Designing Learning Analytics Dashboards for Digital Learning Environments: Investigating Learner Preferences, Usage and Self-Efficacy

Timothy Gallagher

Designing Learning Analytics Dashboards for Digital Learning Environments: Investigating Learner Preferences, Usage and Self-Efficacy

Timothy Gallagher

ISBN: 978-94-6469-839-8

DOI: 10.33540/2227

Printed by: ProefschriftMaken

Design and Layout:

Designing Learning Analytics Dashboards for Digital Learning Environments: Investigating Learner Preferences, Usage, and Self-Efficacy

Het ontwerpen van Learning Analytics-dashboards voor digitale leeromgevingen: onderzoek naar voorkeuren, gebruik en self-efficacy van lerenden
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor aan de
Universiteit Utrecht
op gezag van de
rector magnificus, prof. dr. H.R.B.M. Kummeling,
ingevolge het besluit van het college voor promoties
in het openbaar te verdedigen op

vrijdag 12 april 2024 des middags te 4.15 uur

door

Timothy Gallagher

geboren op 3 juli 1986
te Nambour, Australia

Promotoren:

Prof. dr. L. Kester

Prof. dr. M.F. van der Schaaf

Copromotor:

Dr. B. Slof

Beoordelingscommissie:

Prof. dr. J. Glassey

Prof. dr. T. Van Gerven

Prof. mr. dr. U.R.M.T. de Vries

Prof. dr. R.C. Veltkamp

Prof. dr. T.A.J.M. van Gog

Dit proefschrift werd (mede) mogelijk gemaakt met financiële steun van het Europese programma Horizon 2020 voor onderzoek en innovatie (No. 812716).

Contents

Preface and acknowledgements

The conclusion of this five-year journey would not have been possible without the support and guidance of many individuals. First, I extend my thanks to Bert Slof, whose insightful advice on a range of matters, from research methods to career pathways, has been endlessly helpful. I am equally indebted to Marieke van der Schaaf, whose encouragement to forge connections within the medical educational research community and further, not only enriched my PhD journey but also opened doors to a myriad of research opportunities. I am also incredibly grateful to Liesbeth Kester. Having a professor with such extensive expertise and genuine enthusiasm for a diverse array of subjects as a daily supervisor is an opportunity few doctoral students receive. Our weekly dialogues were not merely supervisory check-ins, but conversations from which I always emerged with new knowledge or a fresh perspective. Be that about the educational sciences, the challenges and motivating aspects of running a department, parenting, or new insights about the Netherlands.

My gratitude extends to the entire education department at UU. The camaraderie shared, even in the remote working era during and after COVID, was a valued aspect of this journey. A special thanks go out to my CHARMING UU companions, Michaela and Yuan Yuan, for their constant support. The entire CHARMING Project team has enriched this journey, offering a rich diversity of perspectives, experiences, and cultures that have broadened my horizons immeasurably.

I am grateful for the collaborative opportunity with Merck, which allowed me to engage in important research alongside a leading science and technology firm.

I am also grateful to those who I was able to collaborate with at University Medical Centre Utrecht, although our plans for collaboration were interrupted by the pandemic, the collaboration we did do together was very valuable for my own development.

I would like to thank all the CHARMING PhD researchers with whom I collaborated with, especially those in Work Package 3, Sofia, Ryo, Yusra and Philippe.

Lastly, I want to thank my family. Pauline, for her faith in me and endless support. She has seen me through a lot of tough moments. Helping me deal with failures and encouraging me to celebrate my successes. My children, Lara and Sam, have also been a major source of joy and motivation. Whenever I was getting too caught up in my work, I was able to shift my focus to them which quickly put everything into perspective. I owe a great debt to my parents, who instilled in me a love for science and history. Their nurturing of my curiosity has significantly influenced the chapters that follow.

CHAPTER 1

1

General Introduction

Chapter 1 General Introduction

Today's global society is undergoing unprecedented digital transformation that is reshaping our economy, thereby influencing our workplaces and educational institutions (Schmidt & Tang, 2020). This transformation necessitates more effective, efficient, and engaging learning solutions to navigate the complexities of an increasingly interconnected world. To address the challenges that this transformation brings, digital learning environments have emerged, offering innovative methods to acquire knowledge and skills crucial for individual and societal advancement. However, this shift is not just about technology; it represents a fundamental change in our perception and approach to learning. It also underscores the need for continuous lifelong learning, tailored to accommodate the swift adaptations and demands of the modern world and workplace (Castro, 2019; Ifenthaler, 2018). In navigating the evolving digital educational landscape, certain tools and methodologies stand out as essential for harnessing its potential.

Central to the digital transformation of education and training is the application of learning analytics, which can be used to analyse educational data to understand and optimise learning and the environments in which it occurs (Siemens & Gasevic, 2012). When integrated into digital learning environments, these analytics can pave the way for more adaptive, personalised learning experiences. A foundational element within the scope of learning analytics is the learning analytics dashboard (LAD) (Verbert et al., 2013). LADs are tools that aggregate various indicators about learners, learning processes, and contexts, into visual representations (Schwendimann et al., 2017). They provide stakeholders with a comprehensive view of educational data, facilitating informed decision-making.

Digital learning environments, equipped with LADs, have potential to foster adaptive learning pathways and elevate the overall effectiveness of education and training (Schmidt & Tang, 2020). In these environments, features like automatic feedback can be integrated, providing the possibility for feedback at any time and any place. After all, in these digitally enhanced settings, learners possess greater control over when and where they engage in learning experiences, setting them apart from those delivered in non-digital environments.

The European Union's significant investment in the European Training Network for Chemical Engineering Immersive Learning (CHARMING) project, including a focus on the presentation of learning analytics for lifelong learning, for instance, stands as a testament to the societal imperative of leveraging the affordances digital learning offers. This momentum towards embracing the digital evolution in industry training and education is mirrored in the emerging specialised teams within higher education institutions. A case in point is the dedicated teams focusing on the conceptualisation and deployment of learning analytics systems in tertiary education, such as the Learning Analytics Team at Utrecht University. This trend underscores the broader movement towards integrating technology and data-driven approaches into the academic landscape.

As learning analytics systems gain prominence in education and training, there is an escalating need to address the nuances of their design (Jivet et al., 2017; Matcha et al., 2020). Currently, in specific educational settings (Garcia Fracaro et al., 2021), there is a gap in understanding both user perceptions of different LAD designs and how these designs separately influence learner self-efficacy and dashboard use (Gasevic et al., 2015; Ruiz-Calleja et al., 2017). In the following, this dissertation elaborates on this knowledge gap and offers four interconnected studies. By doing so it addresses the pivotal research question: How does LAD design influence learner preferences, interaction, and self-efficacy in training and education? Through these studies, empirical evidence is collected and presented. In addition to informing dashboard design practices, the data generated from these studies is expressly aimed at enhancing educational and training experiences for learners through optimised Learning Analytics Dashboards. Furthermore, it provides critical insights to other stakeholders such as teachers, administrators, and educational scientists, contributing to a more holistic understanding of how LAD design affects key factors like self-efficacy and dashboard interaction.

Building on the problem statement, this dissertation introduces a theoretical framework that not only sits at the intersection of technology and education but also directly guides the subsequent empirical studies. Central to this framework is the examination of the interplay between LAD design, self-regulated learning, self-efficacy, social comparison theory, temporal comparison theory, goal origin theory and achievement goal orientation theory. By situating this examination within the context of LADs for digital learning environments, particular emphasis is given to the settings of workplace learning and higher education. The overarching theoretical framework advances broader strategies for LAD design, reinforcing the promise of learning analytics in refining and improving education and training.

1.1 Learning analytics

While the primary role of learning analytics revolves around collecting, measuring and analysing educational data, the fundamental value lies in deriving actionable insights from this data (Susnjak et al., 2022). As learners engage with digital platforms, the data captured can inform more personalised learning pathways, tailoring instruction to individual needs (Gasevic et al., 2015). With the growing prominence of learning analytics tools, especially LADs, there is an emergent emphasis on refining their design to maximise their impact in both educational and workplace settings (Matcha et al., 2020).

In this evolving landscape, learner-facing LADs assume a particularly pivotal role (Valle, Antonenko, Dawson, et al., 2021). Specifically designed to convert learning analytics data into intelligible insights, these dashboards serve the immediate needs of learners as primary stakeholders (Farahmand et al., 2020). They synthesise data-derived insights about educational performance and behaviour, presenting them in an accessible, visual format that allows for actionable interpretation.

The complexity of designing effective learner-facing LADs stems from a multidisciplinary relationship that integrates insights from educational theory, data science, and human-computer interaction (Greller & Drachsler, 2012). This multidisciplinary approach is crucial for addressing the varied considerations that underpin the functionality and pedagogical effectiveness of the dashboards (Marzouk et al., 2016). This dissertation is fundamentally grounded in the field of educational science and focuses on a pedagogically informed analysis of learner facing LADs.

Within this complex design landscape, 'learning analytics reference frames' emerge as indispensable features of learner facing LADs. Defined as the comparison points that guide learners in interpreting their learning analytics data, these frames provide the contextual backdrop against which learners evaluate their performance (Wise, 2014).

1.2 An Operational Framework for Learning analytics reference frames

This dissertation introduces an Operational Framework for Learning Analytics Reference Frames, aimed at the systematic deconstruction of reference frames into discernible components. These components—namely performance outcome, point of comparison, and score delta—enable a focused exploration of their differential impact on key learning-related variables.

