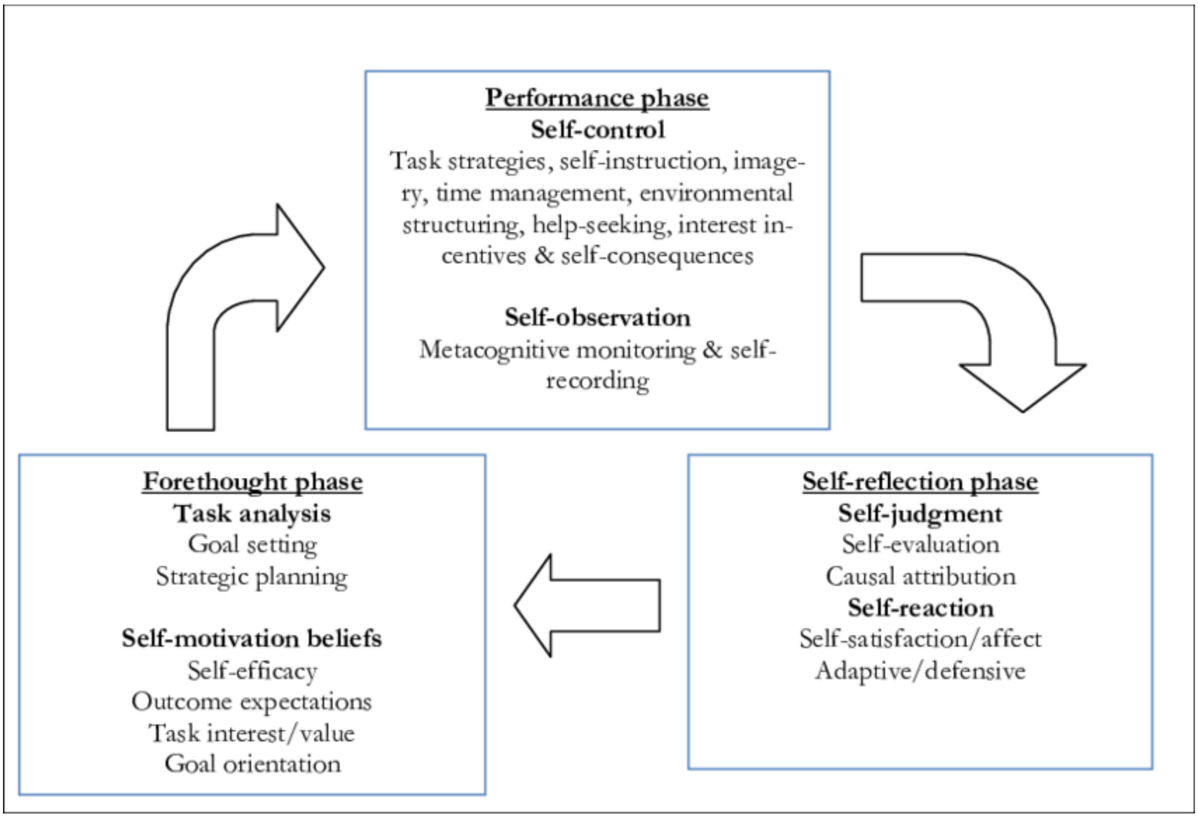
Immersive learning helps to gain a deep understanding of the required learning materials, which would normally take place in the real world (Martirosov & Kopecek, 2017; Salzman et al., 1999). According to Martirosov and Kopecek (2017) VR simulations prove to be effective not only for educational purposes but also in situations that would be costly or potentially dangerous in real life, if done by learners. Immersive learning environments (ILE) can be beneficial in workplace learning. The complexity of a task and lack of predictability of certain events within the task make it hard or costly to practice in real life e.g. fire fighters extinguishing fires in 10+ storey buildings or training of scientists in a laboratory setting with toxic substances, which are expensive or not always available. ILEs reduce the costs and remove the danger, risks, and problems with available locations almost entirely (Harman et al., 2017). ILEs have proven to be successful in providing different learning styles, collaborative knowledge building, and experiential learning (Mikropoulos & Natsis, 2011; Haj-Bolouri, 2020). Also, learning and performance can be enhanced before, during, and after the session which help learners to reflect on the training, and gain better understanding (Salas et al., 2012). This feedback is usually provided through LADs (Sedrakyan et al., 2020). LADs are in-game dashboards that provide information about the progress and the learning process. The information that they provide are based on collected data e.g. amount of clicks, spent time on assignment, scores compared to others or self, etc. However, little research has been done on the types of LADs grounded in learning sciences about how, when, and to who they are presented in an ILE (Haj-Bolouri, 2020; Jivet et al., 2017). Often, LADs have been designed without the aid of educational scientists, leaving learners with top-notch dashboards they were unable to interpret (Verbert et al., 2013). In order for LADs to be useful, it is important to know the learners’ preferences about what feedback they like to receive in an ILE (Jivet et al., 2020). Knowing what information to provide based on learners' preferences could certainly improve the usefulness and usability of LA interventions (Dollinger & Lodge, 2018).

SRL and Reference frames

Changing times and technological developments cause changes in requirements for jobs and puts an increased focus on life-long learning of workers (Dabbagh & Castaneda, 2020; Donovan & Benko, 2016). Life-long learning, the ongoing, voluntary, self-motivated pursuit of knowledge, flexibility, and contextual awareness, demands self-regulated learning (SRL) skills (Knowland & Thomas, 2014). SRL skills are the cognitive, metacognitive, emotional, motivational and behavioural aspects of learning (Panadero, 2017); or the ways in which individuals are able to regulate, control their cognition, understand, and adjust their own behaviour in an educational setting (Pintrich, 2000; Schraw et al., 2006; Zeidner & Stoeger, 2019). This requires transforming mental competencies into academic performance through using strategies and goal-setting (Zimmerman, Schunk, & DiBenedetto, 2015). Several different high quality models of SRL exist in the educational sciences, one of which is Zimmerman's social cognitive model of self-regulation, the SRL model used in this study (Puustinen & Pulkkinen, 2001; Panadero, 2017).