The first component, performance outcome, signifies the extent to which a learner has successfully executed a given task and received corresponding feedback. The second component, the point of comparison, serves to contextualise this performance outcome, thus, assisting learners in evaluating their achievements. Lastly, the component termed 'score delta' quantifies the difference between the performance outcome and the point of comparison. This score delta can possess negative, neutral, or positive values, thereby indicating directions of comparison as upward, lateral, or downward respectively.

The operational framework as presented in this dissertation references four types of reference frames—progress, social, internal achievement, and external achievement. Each type is differentiated by its unique point of comparison.

Progress reference frame

The progress reference frame uses historical performance data as the point of comparison, enabling learners to assess their progress over time.

Social reference frames

The social reference frame incorporates the performance of others as the point of comparison, enabling learners to compare their own performance with that of their peers.

Internal achievement reference frame

The internal achievement reference frame uses a self-set goal as the point of comparison, enabling learners to compare their own performance relative to their personally established benchmarks.

External achievement reference frame

The external achievement reference frame employs an assigned goal as the point of comparison, enabling learners to compare their own performance relative to benchmarks established by their trainer or teacher.

Though the importance of reference frames in shaping learning processes is acknowledged (Davis et al., 2017; Lim et al., 2019; Wise, 2014), a gap exists in the current body of research regarding the impact of different dashboard designs—especially those varying in reference frames—on key learning-related variables. Addressing this gap necessitates an examination grounded in educational theory. Such an inquiry aims to gather evidence to help stakeholders determine how reference frames can be optimally integrated into dashboard designs to align with pedagogical objectives, thereby contributing to the optimisation of these educational tools.

The theoretical frameworks employed in this dissertation fulfill multiple roles. They act as a conceptual scaffold that informs various empirical investigations: these delve into learner preferences concerning specific types of reference frames and probe how different reference frames influence learner self-efficacy beliefs and their interactions with LADs. Additionally, these frameworks provide the foundation upon which the study's hypotheses are constructed and offer a cohesive analytical perspective for interpreting the collected empirical evidence.

Having described the role of the dissertation theoretical framework, the next step is to delve into the specific theories that inform it, starting with the Zimmerman's Cyclical Phases model of SRL (Zimmerman, 2002).

1.2.1 Self-regulated learning

Self-Regulated Learning (SRL) is defined as the control that learners exercise over their cognition, behaviour, emotions, and motivation, utilising personal strategies to achieve established goals (Panadero & Alonso-Tapia, 2014). In this dissertation, Zimmerman's Cyclical Phases model of SRL (Zimmerman, 2002, 2013) serves two main purposes. First, it informs the content and timing of the LADs, ensuring that these interventions are well-aligned with the specific phases of learning. Second, the model provides a detailed account of the learning processes learners undergo before, during, and after task performance. This understanding is instrumental for exploring how different types of reference frames could potentially influence learners' preferences for different LADs, their self-efficacy and their interactions.

A point of emphasis within Zimmerman's model in this dissertation is the self-reflection phase, particularly the self-evaluation sub-processes (Zimmerman, 2015). During this phase, learners assess their performance against various criteria for success, such as mastery, past performance, and normative criteria (Zimmerman & Campillo, 2003). This phase is pivotal in this dissertation because different learning analytics reference frames offer these distinct criteria. Furthermore, it is during this self-reflection phase that learners accumulate evidence to inform their self-efficacy beliefs. Significantly, learners draw upon these very criteria—mastery, past performance, and normative criteria—to formulate their self-efficacy beliefs (Bandura, 2010). This relationship will be further elaborated upon in the subsequent section on Self-Efficacy.

1.2.2 Self-efficacy

Building on Zimmerman's theory of SRL, the construct of self-efficacy emerges as another focal point of this dissertation. Self-Efficacy (Bandura, 1997) serves as a primary outcome variable of interest. The motivation to focus on self-efficacy lies in its established importance across diverse settings, including the workplace and academia (Honicke et al., 2023; Sadri & Robertson, 1993; Stajkovic & Luthans, 1998). For example, self-efficacy has been shown to be a strong correlate for academic and workplace performance (Carter et al., 2018; Schunk & Pajares, 2002). Therefore, understanding how LAD designs can influence self-efficacy becomes a critical pursuit, benefiting all stakeholders interested in enhancing either occupational or academic performance.

Delving into the details of self-efficacy, individuals are said to form their beliefs about their capabilities through a critical analysis of specific sources of self-efficacy information, including mastery experiences and social modelling (Bandura, 1997). Focusing first on mastery experiences information, these serve as the primary and most influential source of self-efficacy information (Bandura, 1997). Such experiences provide direct evidence of an individual's capacity to execute particular tasks. Successful performance can strengthen self-efficacy beliefs, while failure can weaken them.

Turning to social modelling, this secondary source shapes self-efficacy beliefs through social comparison processes (Bandura, 1997; Mcintyre & Eisenstadt, 2011). In settings where objective performance measures are less accessible, individuals often compare their performance against that of their peers. Superior performance can bolster self-efficacy, whereas inferior performance has the potential to diminish it.

This dissertation highlights the interconnection between sources of self-efficacy information—namely mastery experiences and social modelling—and the criteria for success which are part of the self-evaluation sub-process of the self-reflection phase of the SRL cycle. Specifically, mastery experiences align with the mastery and past performance criteria, while social modelling corresponds to normative criteria. This alignment highlights the mechanisms through which learners develop self-efficacy beliefs during the self-reflection phase of the SRL cycle and the potential implications of different LAD designs on self-efficacy.

1.2.3 Social Comparison Theory

Social Comparison Theory highlights the mechanisms by which LADs employing social reference frames may influence learners' self-efficacy, principally through the provision of social modelling information. Social comparison theory posits that individuals evaluate their abilities by comparing themselves with others (Festinger, 1954; Gerber et al., 2018; Suls et al., 2002; Wheeler & Suls, 2020). Comparing oneself to peers can help gauge one's ability to perform specific tasks. Social comparisons can inform one's self-efficacy beliefs as they can act as a source of self-efficacy in the form of social modelling information. Similarly, they also serve as normative criteria for success during the self-reflection phase of the SRL cycle.

1.2.4 Temporal Comparison Theory

Temporal Comparison Theory (Albert, 1977) highlights how LADs utilising progress reference frames can shape self-efficacy by offering learners mastery experiences information. It elucidates how learners find it useful to compare their current performance and abilities against their prior achievements (Möller & Marsh, 2013; Wilson & Shanahan, 2020). Two key propositions from temporal comparison theory are salient in this dissertation. First, temporal comparisons occur during periods of rapid change as a means of evaluating said change. In the context of a learning task, change refers to a changing in ability. When receiving task feedback from multiple time points, learners can more readily detect if a desired change has occurred. Second, temporal comparison theory suggests that individuals have an inherent desire to make future predictions. By evaluating their past performance in comparison with their current performance, learners can assess their progress toward meeting their learning objectives. Temporal comparisons also inform self-efficacy beliefs as a form of mastery experiences information and serve as mastery and past-performance criteria for success during the self-reflection phase of the SRL cycle.

1.2.5 Directional Comparison

The concept of directional comparison (Collins, 1996; Guyer & Vaughan-Johnston, 2018) provides insight into the ways in which a learner's performance, when set against with various points of comparison (i.e., reference frames), might affect their self-efficacy. It functions as an overarching framework for understanding how an individual’s level of performance interacts during social and temporal comparisons for the purposes of self-evaluation. The implications of directional comparison are contingent upon various factors, one of which is the disparity between one's performance and a chosen comparison point (Gerber et al., 2018; Guyer & Vaughan-Johnston, 2018; Zell & Alicke, 2010). Directional comparison encapsulates downward, lateral, and upward comparisons, each with distinct influences on self-evaluation and in turn, self-efficacy beliefs. Importantly, research suggest that neither downward nor upward comparisons consistently yield predictable self-evaluations (Guyer & Vaughan-Johnston, 2018; Wilson & Shanahan, 2020; Zell & Alicke, 2010). This dissertation probes into the specific role that LAD design plays in affecting these outcomes as a means of providing additional support for stakeholders making decisions related to LAD design.

1.2.6 Goal Origin

As another crucial element in this dissertation, Goal Origin Theory (Hollenbeck & Brief, 1987; Seo et al., 2018) aims to connect established pedagogical principles with the designs of existing learning analytics reference frames. Goal Origin Theory explores the mechanisms underlying the setting of goals in learning contexts, with particular relevance to LADs. Goals may originate from either external sources, termed 'assigned goals,' or from internal sources, the learners themselves, termed 'self-set goals' (Hollenbeck & Brief, 1987; Locke et al., 2015). Assigned goals, often set by teachers or trainers, provide precise benchmarks that can serve to minimise ambiguity and guide learners. Such goals have been shown to be particularly effective in specific academic contexts, such as academic writing and proofreading (Ariely & Wertenbroch, 2002). Conversely, self-set goals, generated by the learners themselves, offer a degree of autonomy that can potentially increase intrinsic motivation (Ryan & Deci, 2000). However, the autonomy associated with self-set goals is complex; it can be both beneficial and detrimental, depending on the complexity of tasks and the context in which they are set (Osman, 2012).

In summary, the existing literature indicates that the effectiveness of goal origin—whether self-set or assigned—is influenced by various factors such as context, SRL skills, and additional support mechanisms (Locke et al., 2015). This dissertation specifically aims to highlight the role goal origin can play in learning analytic reference frame design and learner preferences for different designs.

1.2.7 Achievement Goal Orientation

Building upon the examination of Goal Origin Theory (Hollenbeck & Brief, 1987; Locke et al., 2015), this dissertation also incorporates the salient principles of Achievement Goal Orientation Theory (Pintrich, 2000a). At its core, this theory posits that learners' motivations and behaviour are predicated on the types of goals they elect to pursue, which are typically classified as either mastery or performance goals (Pintrich, 2000a). This serves as a valuable lens through which to explore and understand how learners engage with tasks.

In the context of LADs, the theory takes on additional complexity. Rather than merely reflecting individual choices, the orientation toward mastery or performance is often shaped by the design of the LAD itself (Corrin & De Barba, 2014; Fleur et al., 2023). Specifically, the type of reference frame chosen by the dashboard designer can likely influence the goal orientation of learners.

For instance, an LAD employing a progress reference frame, which accentuates temporal comparisons, tends to promote a mastery orientation. This subsequently fosters a learning environment that promotes skill development and personal growth. Such an orientation encourages learners to employ mastery criteria during self-evaluation taking place within the self-reflection phase of the SRL cycle. These criteria, in turn, can serve as valuable 'mastery experiences,' which are pivotal in shaping self-efficacy beliefs.

In contrast, an LAD featuring a social reference frame, designed to facilitate social comparisons, is inclined to instigate a performance orientation. This orientation propels learners toward a competitive mindset, stimulating the desire to outperform their peers. Here, normative criteria come into play, providing potential 'social modelling' information, another significant influencer of self-efficacy, during the self-evaluation stage of the SRL cycle.

1.3 Methodology

To provide a robust empirical foundation for this research, different methodological approaches are employed at various stages. In the initial stage of this research, Comparative Judgement (Pollitt, 2012; Thurstone, 1927; Whitehouse & Pollitt, 2012) and multinomial logistic regression (Agresti, 2002) to analyse learner preferences for learning analytics reference frames and their relation to perceived self-regulated learning skills. Subsequent chapters utilise Bayesian informative hypothesis evaluation (Hoijtink et al., 2019; Van Lissa et al., 2021) to assess evidence for or against competing hypotheses concerning reference frames and learning-related variables.

1.4 This Dissertation

This dissertation presents four interconnected studies that collectively aim to address the key research question: 'How does LAD design influence learner preferences, interaction, and self-efficacy in training and education?' To add a broader context, this research is partially conducted within the framework of the CHARMING project. This initiative targets the fortification of Europe's standing in the global chemical industry via innovative educational strategies and training methods. CHARMING actively explores the potential of immersive learning technologies, such as VR training environments, to better equip employees for the chemical industry. This project was established as an interdisciplinary and inter-sectorial network, combining expertise from leading universities, industry participants, and PhD researchers spanning fields as diverse as chemical engineering, educational sciences and pedagogy, games design, and immersive technology.

The first three studies featured in this dissertation are directly born out of the CHARMING project, concentrating on the design and implementation of LADs within VR digital learning environments specifically tailored for workplace-based training in the chemical industry. Expanding beyond the bounds of the CHARMING project, the fourth study in this dissertation represents a collaborative effort with a learning analytics team at Utrecht University. This final study extends the research scope to the context of higher education, examining the use of LADs in blended learning environments with the goal of investigating the relationships between LAD design and academic self-efficacy.

By creating a bridge between theory and practice in two rapidly emerging contexts of digital learning implementation - workplace-based training and higher education - this dissertation contributes to the growing body of literature on LAD design and addresses the need for more theory-informed design. The research narrative unfolds across the next four detailed chapters, each integral to the dissertation and collectively enriching our understanding of learning analytics in digital learning environments.

Chapter 2 explores a noteworthy gap in the existing literature: the lack of comprehensive understanding regarding workplace learners' design preferences for LADs. The chapter is particularly concerned with how these preferences vary according to different phases of SRL—namely, the forethought, performance, and self-reflection phases. Additionally, the chapter investigates how these design preferences relate with learners' self-assessed SRL skills. This investigation is structured around two specific research questions: ‘In the context of an immersive learning environment, what are workplace learner preferences for learning analytics reference frames in LADs designed for before, during, and after task performance?’ and ‘In the context of an immersive learning environment, how are workplace learner preferences for learning analytics reference frames in LADs related to their perceived self-regulated learning (SRL) skills?’.

Understanding learners’ preferences for dashboard designs that align with distinct SRL phases and skills is crucial. This is because the efficacy of feedback mechanisms within these dashboards can be contingent upon the learner's perception of the information presented.

Chapter 3 answers two research questions that address a gap in our understanding of how the progress and social reference frame in LADs impact occupational self-efficacy during the self-reflection phase of the SRL cycle. The primary research question is: 'When controlling for workplace self-reflection as a phase of the SRL cycle, are there between-group differences in total change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?'.

To address this question, the chapter puts forth three competing hypotheses. Hypothesis 1 posits that the progress reference frame will lead to greater changes in occupational self-efficacy, given its emphasis on mastery experience information. Hypothesis 2 suggests that the social reference will lead to greater changes in occupational self-efficacy, due to its dual provision of mastery experience and social modelling information. Hypothesis 3, meanwhile, proposes that both reference frames will produce comparable effects on occupational self-efficacy.

The exploratory research question is: 'When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in direction of change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?'

The exploratory hypotheses are informed by the same theoretical framework that guided the formulation of our primary research question and hypotheses.

The necessity of this inquiry arises from the limited literature on how LADs with different reference frames influence the self-reflection phase and self-efficacy in VR simulation-based training. This is important, given the increasing prevalence of digital learning environments employing LADs to stimulate SRL, where the self-reflection phase is a core component (Valle, Antonenko, Valle, et al., 2021), and the established link between self-efficacy and job performance (Bandura & Locke, 2003; Carter et al., 2018; Song et al., 2018).

Chapter 4 represents a thematic shift from the focus on occupational self-efficacy in Chapter 3 to an analysis of user interactions with LADs. Using log-file data, this chapter analyses the engagement patterns of chemical plant employees with LADs, set within a virtual reality simulation-based training environment. The chapter extends the investigation of the same two LAD reference frames covered in Chapter 3: the progress reference frame and the social reference frame.

Guided by the overarching research question, 'How do reference frames influence LAD interaction?', this chapter formulates three specific research questions, each bolstered by a set of competing hypotheses: ‘Are there between-group differences in total time spent reviewing LADs with a reference frame?’, ‘Are there between-group differences in total time spent reviewing detailed task feedback?’ and ‘Are there between-group differences in engagement with LADs?’

For each research question, three hypotheses are posited. The hypotheses suggest that either the progress or social LAD group will exhibit a greater, lesser, or equal impact on the respective dependent variables: time spent reviewing LADs with a reference frame, time spent reviewing detailed task feedback, engagement with LADs. The rationale for these hypotheses is rooted in the theoretical understanding that the contrasting effects may be attributable to the stimulation of mastery goal orientation by the progress reference frame and performance goal orientation by the social reference frame.

This chapter expands the scope of inquiry set out in Chapter 3 by scrutinising how dashboard designs influence the interaction patterns of workplace learners. Such exploration offers valuable insights that could inform subsequent efforts to refine LAD implementations.

Chapter 5 represents a significant expansion of the inquiry into LADs, pivoting the focus from workplace learning environments to the setting of higher education. Employing a similar research design to Chapter 3, this chapter evaluates LADs designed with both progress and social reference frames, probing their impact on academic self-efficacy. The distinguishing feature of Chapter 5 is its exploration of an added dimension—the 'direction of comparison.' This concept encompasses downward, lateral, and upward comparisons made relative to a predefined 'point of comparison,' which varies depending on the chosen reference frame—either progress or social.

The guiding research question for this chapter is: 'How do the type and direction of comparison within learning analytics reference frames affect academic self-efficacy among higher education students?' To address this question, the chapter introduces a set of 10 competing hypotheses. These hypotheses account for both the type of reference frame (progress, social) and the direction of comparison (downward, lateral, upward).

The hypotheses state that the mean change in self-efficacy varies systematically depending on the type of comparison under consideration. Specifically, change in academic self-efficacy is anticipated to decrease in a structured manner, generally shifting from downward to lateral and then to upward conditions. The underlying mechanisms for these changes are ascribed to the dominant influence of either mastery experiences or social modelling information. In downward conditions, either mastery experiences or social modelling tend to have an overriding impact, depending on the hypothesis. In upward conditions, again either mastery experiences or social modelling information emerge as the more influential factor. In lateral conditions, both mastery experiences and social modelling information are often expected to have an equal effect on self-efficacy. Some hypotheses propose that learners may discount particular types of information in specific upward conditions, resulting in an equalising effect between mastery experiences and social modelling information.

The decision to examine the 'direction of comparison' in combination with the different types of reference frames (progress and social) is grounded in two principal considerations. First, this approach acknowledges that the 'direction of comparison' within LADs warrants theoretical attention. This variable serves as an element that could substantially impact the interpretation and consequent utility of learning analytics data. Second, it recognises that the influence of the 'direction of comparison' on academic self-efficacy is likely to be non-uniform and may interact with the type of reference frame employed. Given the relationship between academic self-efficacy and academic performance, understanding this interplay becomes valuable. Thus, by investigating the 'direction of comparison' in tandem with the types of reference frames, this chapter aspires to deepen our understanding of how LADs can be effectively configured to enhance academic self-efficacy among higher education students.

CHAPTER 2

2

Learning analytics dashboard design: Workplace learner preferences for reference frames in immersive training in practice

This chapter is based on:
Gallagher, T., Slof, B., Van der Schaaf, M., Arztmann, M., Garcia Fracaro, S., & Kester, L. (2023). Learning analytics dashboard design: Workplace learner preferences for reference frames in immersive training in practice. Manuscript submitted for publication. Department of Education, Utrecht University.