Figure 1

SRL Cycle with 3 Phases (Zimmerman & Moylan, 2009)



Zimmerman and Moylan (2009) distinguished three cyclical phases of SRL (Figure 1): the Forethought phase (FP), the Performance phase (PP), and the Self-reflection phase (SP). Each phase consists of integrated micro-processes belonging to that phase that help learners regulate their learning during that particular phase of the learning task e.g. goal-setting, strategic planning, monitoring, time management, and help-seeking (Zimmerman & Moylan, 2009; Zeidner & Stoeger, 2019). SRL skills are strongly linked to higher academic achievement (Broadbent & Poon, 2015) and learners with high SRL skills show different study behaviour like monitoring goal achievement, higher study motivation, and goal mastery (Kizilcec et al., 2017; Pintrich, 1999). This shows that SRL skills are extremely useful for jobs that require life-long learning, as these meta-skills will help the learner continually monitor and achieve their goals.

Learning Analytics Dashboards (LADs) come in different types but in order to design pedagogical interventions to support student use of learning analytics, learners need a representative reference frame to help them interpret their data (Wise, 2014; Jivet et al., 2017). Jivet et al. (2017) found that these reference frames are rooted in the theory that SRL skills inform the design of LADs. This means that awareness of and reflection on their learning process, as well as goal setting, planning, monitoring and self-evaluation are key elements in LADs. Jivet et al. distinguish three types of reference frames: social (SRF), achievement (ARF), and progress (PRF) in a meta-study on how LADs could be categorised. According to them the SRF focuses on comparison of performance level with other learners/peers in the current state. They found several studies in which learners could compare their data with a whole class or working groups, other learners with similar goals, and even with previous graduates of the same subject. The ARF focuses on comparison of their performance level with self-set goals also known as Internal Achievement (I-ARF), and goals set by the teacher, also known as External Achievement (E-ARF). The purpose of showing the results in relation to the goals was to illustrate mastery and skilfulness achievement. The PRF focuses on comparison of the learners' performance level over time with their historical data.

Jivet et al. (2020) found that learners with higher SRL skills are more inclined to use LADs because they see the relevance, they know how to use the information provided by the LAD, and their "learning goals can shape what reference frames they find relevant for their particular situation" (p.13). Novice SRL learners showed the opposite results, and were less inclined to use LADs because they lacked metacognitive abilities to recognise the potential of these reference frames. However, the study of Jivet et al. (2020) was done in a higher education environment using MOOCs. Furthermore, Jivet et al. (2017) stipulated that online learning environments that use LADs should provide additional tools to ensure that learners carry out all phases of the SRL cycle. Their meta-study found evidence that most LADs only supported the 'reflection and self-evaluation phase of SRL, and neglected others. Also Sedrakyan et al. (2020) suggested that the learning process can be positively influenced when LAD feedback provides information about regulatory mechanisms based on learner profiles e.g. planning, monitoring, and adapting activities so that learners detect inefficient processes in learning. However, no research has been done on how the SRL skills affect the reference frame preference in a ILE using VR during workplace training. Knowing what reference frame learners prefer based on their SRL score, could help us understand how LADs should provide feedback with regard to the learners' needs based on their self-regulatory abilities to foster improvement of monitoring and control processes in learning (Viberg et al, 2020).

Research Question

Because no research has been done in an Immersive Learning Environment on how the SRL skills predict the learners' preference for a Learning Analytics Dashboard reference frame, this study aims to answer the following research question: *How do the learners'* *subjectively measured SRL skills predict their reference frame preferences in an Immersive Learning Environment?*

Method

The Research

This research contributes to learning analytics literature and immersive learning literature by conducting exploratory comparative judgement and explanatory quantitative survey research. In-development learning analytics dashboards and fictional data for the research were used as stimulus for the exploratory part. Participants shared their preferences for the learning analytics feedback design in a Virtual Reality training application through adaptive comparative judgement (ACJ). Also, the self-regulated learning skills were examined to see how they impact the preferences for learning analytics feedback designs. Finally, semi-structured interviews were conducted to gain more insight into why, if so, one learning analytics dashboard design was preferred over the others. The results of the interviews were, however, not used for this present study.

Participants

The participants of this study were employees (n = 77) of Merck KGaA, a German science and technology company. All of the participants take part in Merck's European Training Network for Chemical Engineering Immersive Learning (ETN-CHARMING) project as trainees. They all have a German background with German as their native language, 53 of which were male, 20 female, and 4 chose not to share their gender. The average age was 22,5 (SD=3,15).

Instruments/materials

Demographic Data Questionnaire

The data that were used for this present study is the same data used in an earlier study. The demographic questionnaire was used to gather information about the participants' age, gender, and highest qualifications. In order to reduce cognitive overload or mistakes by the participants, the questionnaire was translated into German.

Adaptive Comparative Judgment

This study used the Adaptive Comparative Judgment (ACJ) method to identify the rank order of the participants' preferred learning analytics feedback designs for each instructional phase. Adaptive Comparative judgment, based on the Law of comparative judgment (Thurstone, 1927), is a method used to identify the preferred stimulus after comparing stimuli with one another, and applying a mathematical formulation to enable a rank order of the stimuli. For example, judging different kinds of packaging for a new product used for a marketing campaign by participants of a market research. In short, comparative judgment is a way to compare stimuli pairwise in order to get a perceived ranking. ACJ as an assessment tool has been proven to be useful when many judges take part in a judgment round, the ease of use is beneficial, and the potential for increased reliability levels (Bartholomew & Yoshikawa, 2018). Because ACJ can be easily be applied to evaluate various artifacts efficiently and effectively in many contexts, it has a high reliability and validity (Bisson et al., 2016).