Acknowledgement of author contributions:
All authors designed the study. Timothy Gallagher contributed to the study's conceptualization, methodology, formal analysis, investigation, data curation, drafting the original manuscript, manuscript review and editing, visualization, and project administration. Bert Slof, Marieke van der Schaaf, and Liesbeth Kester contributed to the conceptualization, methodology, manuscript review and editing, supervision, and funding acquisition. Michaela Arztmann and Sofia Garcia Fracaro contributed to the investigation, resources, and project administration.

Acknowledgement:
The authors would like to thank Dr. Maarten Cruyff for sharing his statistical expertise.

Abstract

Learning analytics dashboards are increasingly being used to communicate feedback to learners. However, little is known about learner preferences for dashboard designs and how they differ depending on the self-regulated learning (SRL) phase the dashboards are presented (i.e., forethought, performance and self-reflection phases) and SRL skills. Insight into design preferences for dashboards with different reference frames (i.e., progress, social, internal achievement and external achievement) is important because the effectiveness of feedback can depend upon how a learner perceives it. This study examines workplace learner preferences for four dashboard designs for each SRL phase and how SRL skills relate to these preferences. Seventy participants enrolled in a chemical process apprenticeship program took part in the study. Preferences were determined using a method of adaptive comparative judgement and SRL skills were measured using a questionnaire. Preferences were tested on four dashboard designs informed by social and temporal comparison theory and goal setting theory. Multinominal logistic regressions were used to examine the relationship between dashboard preferences and SRL. Results show that the progress reference frame is more preferred before and after task performance, and the social reference frame is less preferred before and after task performance. It was found that the higher the SRL skill score the higher the probability a learner preferred the progress reference frame compared to having no preference before task performance. The results are consistent with other findings which suggest caution when using social comparison in designing dashboards which provide feedback.

Key words

Learning analytics dashboards, Reference frames, Workplace learning, Social comparison theory, Temporal comparison theory, Immersive learning environments.

2.1 Introduction

Immersive environments are increasingly being used for workplace-based training (Langley et al., 2016; Li et al., 2017). A main advantage here is the application of learning analytics dashboards (LADs). Learning analytics involves the measurement, collection, analysis and reporting of learner data to support and optimise learning (Siemens & Gasevic 2012). LADs visualise learner data (e.g., performance score) by providing different types of reference frames, which are data comparison points (e.g., time, peers) which orient learners' interpretation of provided analytics data (Wise, 2014). Comparing data from different viewpoints can support learners by providing feedback on task performance (Valle, Antonenko, Dawson, et al., 2021; Wise, 2014) and stimulate self-regulated learning (SRL) behaviours (Jivet et al., 2017; Matcha et al., 2019). SRL has been coined as "the control students have over their cognition, behaviour, emotions and motivation through the use of personal strategies to achieve the goals they have established" (Panadero & Alonso-Tapia, 2014, pp. 1-2). It is a cyclical process which consists of three interrelated phases, which should all be stimulated by the LAD design (Winne, 2017; Zimmerman, 2013):

- forethought phase: taking place before the task and involving learners' implementing task strategies such as goal setting and strategic planning (Panadero & Alonso-Tapia, 2014; Zimmerman, 2002).
- performance phase: taking place during task performance and involving learners to monitor their own performance and optimise their learning efforts (Panadero, 2017; Zimmerman, 2002; Zimmerman & Moylan, 2009).
- self-reflection phase: taking place after task performance and involving learners to self-evaluate and attribute their success or failure to particular causes (Panadero & Alonso-Tapia, 2014; Zimmerman, 2013).

The effectiveness of LADs is, however, affected by how learners perceive and use them (Jivet et al., 2018; Nicol, 2020; Winstone et al., 2017). To address this, we combine insights from LADs, social and temporal comparison theory and goal setting theory. Based on these insights, different types of LAD's will be developed and augmented in an immersive learning environment for a workplace setting. Workplace leaners' preferences for LADs and their perceived self-regulation behaviour will be examined to gain more insight into their interrelationship.

The significance of this research lies in its novel focus on the often-overlooked demographic of workplace learners, exploring their learning preferences within professional, immersive environments. Moreover, it provides a detailed examination of the different phases of the SRL cycle in this context, probing into learner preferences related to each phase. The insights gained from this research will enhance our understanding of workplace learner perceptions of LADs and also contribute to the broader literature on learning analytics and self-regulated learning.

2.1.1 Designing LADS

Although the use of reference frames might stimulate performance and SRL behaviour, clear rules of thumb for designing them are lacking (Janson et al., 2022; Wilson & Shanahan, 2020). A first step in this direction is examining what is known about the use of comparisons in general and more specifically, into the relationship between SRL behaviour and learners' preferences. These generic insights could be valuable for designing LADs for the workplace setting this study is targeting. Social comparison, temporal comparison and goal setting theory are explored to gain insight into which types of references frames might be valuable. Furthermore, generic insight into learner preferences are provided.

Types of reference frames

Comparing performance scores might stimulate learners to determine how much progress they have made, if this was satisfactory for them, and utilize opportunities to improve themselves (Fleur et al., 2023; Suls et al., 2002). Comparisons can be made from a temporal as well as social viewpoint (Barreiros et al., 2023). Temporal comparison theory deals with comparing oneself at different points in time (Albert, 1977; Wilson & Shanahan, 2020). For example, one can make judgements about their performance on a particular task by comparing it with their past level of performance on a similar task. Social comparison theory deals with comparing oneself to others (Festinger, 1954; Wilson & Shanahan, 2020). For example, one can make judgements about their performance by comparing it to those of others on a similar task. By taking note of the achievement of others, people are able to gauge their own task efficacy, with varying degrees of accuracy (Cleary, 2009). Learners might be more inclined to use social comparisons when information of prior task performance is lacking (Bandura, 1997; Wilson & Shanahan, 2020).

Another perspective on comparisons is offered by goal setting theory. A goal is “the object or aim of an action, for example, to attain a specific standard of proficiency” (Locke & Latham, 2002, p. 705). Goal setting is used to stimulate behavioural change which, in turn, may enhance one’s performance” (Epton et al., 2017; Locke & Latham, 2002). According to Zimmerman (2013), goal setting is a forethought phase process. The targeting of goals can be self-set by the learner or set for them by someone else (Hollenbeck & Brief, 1987). The effects of both types of goal setting are mixed (Osman, 2012). Assigned goals can lead to greater task performance in the context of academic writing and proof reading (Ariely & Wertenbroch, 2002), while in the context of improving athletic performance, assigned goals do not differ from self-set goals (Fairall & Rodgers, 1997). A plausible reason for this may be that learners differ in goal setting skills and might find it difficult to set challenging goals and achieve them (Epton et al., 2017). Self-set goals may satisfy learners need for autonomy which could promote their intrinsic motivation (Ryan & Deci, 2000). Interventions (e.g., setting aside a specific amount of time on specific days to complete coursework) aimed at encouraging commitment could promote the achievement of self-set goals (Seo et al., 2018). Goals assigned to a learner, for example, by a trainer, have the potential to be more appropriately challenging (Epton et al., 2017; Latham & Locke, 1991). A trainer presumably has more insight into what the criteria for good performance are.

2.1.2 Learner preferences

As indicated above, the effectiveness of LADs also depends on learner preferences. We can look at prior studies to gain valuable insights. For example, Ruble & Flett (1988) examined leaner preference for reviewing one's own or a peers performance score on a math exam. Results indicate learners preferred social comparison. Especially low ability learners were less inclined to opt for temporal comparison. Interestingly, the appreciation of temporal comparison appeared to increase as learners got older.

Tuning into settings where learning analytics was used also revealed interesting findings. Konert et al. (2016) augmented Moodle with an LAD aimed at improving learner SRL skills by enabling them to set goals, keep track of knowledge gained from the course and time investment. A social comparison feature enabled them to compare their own knowledge level and time investment with their peers. The evaluation of the LAD indicated that learners were most positive about the social comparison feature. Tabuenca et al (2015) asked learners to rate their preferences with regard to personal learning analytics, social analytics and teacher estimations. Each of these included a form of progress, social and external achievement reference frame respectively. Results indicated that the progress reference frame was preferred over the social and external achievement reference frame.

Guerra et al (2016) evaluated a learning analytics system augmented with temporal and social comparison features. Results indicated that student engagement, efficiency and effectiveness were positively affected by social comparison features. Furthermore, results of a usability and usefulness survey, indicated that both comparison types were appreciated by learners, particularly those who were highly motivated. Gasevic et al (2013) analysed workplace learner perceptions of usefulness of a learning environment augmented with learning analytics. Results indicated that workplace learners use social comparisons when engaging in SRL processes such as goal setting. The study suggests that interventions aimed at supporting self-regulatory processes in the workplace should account for social and organisational elements, such as the alignment of learning goals and activities of colleagues and organisational goals. Gallagher et al (2022) conducted a study with an immersive training environment for a workplace. They compared trainee engagement with two LADs, one designed with a progress reference frame and one with a social reference frame. Results indicated that trainees receiving a progress reference frame were more likely to engage with the LAD than those receiving the social reference frame.

Finally, there are indications that learner preferences for a specific type of reference frame may also depend on their SRL skill level and the specific SRL activities they are involved in (Zimmerman, 2013). SRL skill level affects one's ability to set their own goals (Epton et al., 2017; Latham & Locke, 1991; Zimmerman, 2002). A higher skilled SRL learner may prefer setting their own goals because they recognise their ability to successfully execute goal related SRL processes. On the other hand, lower SRL skilled learners may prefer goals assigned by their trainers or managers as they may believe they have more insight into what goals are best.