For the ACJ, a comparing tool called RM Compare was used; a valid, reliable, and user-friendly instrument. The interface of RM Compare and the learning analytic feedback designs were in English but the questions that were asked to the participants were written in German. All the raw data needed for further analysis were provided by RM Compare, which included the individual results of the ACJ of each participant. RM Compare did not give reference frame preferences for each individual, just the result of each comparison made.

The amount of comparisons made by every participant depended on whether the preference for a reference frame was clear, though no more than 7 comparisons were made for each instructional phase. The designs were compared separately for each instructional phase, and were not mixed between the phases. Through the repeated comparisons the adaptive algorithm of RM Compare created a highly reliable scale or rank order of the preference for a learning analytics reference frames. After every frame had been judged once, the system generated new pairs were close in quality so that after several rounds of revision enough data had been collected to produce an accurate rank order (Pollitt, 2012; Bartholomew & Yoshikawa, 2018). Thus, despite not every pair was seen by every judge, a reliable total rank order could be created, saving time and effort of the participants.

This algorithm pairs items with similar rank scores based on previous judgments, making the results more reliable than normal comparative judgment, and immediately creates a rank order (Politt, 2012; Bramley, 2015; Wright & Masters, 1982). The results of the ACJ provided a single rank order of the learning analytics feedback design from the perspective of reference frames for the three instructional phases. Other data collected by RM Compare included duration of each answer, probability scores, misfit scores, etc. but they were not used in this present study as these data only say something about the single rank order, and this study only used the raw data of the reference frame preferences of each individual participant. These data were used for different studies that used the same dataset.

This study used learning analytics feedback dashboards that were designed for a chemical process industry virtual reality training simulator. The dashboards were designed by educational scientists and instructional designers from Utrecht University in collaboration with experts from the field of chemical engineering, who provided the content. In total 9 different dashboard mock-ups were used, representing three sub-categories: the before the task, during the task, and after the task. These sub-categories are the intended instructional phases in which the learning analytics dashboards will appear while operating the virtual reality training simulator. The before task dashboard is presented to the user before commencing with the simulator task, and it shows data on the user’s previous attempts. The during task dashboard is presented when a user pauses the simulator during the task, and this dashboard provides information on the user’s previous attempts as well as their present data. The after task dashboard is presented when the user has completed the task, and shows performance data on the user’s previous attempts as well as the attempt just completed.

The designs used all looked very similar to one another, except the reference frame that was used. This was done so that the participants would judge the designs based on their reference frames instead of their aesthetic design or other features. The four reference frames that were used are the progress-, the social-, the internal achievement-, and the external achievement reference frame. The progress reference frame compared the user's data with their own performance on previous attempts at the simulator. The social reference frame compared the user's data with data from their peers. The internal achievement reference frame compared the user's data with self-set achievement goals. The external achievement reference frame compared the user's data with the achievement goals set by a trainer.

Figure 2

Dashboard design with 3 different elements.

Each dashboard design consisted of three elements (Figure 2). Element 1 contextualised the learning analytics dashboard by showing what instructional phase they are in, highlighted the reference frame, and encouraged the participant to review the data displayed. Element 2 presented the user with the simulator task, which showed fictional overall score on the task as well as the scores for each phase, providing feedback on performance behaviours. Element 3 displayed the fictional data based on the time spent practicing in the simulator, which provides feedback related to self-regulated learning behaviours.

Self-regulated Online Learning – Questionnaire – Revised (SOL-Q-R)

The instrument used to gather data about the self-regulated learning skills was the SOL-Q-R questionnaire (Appendix A). The questions of the survey were applied to the context of Merck KGaG, and was also translated into German by professional translation agency to ensure that every participant fully understood the questionnaire. The questionnaire was then translated back to English by a third party for any mistranslation, and then back into German. The SOL-Q-R is a survey comprised of 42 items, that made the participants rate statements about their self-regulated learning behaviour in online educational contexts (Jansen, et al., 2018). Five different scales are measured: metacognitive skills (20 statements), time management (5), environmental structuring (4), persistence (6), and help seeking (6). The metacognitive skills scale consists of three subscales: metacognitive activities before (7/20 statements), during (7/20), and after a learning task (6/20). The questionnaire was presented to the participants with the items in randomised order, and for each item a 7-point Likert scale was used, ranging from “not at all true for me” (=1) to “very true for me” (=7). Jansen et al. (2018) made the SOL-Q-R by revising an earlier version of the questionnaire. The updated version showed a high validity, reliability, and usability, and is therefore considered a valuable tool to measure SRL in online education.

Procedure

One week before the data collection, the participants signed informed consent, and declared that the study was voluntary and that no form of extrinsic incentive was used to participate in this study. The participants could withdraw from participating any time they wanted without reason or any consequences. The participants were aware that all results would be anonymised and stored securely on the protected file store service. The entire data set was held by an independent researcher, who has the keys which identify the participants to match the different instruments. The participants were informed that the data were reused for different studies. Because of COVID-19 restrictions, the participants were given instructions about the research in groups of 20 by one of the researchers who leads the ETN-CHARMING project through an online webinar. Participants could sign up to one of the six webinars held in German, in which they were guided through the research procedure.

This is how the data collection was conducted. First, all participants received an email with a login code so they could take part in the study anonymously. Then they were asked to complete the demographic data questionnaire to gather information about their age, gender, and their education. These data were gathered using an online programme called Qualtrics.

The participants were then asked to comparatively judge which learning analytics dashboard design they preferred for each instructional phase using RM Compare. Each participant received a pre-set anonymised username and password to log in so that only the independent research could trace the results back to the participants. The participants in this research took part in three separate comparative judgment rounds, for each of the instructional phases. The participants were asked to comparatively judge between learning analytics dashboards in the following order: first for the 'before the task' phase, then for the 'during the task' phase, and finally for the 'after the task' phase. In total 21 comparisons were made by each participant, and the results were automatically saved in RM Compare

After this the SOL-Q-R was administered to all participants in an online environment. Each participant was asked to score themselves on the 42 statements about their self-regulated learning behaviour. The questions were presented to each participant in randomised order. Once submitted, the results were automatically saved and became part of the dataset. Finally, the data collection ended with a semi-structured interview in which the participants were asked to elaborate on their choices made in the ACJ and the SOL-Q-R. The data collection took about 20 minutes per participant.