Reference frames may provide valuable comparisons, but it is up to the learner to navigate the cyclical SRL phases and act upon the provided information (Panadero, 2017; Zimmerman, 2002). Learners make choices about how to execute SRL processes based on their self-evaluations derived from the provided comparisons. Interesting here is that the same performance outcome can look like success or failure depending on the point of comparison being used to contextualise feedback. For example, if a learner scores 85% on a training task and is offered a social reference frame with a point of comparison of 75%, this would likely trigger positive self-evaluations because their performance outcome is greater than the point of comparison. However, if a different reference frame, such as an external achievement reference frame, with a point of comparison of 90% was offered, then this may trigger negative self-evaluations because their performance outcome is lower than the point of comparison. This illustrates that although the performance outcome in both examples remains constant at 85%, the point of comparison can differ and in turn potentially trigger either positive or negative self-evaluations. Due to the cyclical nature of SRL, this has implications on the other phases of the cycle because positive or negative self-evaluations likely differentially affect forethought phase processes, which in turn affect performance phase processes.

2.1.3 This study

This study examines learner preferences for four mock LADs aimed at stimulating SRL behaviour. The LADs are augmented in a virtual reality based immersive learning environment developed for the chemical process industry. The environment will serve a dual purpose. Firstly, it will train employees on the procedural steps involved in the production specific chemical compounds. Secondly, it will equip operators with the necessary skills and knowledge to respond effectively to emergency situations. By doing so, the platform aims to increase safety and efficiency within the chemical production process. Three primary topics will be explored to gain a more in-depth understanding of learner preferences for LAD design within immersive learning environments. Firstly, we explore whether or not learner preferences for reference frames in LADs differ depending on if the LAD is designed for before, during or after task performance. It may be the case that preferences differ because different SRL skills are required depending on the phase of the SRL cycle. For example, learners may prefer a progress reference frame during the 'before task phase' more than any other reference frame, because they may think they can use of temporal comparison more effectively. This could see temporal comparison as advantageous because it offers them insight into if they are progressing compared to prior tasks attempts or not, which may be valuable goal setting and/or strategic planning information.

Secondly, we explore whether or not a workplace learner's perceived SRL skills relates to their preference. For example, it is unclear if higher skilled self-regulated learners prefer one reference frame type while lower skilled self-regulated learners prefer another. For example, we speculate that it may be the case that high skilled self-regulated learners are better supported by a progress reference frame because they are able to make use of temporal comparison during the three phases of the SRL cycle (Winne, 2017; Zimmerman, 2013).

By doing so, we will answer two research questions, namely:

Research Question 1: In the context of an immersive learning environment, what are workplace learner preferences for learning analytics reference frames in LADs designed for before, during and after task performance?

Research Question 2: In the context of an immersive learning environment, how are workplace learner preferences for learning analytics reference frames in LADs related to their perceived SRL skills?

2.2 Method

2.2.1 Participants

The participants (*N=70*) were employees of a science and technology company located in Germany and were all trainees of the company's chemical process apprenticeship program. The employees' working language was German. See Table 1 and 2 for demographic details.

Table 1

Gender of Participants

Gender	Number of Participants
Male	48
Female	15
Not specified	7
Total	70

Table 2

Age of Participants

Age	Number of Participants
18-21 years	30
22-25 years	25
26-29 years	10
Over 30 years	3
Not specified	2

This sample size was chosen for contextual reasons. These 70 participants represented the available members of the apprenticeship program at the time of the study. More importantly, the immersive learning environment, for which the LADs were being designed, was specifically intended for this program. Thus, the choice of participants aligns directly with the real-world application of our research, providing valuable, context-specific insights.

Out of the total 70 participants, data from everyone was utilised for group level results. However, for individual level results, data was only used from a subsection of the participants. We collected data on individual reference frame preference for before task performance from 54 participants, for during task performance from 62 and for the after task performance from 58. There were three reasons why some participants did not have their data collected. Firstly, due to technical issues (i.e., the system did not offer one or more reference frames to the participants for comparison which led to researchers being unable to determine a reference frame preference) (before: 9, during: 2, after: 7). Secondly, due to participants dropping out of the comparisons and completing less than 6 rounds of comparison (before: 4, during: 3, after: 3). If participants completed less than 6 rounds of comparisons, it was not possible to determine a reference frame preference. Finally, because the participants did not complete the SRL questionnaire (n=4).

Participation in this study was entirely voluntary; participants were neither compensated financially nor in credits for the apprenticeship program. The researchers asked for active consent and informed the participants that they could withdraw at any moment without a given reason. There were no participants who asked to withdraw from the study, however, as described earlier, some dropped out while completing aspects of the study. The responses to the questionnaires were pseudonymised which concealed the identity of the participants.

2.2.2 Design

A within-group design was used in which learner preferences were tested using adaptive comparative judgement. Adaptive comparative judgement enables us to place each reference frame design (i.e., social, progress internal achievement and external achievement) in a rank order from most to least preferred for the entire group of participants. In addition, a parameter value is assigned to each reference frame to give an indication of how much more or less one reference frame is preferred over another. We also determine each participant's reference frame design preference before, during and after task performance, which is used in a within group design to test the relationship between SRL skill and reference frame preference.

2.2.3 Materials

Learning analytics dashboards

The learning analytics feedback used for this research came in the form of 12 distinct mock LADs with fictious data; four per task stage (before, during and after task performance) designed for three task phases. Each dashboard differed by which reference frame it was designed with (i.e., social reference frame, progress reference frame, internal achievement reference frame, external achievement reference frame) and which task stage it was designed for (i.e., before task, during task, after task).

To support the design choices made for the mock LADs and the study design, we draw on design recommendations for learner facing LADs which were the result of a literature review by Jivet et al. (2018). Table 3 restates these recommendations as design guidelines and explains how these manifest in the mock LADs presented to participants and the study design.

Table 3

Design Guidelines Based on Recommendations from Jivet et al. and Their Manifestation in the Mock LADs and Study Design.

Design guidelines	Manifestation of recommendations in mock LADs or study design
Design LADs as pedagogical tools to enhance awareness and reflection, and to promote changes in cognitive, behavioural, and emotional competences.	• By offering feedback on performance before, during and after task performance, learners are made aware of their strengths and weaknesses related to their task, which can develop cognitive, behavioural and emotional competencies. • Self-reflection is enhanced through the offering of a reference frame, which learners can use to better understand how their performance compares to a particular point of comparison.

Use educational concepts from learning sciences to guide design decisions.	• The decision to offer LADs before, during and after task performance is motivated by theories of SRL and the three phases of the SRL cycle. • The design of each reference frame is informed by temporal comparison theory, social comparison theory and goal setting theory.
Use social comparison cautiously.	• One of the motivating factors for comparing the four dashboard designs is to collect additional evidence related to this recommendation.
Customise LADs to cater to different groups of users with the aim of providing equal support to all.	• One of the motivating factors for examining the relationship between SRL skills and reference frame preference is derived from this recommendation.
Integrate the dashboard seamlessly into the learning environment and learners' usual activities.	• This recommendation informed the decision to present the LAD before, during and after task performance because if an LAD is to support SRL, it seems beneficial for it to be presented at each phase of the SRL cycle.

The mock LADs were tailored for a chemical process industry immersive learning environment and were the result of a collaboration between educational scientists and instructional designers and content matter experts from the field of chemistry and chemical engineering. For illustrative purposes we briefly describe five mock LADs here.

Figures 1, 2, 3 and 4 are the mock LADs designed with a progress, social, internal achievement and external achievement reference frame for the 'after task phase' before being translated into German. All mock LADs can be found in the supplementary materials. Figure 5 is the mock LAD for the progress reference frame for the 'during task phase'. Fictitious performance data was used for all dashboards.

Each dashboard includes a sentence highlighting the type of comparison within the LAD (i.e., "Your score is being compared with your score on previous attempts"). This helps ensure participants were able to easily distinguish between reference frames. The score is a performance measure and refers to the percentage of the task completed correctly. Each dashboard includes an indicator highlighting the stage of the task the learner is in (i.e., “session ready”, “training paused”, “session complete”).

For the before and after task phase, an indicator titled “Training Effort” highlights how much time the learner has spent in the simulator practicing the task. The equivalent indicator for the during task phase provides a task progress bar indicating how much of the task has been completed and how much is left to complete.

For the before task phase, the task that is to be performed is stated and for the after task phase the task which has just been completed is stated. For the during task phase, an indication of sub-steps within the phase is shown. Task performance feedback is indicated by each dashboard using a bar chart for the before and after task phase which is presented alongside the point of comparison while in the during task phase, the percentage is stated in numbers.

Figure 1

Learning Analytics Dashboard with a Progress Reference Frame for the 'After Task Phase' Before Translation.

Figure 2

Learning Analytics Dashboard with Social Reference Frame for 'After Task Phase' Before Translation.

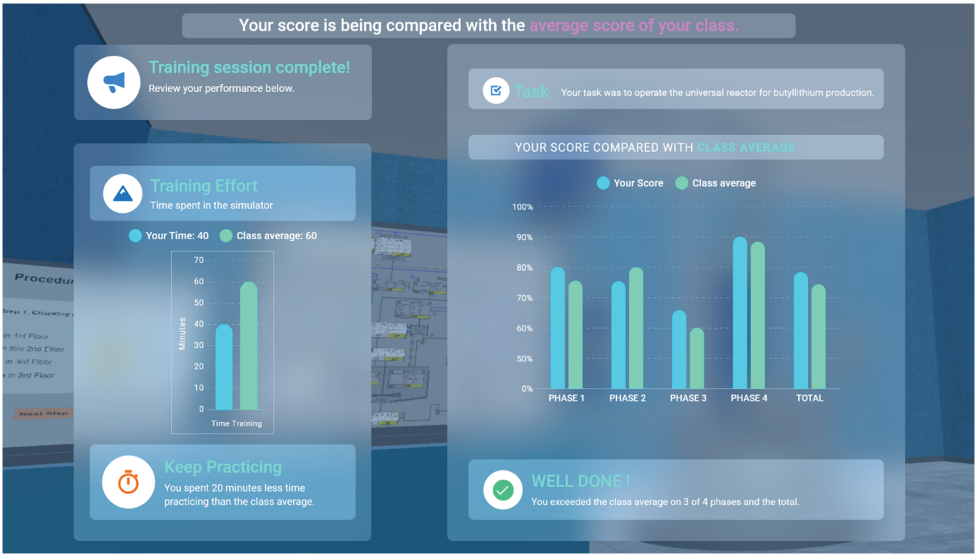


Figure 3

Learning Analytics Dashboard with an Internal Achievement Reference Frame for 'During Task Phase' Before Translation.

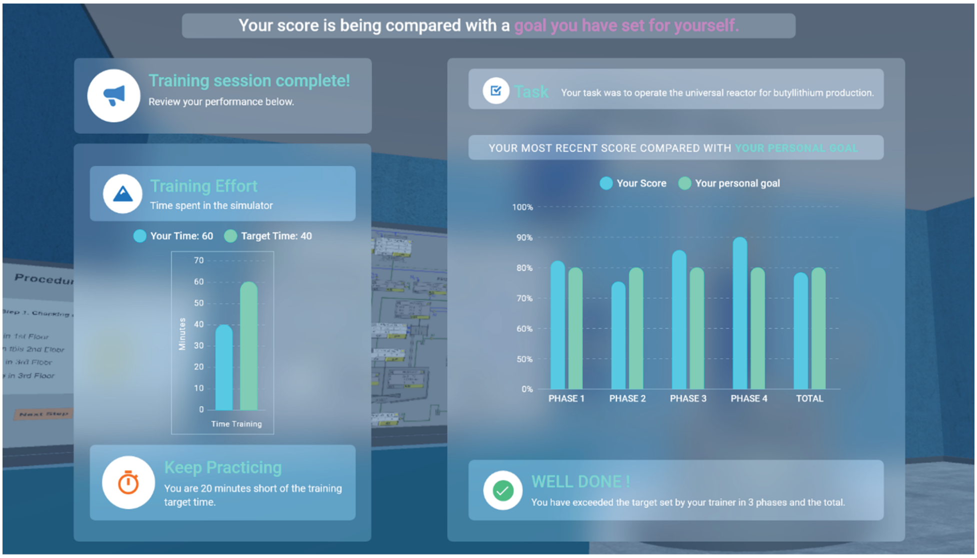


Figure 4

Learning Analytics Dashboard with an External Achievement Reference Frame for 'After Task Phase' Before Translation.

Figure 5

Learning Analytics Dashboard with Progress Reference Frame for 'During Task Phase' Before Translation.

The LADs were designed to appear within the immersive learning environment, however, due to the feasibility of sharing the designs in that format, screenshots of the mock LADs were taken and shared with participants. All mock LADs can be found in the supplementary materials.

2.2.4 Instruments

Demographic data questionnaire

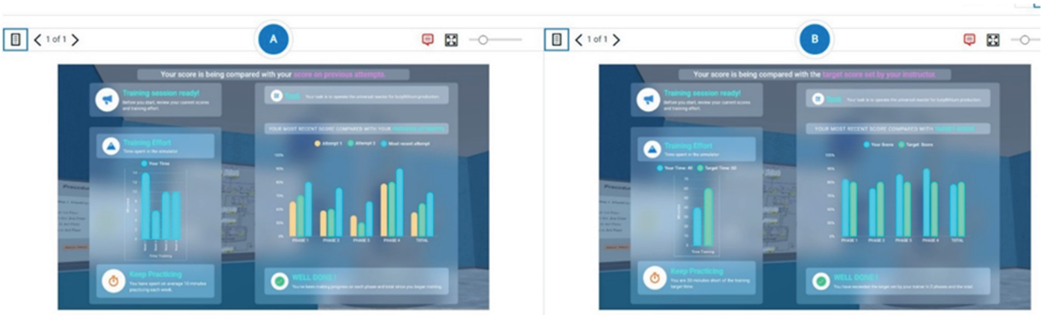
A demographic questionnaire was completed by all participants which asked them their age and gender. This questionnaire was distributed using Microsoft Forms.

Adaptive Comparative Judgement Software

RM Compare was used to investigate learner preferences for the learning analytics reference frames (https://www.compare.rm.com/). RM Compare is an online tool for conducting adaptive comparative judgement which is a method by which individuals are given a series of dichotomous choices between stimuli, from a larger pool of stimuli, and asked which one they prefer based on a given question (Pollitt, 2012; Verhavert et al., 2019). In the case of this research, the stimuli are screenshots of the mock LADs. RM Compare's design is informed by the theory of comparative judgement (Pollitt, 2012; Thurstone, 1927) that states when presented with a series of stimuli in which an individual must make subjective judgements on which stimulus is better or worse compared to another stimulus, it is rational to compare each stimulus with each other stimulus. After a group of individuals have made these subjective judgements, it is then possible to apply a mathematical model which takes into account each individuals judgements to calculate the rank order and parameter value of the compared stimuli for the entire group (Pollitt, 2012).

Figure 6

Participant View When Performing Adaptive Comparative Judgement in RM Compare.



Self-regulated online learning - Questionnaire

Learners' SRL was measured using the SOL-Q-R which was designed to measure SRL behaviours in online educational contexts (Jansen et al., 2017). Of the many different instruments available to measure SRL, the SOL-Q-R, was selected due to its validity (RMSEA = <0.102) and reliability (>α=0.740) for researching SRL (Jansen et al., 2017).

Participants completed the SOL-Q-R which is made up of 42 items. Each item is presented in randomised order and answered on a 7-point Likert scale, ranging from "not at all true for me" (=1) to "very true for me" (=7). It provides an overall score for perceived SRL skill and also provides sub-scale scores. For this research we chose to use the overall score.

2.2.5 Translation process (English to German): Questionnaire Translation Process–Front-translation, back-translation, reconciliation

Both questionnaires were originally developed in English, therefore, translation was required and was done following a front-translation, back-translation, and reconciliation process.

The mock LADs as well as the questions posed for each round of the adaptive comparative judgement were originally developed in English and were translated into German following a standard front translation process.

2.2.6 Procedure

After signing informed consent, participants were invited to attend one of six research sessions. The introduction to the research and data collection was completed online using a video conferencing tool. During these sessions, participants were guided through the research procedure by one of the authors who is a native German speaker. This research procedure included an introduction to the research which provided a brief overview of the research aims and details on how they could log into the RM Compare website using usernames and passwords which were assigned to them. Participants were shown a video of the immersive learning environment which was being developed for the company they worked for and were told that the research was focused on improving the design of this prototype, in particular, the LADs. As it was expected that LADs as a concept was new to the participants, a short presentation on LADs and reference frames was shared. During the research session, participants were told that they were to make their judgements based exclusively on the different reference frames they preferred and asked to do their best to ignore any aesthetic preferences they may have. Once the introduction to the research and data collection was complete, participants were able to ask any questions.

The order in which the instruments were administered was as follows. First, the demographic data questionnaire was completed and was followed by a Goal Orientation questionnaire. The goal orientation questionnaire was used by a collaborating researcher and the results from this questionnaire are not reported here. Next, participants were asked to log into the RM Compare website to begin the adaptive comparative judgement session.

Once participants had logged into RM Compare, they were presented with the first round of comparisons, which asked them to compare the four different LAD reference frame designs, designed for before task performance. In this first round, before making each comparative judgement, they were prompted with the question: "*Which design do you think will help you best with planning, goal setting and motivation?"*. Note that this question aligns with the forethought phases processes of SRL cycle.

Upon completion of this first round focusing on before task performance, participants then moved on to the 2nd round focusing on LADs designed for during task performance. For this task stage they were prompted with the question: *"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?"*. Note that this question aligns with the performance phases processes of SRL cycle.

For the third and final round, participants were prompted with the question: "*Which design do you think is best for helping you reflect upon your performance and judging or assessing your own performance?"*. Note that this question aligns with the self-reflection phases processes of SRL cycle.

After participants had completed the adaptive comparative judgment, they were then asked to complete the SRL online learning questionnaire which was administered using Microsoft Forms.

2.2.7 Scoring and Data analysis

Self-regulated Online Learning

JASP version 0.16 is used for the statistical analysis of the validity and reliability of the SOL-Q-R. A confirmatory factor analysis is used to determine validity and a unidimensional reliability test to calculate Cronbach's α. Four statistical tests measure validity, including the Root Mean Square Error of Approximation (RMSEA) with values below .08 indicate adequate fit (Chen et al., 2008; Maccallum et al., 1996), the standardized Root Mean Residual (SRMR) with values below .05 indicating good fit, the Comparative Fit Index (CFI) with values above .90 indicating good fit and the Tucker Lewis Index (TLI) with values above .95 indicating good fit (Hu & Bentler, 1999).

For the SOL-Q-R, the confirmatory factor analysis thresholds for validity were partially met (RMSEA = 0.127, SRMR = 0.047, CFI = 0.949, TLI = 0.924), and the reliability threshold was achieved (α = 0.905). The RMSEA value did not fall within the preferred range. However, considering the sample size and the other fit measures being within acceptable ranges, the results can still be deemed satisfactory.