Data analysis

The survey datasets of the demographic data collection and SOL-Q-R, as well as the ACJ dataset gathered by RM Compare were analysed by using the SPSS statistical software program to see how SRL skills predict their reference frame preferences in an Immersive Learning Environment.

Because RM Compare only provided reference frame preference for the entire population and not for each individual, the individual preferences per instructional phase had to be decided on before moving on to further analysis. Because RM Compare used Adaptive CJ, not all reference frames were necessarily compared as the algorithm pairs items with similar rank scores based on previous judgments. This could also cause participants to compare the same 2 reference frames twice or more. Next to the 4 reference frames, 2 more categories were created when a single preference could not be derived from the data: Contradictory Scoring and Missing Comparison. If no clear preference could be found because of contradicting scores, this was called ‘Contradictory Scoring’. If no clear preference could be found because two reference frames would be preferred equally, and when these two reference frames had not been compared, this was called ‘Missing Comparison’. Thus, there were 6 possible outcomes for reference frame preference.

A multinomial logistic regression (MLR) analysis was carried out to predict the Learning Analytics reference frame preferences based on the SRL scores. The data of the ACJ was used to distinguish the preferred reference frames for each participant for each instructional phase. From the data of the SOL-Q-R, the grand total of all the mean scores of the 5 SRL-scales for each participant was calculated to get individual SRL scores. Although the SOL-Q-R used Likert scales, which collect ordinal data, it is nevertheless common practice that these scores can be treated and analysed as continuous outcomes in a regression analysis (Norman, 2010).

All assumptions for a MLR were checked. Because there is an nominal variable involved, Tabachnick et al. (2007) argue that the regression does not require the normal assumptions of a regression analysis. They say that linear relationship between the dependent and independent variable is not needed. Further, the residuals do not need to be normally distributed and homoscedasticity is not required. Instead the regression analysis was checked on the following assumptions as described by Schreiber-Gregory (2018). In order to do a multinomial logistic regression analysis, the dependent variable needs to be ordinal or nominal. Also, the analysis was checked on whether observations were independent from each other. The observations in MRL should not come from repeated measures or matched data. (Starkweather & Moske, 2011) Further, the regression analysis was checked on multicollinearity among the independent variables to see if they were too highly correlated so Pearson's Bivariate Correlation matrix was used. Next, the logistic regression analysis was checked on linearity of independent variables and log odds. Although linearity is not required in MRL, the independent variables should be linearly related to the log odds. This can be done by inspecting a graph and testing whether the log odds are linear. Finally, the regression was checked on the sample size. A logistic regression analysis requires a minimum of 10 cases for the least frequent outcome of each independent variable (Schwab, 2002).

Originally, the data collected also investigated three phases from Zimmerman's cyclical model of SRL, but they were omitted in this study because of the limited requirements of an academic master thesis. Therefore, the data of the three phases were treated as if the phases were irrelevant, and they were put in one dataset. Each participant is therefore represented three times but treated as if they were different individuals, resulting in 192 scores on reference frame preference. This was done just to see if the SRL scores predicted for a Reference frame preference.

In this study, the SRL scores were the independent variables, and the dependent variable was the Learning Analytics reference frame preference (Progress, Social, Internal Achievement, External Achievement, Missing Comparison, Contradictory Scoring).

Results

Table 1 shows the 192 scores on the LAD reference frame preferences when the results of the three instructional phases are treated as one group.

Table 1

Descriptive Statistics for the Reference Frames Preferences

		N	Marginal Percentage
Preference	Contradictory scoring	54	28,1%
	External	43	22,4%
	Internal	14	7,3%
	Missing comparison	18	9,4%
	Progress	50	26,0%
	Social	13	6,8%
Valid		192	100,0%
Missing		0	
Total		192	
Subpopulation		64[a]	

a. The dependent variable has only one value observed in 11 (17,2%) subpopulation

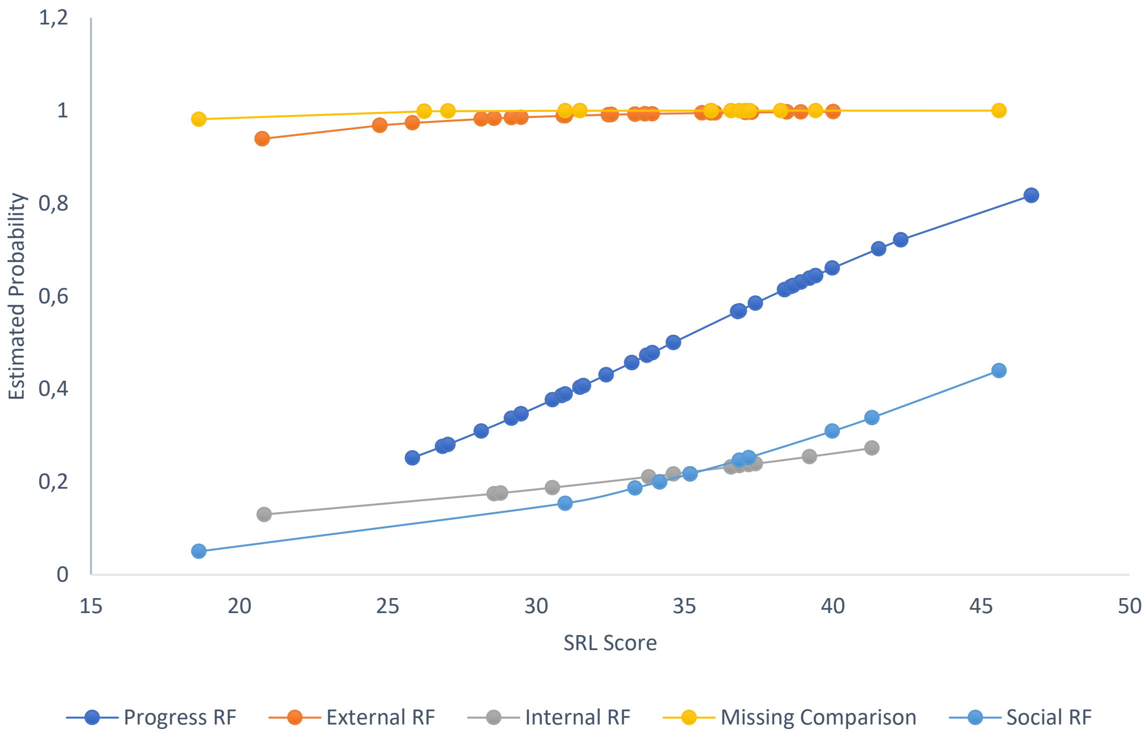
The three most preferred reference frames were Contradictory scoring (N=54), the Progress reference frame (N=50), and the External achievement reference frame (N=43). The subpopulation of *n* = 64 indicates that the population has something in common, and this is due to the fact that every SRL score is represented three times, as if they were different individuals.