Adaptive Comparative Judgement (RM Compare)

This study analyses the results from the adaptive comparative judgement on a group level and individual level. Group level results refer to how all participants valued the reference frames from most preferred to least preferred. Individual level results refer to which reference frame (if any) was most preferred by each individual. Individual level data was required to analyse the relationship between SRL skill and reference frame preference.

Therefore, we report results from the adaptive comparative judgement in two categories: group level results and individual level results.

Group level results

Parameter value along with the standard error (SE) is used to answer Research Question 1 and was automatically calculated by RM Compare.

The parameter value indicates to what degree a reference frame was preferred by the group over another reference frame and is calculated based on the win/loss record of each reference frame (Bartholomew et al., 2018; Whitehouse & Pollitt, 2012). Scores closer to 1.0 indicate stronger preference while scores closer to -1.0 indicate weaker preference. The difference between two parameter values is the likelihood that the reference frame with the higher parameter value is preferred in judgement in comparison with the reference frame with the lower parameter value.

We used parameter values in concert with the SE to interpret reliability of the parameter values. The SE estimates the potential error for a particular parameter value and therefore indicates how confident we can be of the parameter value assigned to each reference frame. To determine if two parameter values are meaningfully different, we will check the standard error of each parameter value and compare it with the standard error of each other parameter value. If there is any overlap between two reference frame SE ranges than we will treat those reference frames and not being meaningfully different.

The scale separation reliability score (SSR) helps us interpret internal consistency of the results (Verhavert et al., 2018). The closer SRR is to 1, the more likely it is that if the comparisons were done again, we would get the same results. An SSR value greater than 0.7 suggests good internal consistency (Marshall et al., 2020). In that way, SSR is analogous to Cronbach's alpha (Marshall et al., 2020; Pollitt, 2012).

The internal consistency which is represented by the SRR score shows good levels of consistency for results examining before (0.82) and after (0.9) task performance preferences, but not during (0.58) task performance preferences. Therefore, we are confident that the results for the before and after task performance preferences are reliable, however, the relatively low SSR score for during task performance preferences results means we cannot confidently draw conclusions for this phase.

Individual level results

There is one variable from the individual level results for each task stage (before, during, after task performance). This variable is the learner reference frame preference. The reference frame preference variable indicates if the learner holds a reference frame preference or not and if there is a preference, what that preference is.

The reference frame preference variable was coded as one of the following options: progress, internal achievement, external achievement, social, no preference.

Coding the reference frame preference variable was done manually by looking at each pair-wise comparison the learners made and inferring which reference frame they preferred (if any). For example, if the results of one participant showed that the social reference frame was preferred over the progress reference frame in one comparison, and in the next two comparisons, the progress reference frame was preferred over the external achievement reference frame and the internal achievement reference frame, it was inferred that the social reference frame was the most preferred for that participant. If, however, there was a contradiction in comparisons, no preference was inferred. For example, if the social reference frame was preferred over the progress reference frame, the progress reference frame was preferred over the internal achievement reference frame and the internal achievement reference frame was preferred over the social reference frame, we would conclude a logical inconsistency had occurred and that the learner holds no clear reference frame preference.

Multinominal logistic regressions

To examine the relationship between learner preferences and SRL skills, three multinominal logistic regressions (before, during and after task performance) were conducted. In SPSS reference frame preferences were set as the dependent variable and SRL score obtained from the SOL-Q-R as the independent variable. Before conducting the analyses, the independent variable (SRL skill) was tested a priori to verify there was no violation of the assumptions required for multinominal logistic regressions and none were violated.

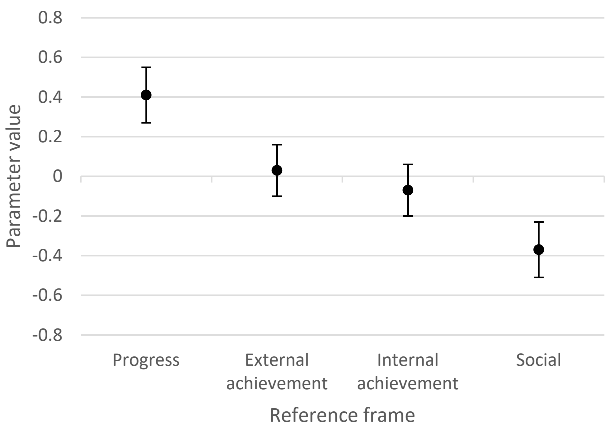
2.3 Results

2.3.1 Preferences for LADs

We answer Research Question 1 by taking into account each parameter value with the SE, as well as the SSR score, all of which were yielded by the adaptive comparative judgement. For before and after task performance the parameter values paired with the SE (Figure 7) indicate that the opportunity to compare one's own performance over time was substantially preferred over all other types of comparison (0.41, SE=0.14). The parameter values paired with SE fail to indicate that comparison with a trainer assigned goal (0.03, SE=0.13) is more preferred than a self-set goal (-0.07, SE=0.13). The results indicate that both comparisons with a trainer assigned goal and self-set goals are substantially preferred over the comparison with peer performance (-0.37, SE=0.14).

Figure 7

Before Task Phase Parameter Values and Standard Errors

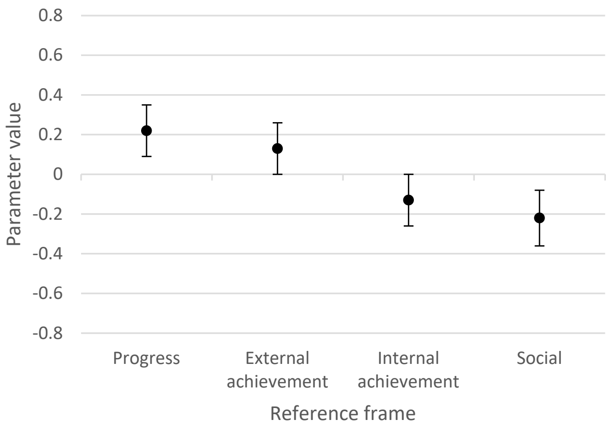


Note. SSR = 0.82 (SSR > 0.7 = good internal consistency)

For during task performance (Figure 8), the parameter values and SE show that comparison with one's own performance over time (0.22, SE=0.13) is not substantially preferred over the comparison with goals assigned by one's trainer (0.13, SE=0.13). Both are substantially preferred over the comparison with self-set goals (-0.13, SE=0.13) and comparison with the performance of others (-0.22, SE=0.14). The comparison with self-set goals is not substantially preferred over the comparison with peer performance when taking into account the parameter values paired with the SE. However, as indicated earlier, the SSR score (0.58) indicates that the reliability of these results fail to meet the threshold of good reliability.

Figure 8

During Task Phase Parameter Values and Standard Errors

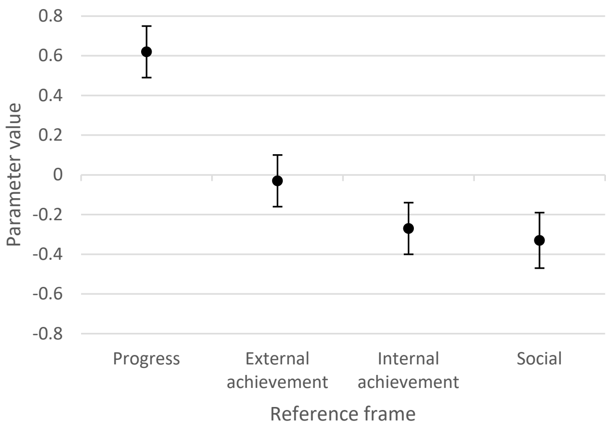


Note. SSR = 0.58 (SSR > 0.7 = good internal consistency)

For after task performance (Figure 9), the parameter values and SE show that the comparison with one's own performance over time (0.62, SE=0.14) is substantially preferred over all other comparison types. Comparisons with trainer assigned goals (-0.03, SE=0.14) are not substantially preferred over comparisons with self-set goals (-0.27, SE=0.14), however, they are substantially preferred over comparisons with peers (-0.33, SE=0.14). It is worth noting that the overlap of the lowest point of the SE for trainer assigned goal comparisons and highest point of the SE for the self-set goal comparisons is marginal (0.04). Self-set goal comparisons are not substantially preferred over comparison with peers.

Figure 9

After Task Phase Parameter Values and Standard Errors



Note. SSR = 0.90 (SSR > 0.7 = good internal consistency)

2.3.2 Interplay between learner preferences and their perceived SRL skill

We answer research question 2 by conducting three multinominal logistic regressions; one for each set of preferences before, during and after task performance. We report descriptive statistics for before, during and after task performance preferences in Table 4, which are followed by the results from the multinominal logistic regression for before task performance in Table 5, during task performance in Table 6 and after task performance in Table 7.

Table 4

Descriptive Statistics for Reference Frame Preferences for Each Task Phase

Task reference frame preference	Task phase					
	Before		During		After	
Progress	n = 17	31.5%	n = 14	22.6%	n = 20	34.5%
Internal achievement	n = 10	18.5%	n = 2	3.2%	n = 4	6.9%
External achievement	n = 10	18.5%	n = 20	32.3%	n = 13	22.4%
Social	n = 3	5.6%	n = 7	11.3%	n = 3	5.2%
No preference	n = 14	25.9%	n = 19	30.6%	n = 18	31%
Total	n = 54		n = 62		n = 58	

Note. *n = Number of participants. The reference frame with the highest n does not necessarily have the highest parameter value because the rank order is also a factor. This is why External achievement reference frame has the highest n in Table 4 but not the highest parameter value in figure 8.*

The results provide information comparing each reference frame preference group (Progress, Social, Internal achievement, External achievement) against the Reference Category (No preference). Specifically, the regression coefficients indicate the odds ratios change as the score on SRL skill increased by one unit.