A multinomial logistic regression analysis was conducted to assess how SRL score predict the preference for an LAD reference frame. The dependent variable (RF Preference) was nominal, meaning there were multiple dependent variables, and the independent variable (SRL score) was a continuous variable. The model was checked on whether the assumptions had been violated.

The assumption of observation independence was violated because the data of the same participants of three separate measurements were used, and treated as one dataset. No multicollinearity was found because the only independent variable was SRL skills, and more independent variables are needed for multicollinearity to occur. The linearity of independent variables and log odds was inspected visually, and they appeared to be linear (Figure 3). The predicted probabilities were calculated for each SRL score for each Reference frame preference manually and put in a graph. The sample size was sufficient with the least frequent outcome being the Internal Reference frame with (N=14).

Figure 3

Predicted Probabilities of Choosing a Preferred Reference Frame Versus SRL Scores with Contradictory scoring as a reference category



The validity of the final MRL model was inspected with the Odds Ratio Test and it was found statistically significant ($\chi 2$ = 13.581; p = .019) meaning that the full model predicts better than the null model. With a significant result the null hypothesis has to be rejected that there is no difference between the model with the predictor variables and without the predictor variables. This was inspected by looking at the *B* coefficients which take different values for each category (Table 2). Because Contradictory scoring was the category with the most results, it was used as the reference category in the MLR.

If the *OR* shows a value >1, it means that the independent variable has an effect on the dependent variable. It means that the higher the value of the independent variable, the more likely it is that a certain category is chosen. In this study the Odd ratios of all the

Reference frames were >1, however only the Progress RF showed a significant result (Wald's $\chi 2$ = 10.182; p = .001) whereas all the other Reference frames showed no significant results.

Table 2

Summary of Multinomial Regression analysis SRL as a predictor for a LAD reference frame preference

Variable	*B*	*SE*	*OR*	95% CI	Wald's $\chi 2$	*p*
External RF	.017	.036	1.018	[0.95, 1.09]	.232	.630
Internal RF	.045	.054	1.046	[0.94, 1.16]	.691	.406
MC RF	.033	.049	1.033	[0.94, 1.14]	.454	.500
Progress RF	.124	.039	1.132	[1,05, 1.22]	10.182	.001
Social RF	.100	.059	1.105	[0.98, 1.24]	2.848	.092

a. The reference category is: Contradictory scoring.
Note. CI = confidence interval for odds ratio (OR).

From these results it can be inferred that the higher the individual's SRL score, the more likely they are to prefer the Progress Reference frame as a way of getting feedback. This effect was not visible for the other Reference frames.

Discussion

This study aimed to explore how SRL skills predict the learners' preference for a Learning Analytics Dashboard reference frame. Knowing what kind of LAD Reference frame preference a learner has, based on their SRL skills, could help understand how feedback should be provided in immersive learning environments to foster improvement of monitoring and control processes in learning (Viberg et al., 2020). Therefore, this study investigated how learners' subjectively measured SRL skills predict the LAD reference frame preference of the learner in an immersive learning environment.

The research question was focused on the effect SRL skills would have on one of the three Reference frames (Progress RF, Social RF, Achievement RF) described by Jivet et al.

(2017). In this study the Achievement RF was divided up into two categories: Internal Achievement (Internal RF) and External Achievement (External RF). In addition to the four reference frames, two more were added: Missing Comparison and Contradictory Scoring. The latter two were added so that participants with unclear outcomes could also be used, although without a clear preference for a Reference frame. The results show that only the Progress RF showed a significant effect, meaning that the higher the SRL score of a learner, the more likely they are to prefer the Progress RF as LAD feedback. A possible explanation could be that because the Progress RF requires self-observation, learners can vary the use of task strategies and make adjustments based on the outcomes in order to make improvements (Boekaerts et al., 1999). A higher SRL score with a preference for the Progress RF may also suggest that one is more inclined to improve their own learning to master the task and develop competence (mastery goal-oriented), for which self-observation and self-referential task improvement is needed (Cerasoli & Ford, 2014).

Previous research shows that it is necessary to provide equal opportunities for both learners with higher and lower SRL skills (Jivet et al., 2020). According to Jivet et al. (2020) not catering for learners with lower SRL skills will usually result in widening the gap between learners with lower and higher SRL skills. Learners with lower SRL skills often lack the meta-cognitive abilities to recognise the potential of LADs, and are therefore less likely to use them properly. In contrast, learners with higher SRL skills are more inclined to use LADs, are transparent and use follow-up actions. The results of this current study may align with the findings of Jivet et al., as the participants with high SRL scores preferred a Reference frame that is linked to self-observation and possibly to mastery-goal orientation; a particular trait related to high SRL scores (Pintrich, 2000).

Limitations and Future Research

This present research has its limitations, which have to be taken into account before making generalisations. The first major limitation was the lack of independence that was violated in the multinomial logistic regression. For this recent study the phases of the SRL cycle were taken out, and the data for each phase was put into one dataset as if it had been one data collection with different participants. Therefore the SRL score of each participant was represented three times, albeit with different results if a participant had different preferences in different phases. This may have affected the outcomes of this study e.g. some SRL scores may be overrepresented for a reference frame driving up the average SRL score for that reference frame (Quinn & Keough, 2002). Or it may have caused effects that are not really there like the significant score for the Progress reference frame or not significant scores for the other reference frames because the triple used SRL scores distorted the actual effect.

Also, although the SOL-Q-R is considered a valid and reliable instrument (Jansen et al., 2018), it makes use of the measurements of self-reported SRL scores. This means that the participants may have been biased when they filled in the questionnaire e.g. socially desirable answers, not having understood the questions correctly, correlations between answers because prior questions elicited memories that affected other answers, and participants being wrong about their own attitudes and thus reporting something they believe is true which in fact is not (Bertrand & Mullainathan, 2001). The data of this present study must therefore be interpreted with caution because the findings may not be causal due to bias in participant self-report.

Next, because RM Compare only provided comparisons until there was a clear preference for the single rank score of the entire population, the individual preferences were not always clear. Had every participant compared every reference frame until one final preference for a reference frame was established, the categories ‘Missing Comparison’ and ‘Contradictory scoring’ of the dependent variable could have been omitted. These two categories together got 72 out of 192 scores (37,5% of total), and had these scores been preferences for one of the original reference frames, it may have affected the outcomes drastically.

Further, this study did not measure whether the participants could all make sense of what they were tested on because they e.g. lacked the ability to identify why they preferred one reference frame above the other. This could be a reason why there were so many participants without a clear preference for a reference frame. Finally, the present study used fictional data, which may not be quite as good a representation as an actual ILE.

Further research is needed to get a better understanding how SRL skills predict the preference of a LAD reference frame so that learners receive the best feedback possible in an ILE based on their needs (Jivet et al., 2020). In a future study a Mixed Model should be used to identify whether there are also differences between the different phases. A multilevel multinomial logistic regression analysis could shed a better light on how the LAD preferences differ between the phases for the SRL scores. In addition, a larger sample is needed for the generalisation of the finding. Also, a ready designed ILE should be used to gather data and have the participants make comparisons while actually immersed in the learning environment, rather than using screenshots of an ILE. This is costly but it might give more valid results as participants who could not make sense of what was investigated and what it was for in the present study, will be able to give better judgments of their preference.

Implications

Previous research has discussed the need for well-designed LA tools in order to foster the learning process in ILEs (Jivet et al., 2020). The role of SRL was found to be important in ILEs to help learners of all levels come to their potential through the use of LADs (Jivet et al., 2017). To our knowledge, the present study was the first to explore the role of SRL skills as a predictor for a LAD Reference frame preference. This study found that learners with a higher SRL score are more likely to prefer the Progress RF.

This study can help in the development of LADs, and can inform designers and policymakers how to design LADs for all levels of learners to help to come to their full potential and to bridge the gap between the learning levels in terms of SRL scores. Because ILEs can truly help learners to with lifelong learning (Dabbagh & Castaneda, 2020) it is advised that more research should be done on LADs as they are helpful. The focus of the feedback should, however, be on the learners and their capabilities instead of what designers might think is suitable for them; often there is a strong emphasis on comparison with peers (Jivet et al., 2017) A grounded theory is therefore needed, and this research contributes to how the self-regulatory skills should play an important part in designing feedback in ILEs.

Conclusion

This study found a significant effect of SRL skill as a predictor for a LAD Reference frame preference. The higher the SRL score the more likely the learner is to prefer the Progress RF in an ILE as feedback frame but no effect was found for any of the other Reference frames. One should be careful to make any generalisations based on this study as it has certain limitations. However, this research does contribute to the field of educational sciences as it gained more insight in the preferences for LAD Reference frames and how the learners needs and skill level play a role in immersive learning. Hopefully, this research will help LAD designers improve the LA feedback tools in ILEs from what were once passive displays to adaptive feedback tools that integrate support and foster learning.

References

Allison, C., Miller, A., Sturgeon, T., Nicoll, J. R., & Perera, I. (2010, October). Educationally enhanced virtual worlds. In *2010 IEEE Frontiers in Education Conference (FIE)* (pp. T4F-1). IEEE.

Bartholomew, S. R., & Yoshikawa, E. (2018). *A systematic review of research around Adaptive Comparative Judgment (ACJ) in K-16 education.* 2018 CTETE Monograph Series. https://doi.org/10.21061/ctete-rms.v1.c.1.

Bisson, M. J., Gilmore, C., Inglis, M., & Jones, I. (2016). Measuring conceptual understanding using comparative judgement. *International Journal of Research in Undergraduate Mathematics Education, 2*(2), 141-164.

Boekaerts, M., Zeidner, M., & Pintrich, P. R. (Eds.). (1999). *Handbook of self-regulation.* Elsevier.

Bramley, T. (2015). Investigating the reliability of adaptive comparative judgment. *Cambridge Assessment research report, Cambridge, 36.*

Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement inonline higher education learning environments: A systematic review. *The Internet and Higher Education,* 27, 1-13.

Cerasoli, C. P., & Ford, M. T. (2014). Intrinsic motivation, performance, and the mediating role of mastery goal orientation: A test of self-determination theory. *The Journal of psychology, 148*(3), 267-286.

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

Dabbagh, N., & Castaneda, L. (2020). The PLE as a framework for developing agency in lifelong learning. *Educational Technology Research and Development, 68*(6), 3041-3055.

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

De Freitas, S., Rebolledo-Mendez, G., Liarokapis, F., Magoulas, G., & Poulovassilis, A. (2010). Learning as immersive experiences: Using the four-dimensional framework for designing and evaluating immersive learning experiences in a virtual world. *British Journal of Educational Technology*, *41*(1), 69-85.

Dollinger, M., & Lodge, J. M. (2018, March). Co-creation strategies for learning analytics. In *Proceedings of the 8th international conference on learning analytics and knowledge* (pp. 97-101).

Donovan, J., & Benko, C. (2016). AT&T's talent overhaul. *Harvard Business Review*, 10, 68-73.

Jansen, R. S., Van Leeuwen, A., Janssen, J., & Kester, L. (2018, September). Validation of the revised self-regulated online learning questionnaire. *In European Conference on Technology Enhanced Learning* (pp. 116-121). Springer, Cham.

Haj-Bolouri, A., Winman, T., & Svensson, L. (2020, December). Meta-requirements for Immersive Collaborative Spaces in Industrial Workplace Learning: Towards a Design Theory. In *International Conference on Design Science Research in Information Systems and Technology* (pp. 339-346). Springer, Cham.

Handley, K., Clark, T., Fincham, R., & Sturdy, A. (2006). *Researching Situated Learning: Participation, Identity and Practices in Management Consultancy*. Durham, UK: Durham University.

Harman, J., Brown, R., & Johnson, D. (2017, September). Improved memory elicitation in virtual reality: new experimental results and insights. In *IFIP Conference on human-computer interaction* (pp. 128-146). Springer, Cham.

Holley, D., & Oliver, M. (2010). Student engagement and blended learning: Portraits of risk. *Computers & Education, 54*(3), 693-700.

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017, September). Awareness is not enough: Pitfalls of learning analytics dashboards in the educational practice. In *European conference on technology enhanced learning* (pp. 82-96). Springer, Cham.

Jivet, I., 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.

Kitto, K., Cross, S., Waters, Z., & Lupton, M. (2015, March). Learning analytics beyond the

LMS: the connected learning analytics toolkit. *In Proceedings of the fifth international conference on learning analytics and knowledge* (pp. 11-15).

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & education*, 104, 18-33.

Knowland, V. C., & Thomas, M. S. (2014). Educating the adult brain: How the neuroscience of learning can inform educational policy. *International Review of Education, 60*(1), 99-122.

Martirosov, S., & Kopecek, P. (2017). Virtual reality and its influence on training and education literature review. *Annals of DAAAM & Proceedings*, 28.

Mikropoulos, T. A., & Natsis, A. (2011). Educational virtual environments: A ten-year review of empirical research (1999–2009). *Computers & Education, 56(3),* pp.769–780. DOI: 10.1016/j.compedu.2010.10.020

Norman, G. (2010). Likert scales, levels of measurement and the ‘‘laws’’ of statistics. *Adv Health Sci EducTheory Pract., 15*(5):625–632

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in psychology*, *8*, 422.

Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International journal of educational research, 31*(6), 459-470.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In *Handbook of* *self-regulation* (pp. 451-502). Academic Press.

Pollitt, A. (2012). The method of adaptive comparative judgement. *Assessment in Education: Principles, Policy & Practice, 19*(3), 281–300.

Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269-286.

Quinn, GP., & Keough, MJ. (2002). *Experimental Design and Data Analysis for Biologists.* Cambridge, UK: Cambridge University Press.

Salas, E., Tannenbaum, S. I., Kraiger, K., & Smith-Jentsch, K. A. (2012). The science of training and development in organizations: What matters in practice. *Psychological Science in the Public Interest, 13*(2), pp.74–101. DOI: 10.1177/1529100612436661

Salzman, M. C., Dede, C., & Loftin, R. B. (1999, May). VR's frames of reference: A visualization technique for mastering abstract multidimensional information. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems* (pp. 489-495).

Schraw, G., Crippen, K. J., & Hartley, K. (2006). Promoting self-regulation in science education: Metacognition as part of a broader perspective on learning. *Research in science education, 36*(1), 111-139.

Schwab, J. A. (2002). *Multinomial logistic regression: Basic relationships and complete problems*. http://www.utexas.edu/courses/schwab/sw388r7/SolvingProblems/

Schreiber-Gregory, D. N. (2018). Ridge Regression and multicollinearity: An in-depth review. *Model Assisted Statistics and Applications, 13*(4), 359-365.

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.

Starkweather, J., & Moske, A.K., (2011) *Multinomial logistic regression.*

Tabachnick, B. G., Fidell, L. S., & Ullman, J. B. (2007). *Using multivariate statistic*s (Vol. 5, pp. 481-498). Boston, MA: Pearson.

Thurstone, L. L. (1927). A law of comparative judgment. *Psychological review, 34*(4), 273 286.

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning analytics dashboard applications. *American Behavioral Scientist*, 57(10), 1500-1509.

Viberg, O., Khalil, M., & Baars, M. (2020, March). Self-regulated learning and learning analytics in online learning environments: a review of empirical research. In *Proceedings of the tenth international conference on learning analytics & knowledge* (pp. 524-533).

Wright, B. & Masters, G. (1982). *Rating scale analysis: Rasch measurement* (1st ed. Ed.) Chicago, IL: MESA Press.

Zeidner, M., & Stoeger, H. (2019). Self-Regulated Learning (SRL): A guide for the perplexed. *High Ability Studies, 30*(1-2), 9-51.

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In *Handbook of metacognition in education* (pp. 311-328). Routledge.

Zimmerman, B. J., Schunk, D. H., & DiBenedetto, M. K. (2015). A personal agency view of self-regulated learning. *Self-concept, motivation and identity: Underpinning success with research and practice*, 83-114.

Appendix A: Self-regulated Online Learning Questionnaire Revised (SOL-Q-R)

Metacognitive activities before learning

1. I think about what I really need to learn before I begin a task in this online course.
2. I ask myself questions about what I am to study before I begin to learn for this online course.
3. I set short-term (daily or weekly) goals as well as long-term goals (monthly or for the whole online course).
4. I set goals to help me manage my studying time for this online course.
5. I set specific goals before I begin a task in this online course.
6. I think of alternative ways to solve a problem and choose the best one in this online course.
7. At the start of a task I think about the study strategies I will use in this online course.

Metacognitive activities during learning

8. When I study for this online course I try to use strategies that have worked in the past.
9. I have a specific purpose for each strategy I use in this online course.
10. I am aware of what strategies I use when I study for this online course.
11. I change strategies when I do not make progress while learning for this online course.
12. I periodically review to help me understand important relationships in this online course.
13. I find myself pausing regularly to check my comprehension of this online course.
14. I ask myself questions about how well I am doing while learning something in this online course.

Metacognitive activities after learning

15. I think about what I have learned after I finish working on this online course.
16. I ask myself how well I accomplished my goals once I'm finished working on this online course.
17. After studying for this online course I reflect on what I have learned.
18. I find myself analyzing the usefulness of strategies after I studied for this online course.
19. I ask myself if there were other ways to do things after I finish learning for this online course.
20. After learning for this online course, I think about study strategies I used.

Time management

21. I make good use of my study time for this online course.
22. I find it hard to stick to a study schedule for this online course.

23. I make sure I keep up with the weekly readings and assignments for this online course.
24. I often find that I don’t spend very much time on this online course because of other activities.
25. I allocate studying time for this online course.

Environmental structuring

26. I choose the location where I study for this online course to avoid too much distraction.
27. I find a comfortable place to study for this online course.
28. I know where I can study most efficiently for this online course.
29. I have a regular place set aside for studying in this online course.

Persistence

30. When I am feeling bored studying for this online course, I force myself to pay attention.
31. When my mind begins to wander during a learning session for this online course, I make a
special effort to keep concentrating.
32. When I begin to lose interest for this online course, I push myself even further.
33. I work hard to do well in this online course even if I don’t like what I have to do.
34. Even when materials in this online course are dull and uninteresting, I manage to keep working until I finish.
35. Even when I feel lazy or bored when I study for this online course, I finish what I planned to
do.
36. When work is difficult in this online course, I continue to keep working.

Help seeking

37. When I do not fully understand something, I ask other course members in this online course
for ideas.
38. I share my problems with my classmates in this course online so we know what we are
struggling with and how to solve our problems.
39. I am persistent in getting help from the instructor of this online course.
40. When I am not sure about some material in this online course, I check with other people.
41. I communicate with my classmates to find out how I am doing in this online course.
42. When I have trouble learning, I ask for help.

Items are answered on a 7-point Likert scale, ranging from ‘‘not at all true for me’’ (= 1) to ‘‘very true for me’’ (= 7). All items are presented in randomized order.

Appendix B: Assignment 4

This section is a reflection on the possible issues, risks, and/or dilemmas of this study.

Sample characteristics and consent procedures

Because this study is exploratory it will be hard to make generalisations but moreover, because the sample consists only of around 30 people. Due to a lack of available contexts in which immersive learning in combination with SRL can be studied, this is the best we can do for now. Hopefully this study can function as a trigger to investigate the relation between SRL and ILE more, and can it contribute to the LA reference frame design. Again, this study will not provide generalisable results for large populations but it can function as a nice start in this field of research.

The consent procedures were all taken care of by an independent party. The participants are voluntarily taking part in this study, and there was no extrinsic incentive. This study re-uses the data of a bigger study in this field, and the participants are aware of this.

Choice of instruments and possibly sensitive questions

This study uses comparative judgement to find the reference frame preferences of the participants, and for the SRL skills, the participants use the SOL-Q-R questionnaire. Both instruments ask about either preferences or the perceived score on a statement. The former does not have any sensitive questions, and neither does the latter because it only asks the participants to score themselves on their SRL skills. There is a possibility the participants give socially desirable answers, although they are instructed not to before the commencing. A reason could be that employees feel that their results could be used against them at a later moment e.g. when evaluating how well they function in the company. They are pointed out that this is an exploratory study, and there is no good or bad, and that their answers will benefit their own learning experience. The participants will also be assured that all results are anonymised so they cannot be traced back to the participant.

Effort required

Very little effort is required from the participants. The participants take part in all tests online, and it will not require more than 30 minutes. Because of COVID-19, the interviews that will be held for the qualitative part of the original study, will not happen in a face-to-face setting but using Teams/Zoom. However, the qualitative data are not part of this present study but it is part of the data collection. The results can be used to improve the training at their company so this outweighs the effort of taking part in two little tests by far. Also, the materials have all been translated into the native language of the participants to get more reliable results, and to make it easier for the participants.

Data handling and storage

All the data is collected in an online environment. For the comparative judgement Comproved is used, and the SOL-Q-R will be conducted through another online survey. The data will be collected by an independent research data collector. The researchers will get an anonymised dataset for analyses.

Appendix C: Timetable of the research.

	Use feedback on Research plan to improve the Introduction and the Method section	
17 February	Improving Introduction	Kick-off meeting: describe RQ, topics, interview questions
24 February	Improving Introduction	Ask peer-feedback about analysis, coding and conclusion concerning theoretical sampling
3 March	Improving Introduction	
10 March		Roundtable: describe coding and sampling decisions, ask for feedback
17 March		
24 March	Hopefully the data will be available.	Ask peer-feedback on coding, sampling, conceptual model
31 March	Analysing data	
7 April	Analysing data	Prepare a focused issue for the supervisor feedback
14 April	Analysing data	
21 April	Results section	Assignment 4: peer-feedback on how well are the findings illustrated with examples and citations of the interviews, are the conclusions logic and is the line of reasoning well described
28 April	Results section + Discussion Prepare questions for the supervision meeting	
5 May	Discussion section	
12 May	Discussion section	Ask peer-feedback on the outline of the draft version
17 May	Submit Draft	
26 May	Discuss weaknesses of the draft and make adjustments	Supervision meeting

2 June	Prepare presentation	Practice presentations with peers, discuss improvements of the draft version
7 June	Submit Final Master’s thesis	
9 June	Prepare presentation	
16 June		Master’s thesis conference