The result for before task performance shows that the model (Table 5) is approaching significance (p=.062) and taking into account the small sample size, further investigation would be fruitful. Table 5 indicates to a statistically significant degree (*p=0.013)* that the higher the SRL skill score the higher the probability a learner had of being in the progress reference frame preference group. The results for during task performance (Table 6) and after task performance (Table 7) indicate there are no statistically significant results.

Table 5

Multinominal Logistic Regression for Before Task Phase with No Preference Set as Reference Category.

		b (SE)	*p*	95% CI for Odds Ratio		
				Lower	Odds Ratio	Upper
External vs No preference	SRL Skill	.122 (.528)	.832	.397	1.118	3.149
Internal vs No preference	SRL Skill	.389 (.546)	.476	.506	1.476	4.305
Progress vs No preference	SRL Skill	1.419 (.573)	.013*	1.343	4.133	12.718
Social vs No preference	SRL Skill	.496 (.858)	.563	.306	1.643	8.826

$X^2 = 8.95$, df = 4, p = 0.062, **p* < .05, *b* = estimated coefficient for each predictor variable, SE = standard error

Table 6

Multinominal Logistic Regression for During Task Phase with No Preference Set as Reference Category.

		b (SE)	*p*	95% CI for Odds Ratio		
				Lower	Odds Ratio	Upper
External vs No preference	SRL Skill	.039 (.389)	.920	.485	1.040	2.229
Internal vs No preference	SRL Skill	-.010 (.899)	.991	.170	0.990	5.766
Progress vs No preference	SRL Skill	.687 (.469)	.143	.793	1.988	4.985
Social vs No preference	SRL Skill	.451 (.567)	.427	.516	1.570	4.773

$X^2 = 3.06$, df = 4, p = 0.549, *b* = estimated coefficient for each predictor variable, SE = standard error

Table 7

Multinominal Logistic Regression for After Task Phase with No Preference Set as Reference Category.

		b (SE)	*p*	95% CI for Odds Ratio		
				Lower	Odds Ratio	Upper
External vs No preference	SRL Skill	.216 (.478)	.652	.486	1.241	3.167
Internal vs No preference	SRL Skill	.713 (.776)	.358	.446	2.040	9.333
Progress vs No preference	SRL Skill	.706 (.456)	.122	.829	2.026	4.955
Social vs No preference	SRL Skill	1.609 (.973)	.098	.743	4.999	33.631

$X^2 = 4.94$, df = 4, p = 0. 293, *b* = estimated coefficient for each predictor variable, SE = standard error

2.4 Discussion

Set within a workplace learning context, this study set out to investigate learner preferences for four mock LAD designs aimed at stimulating SRL behaviour. The LADs provide feedback on an immersive learning environment task designed for training apprentices in the chemical process industry. The four LAD designs differed by reference frame (i.e., social, progress, internal achievement and external achievement) and therefore, the point of comparison used to help learners make sense of their feedback (i.e., social comparison, temporal comparison, comparison with self-set goal, comparison with assigned goal). The main finding of our study lies in the comparison of these different reference frame designs rather than the specific information shown in the dashboards. The study also sought to determine if there is a relationship between learner preferences for a particular reference frame before, during and after task performance and overall SRL skills in the context of a workplace learning environment.

The first research question examined learner preferences for reference frames before, during and after task performance. Results indicate a learner preference for the progress reference frame. This may be because temporal comparison, stimulated by the progress refence frame, is perceived by learners to be supportive of their forethought and self-reflection processes (Zimmerman, 2002). Strategic planning is one such process which helps learners master skills and perform optimally (Zimmerman, 2000a, 2013). As one develops skills on a particular task, initial strategies to acquire those skills can decline in effectiveness and new strategies are required for further improvement. If it becomes apparent that a learner's performance has plateaued or declined, this potentially indicates that the chosen strategies are insufficient, and they may benefit from planning and executing new ones.

Another possible explanation as to why the progress reference frame was preferred over other types of references frames before and after task performance is because participants of the study may have held mastery goal orientations. This aligns with findings from Jivet et al. (2020); learners with a mastery orientation, rated temporal comparison feedback to be very relevant.

Our findings indicate there were no substantial preferential difference for before and after task performance between the external achievement (assigned goal as a point of comparison), and the internal achievement (self-set goal as a point of comparison) reference frames. This aligns with research (Epton et al., 2017) indicating that goal achievement was not affected by the type of goal setting (self, assigned). Learners who have experienced success as well as failure with both types of goal setting, may have no preference for either type of reference frame. It contradicts a finding by Jivet et al., (2020) indicating that learners preferred an LAD feature labelled as “seeing requirements for passing the course”, above the feature “seeing my performance in comparison to my goals”. However, this study did not directly test if this perception of relevance was significantly different.

The social reference frame was the least preferred reference frame before task performance and equally least preferred during and after task performance. This may be because the participants perceive social comparisons to be detrimental to self-motivational beliefs. Self-motivational beliefs play an important role in the SRL cycle and are key in motivating learners to execute task analysis strategies such as goal setting and strategic planning (Zimmerman, 2013). For example, a study by El-Beheiry, McCreery, & Schlachta (2017) indicated that 76 percent of 25 participating first year surgical residents would not be motivated by social comparison in the form of a leader board intervention. However, contradictorily, the presence of leader board did in fact enhance residents practice with an immersive learning environment. This suggests that while learners may not prefer social reference frames over other types, such as progress reference frames, social reference frames may still result in positive learning behaviours. Other studies have obtained findings indicating that learners in some contexts do positively react to social reference frames. Guerra et al. (2016) found that learner engagement, efficiency and effectiveness were positively affected by social comparison features of an LAD and the progress as well as the social reference frame were appreciated by learners. Brusilovsky et al. (2016) found that LADs designed with both a progress and social reference frame were much more successful at engaging learners than those designed with only a progress reference frame. These inconclusive findings are supported by a literature review by Jivet et al. (2018), examining learner evaluations of LADs, which found that while in some cases social comparison via a social reference frame can have a positive effect on motivation, in other cases it can have negative effects.

Perhaps these inconclusive findings can be accounted for by differences between learners. Learners might cope in different ways with the feedback offered by the comparisons (Pintrich, 2000a; Zimmerman, 2013). If, for example, a learner receives a social reference frame which they dislike, they may then ignore the feedback within the LAD all together, which may lead them to poorly execute the SRL phases for which the LAD was intended to support. On the other hand, if a learner receives a reference frame which matches their preference, we may see beneficial effects on the execution of SRL processes. For example, Janson et al. (2022) conducted a study in examining the effects of learners' own preferences for temporal vs social comparison on learning persistence and performance in a digital learning environment and found a beneficial effect of matching feedback design with comparison preferences. This result suggests that we should be concerned with the mismatch between design choices and preferences and warrants further research into how the different types of reference frames affect learner engagement with feedback delivered via LADs at the different SRL phases.

The second research question examined the interplay between learner preferences for reference frames and their perceived SRL skills. Our findings indicate an interplay for one pair of before task performance preferences (progress reference frame vs no preference); learners with higher perceived SRL skills are more likely to prefer the progress reference frame than having no preference for before task performance LADs. A possible explanation is that higher SRL skilled learners have a better idea about the type of reference frame they need to help them with forethought phase processes. However, this reasoning fails to hold for during and after task performance preferences. The inconsistent results for before, during and after task performance could also be due to the methodological approach undertaken and the number of participants in the study. A qualitative investigation could shed further light on the results. For example, an interview could gain more insight into the rationale behind the indicated preferences.

In recognising the novelty of our work, it is important to underline a few key aspects. Firstly, our research uniquely targets the often-overlooked demographic of workplace learners (Poquet et al., 2022). The nature of learning within a professional environment differs notably from traditional educational settings (Ruiz-Calleja et al., 2017). As such, our study unveils valuable context specific insights into the preferences of this specific group, thereby expanding our understanding of the learner population. This intentional choice of our sample deserves emphasis. Given this intentional focus, the pool of available participants was inherently limited, leading to a smaller sample size of our study. However, despite its size, this select group offers critical information, providing us with a deeper understanding of learning preferences within professional environments. Furthermore, while the relationship between LADs and SRL is well-researched (Matcha et al., 2019), our study distinguishes itself by taking a granular look at the different phases of the SRL cycle. This nuanced approach allows us to better discern learners' distinct preferences across various stages of learning, with implications for more personalised and effective LAD designs, an area of research that has been called upon by various research in the learning analytics field (Wong et al., 2022).

In addition, our study examines the learner preferences among four distinct reference frame types - namely, progress, social, internal achievement, and external achievement. To our knowledge, no other research to date has comprehensively evaluated learner preferences across these four distinct categories. This new comparison offers valuable insights into the respective appeal of different LAD designs, which can directly inform the development of future learning analytics tools. Furthermore, our approach not only advances our understanding of learner preferences but also potentially opens the door for future studies on how these preferences may evolve or vary under different conditions or learning contexts.

Lastly, the immersive learning environment, the backdrop for our study, represents a relatively uncharted territory in LAD research (Beck et al., 2023). As this mode of learning continues to gain traction, our study provides early insights into learner preferences within this specific context. This knowledge can guide the development of LADs that cater to the unique needs of learners in immersive environments, thereby advancing our practice in this emerging field.

In summary, while working within the established domains of LAD and SRL, our research shines a light on under-researched areas and provides fresh perspectives on well-known concepts, thereby pushing forward our collective understanding of learning analytics and self-regulated learning.

2.4.1 Limitations

When interpreting the obtained findings, one should also take its limitations into account. Firstly, the study is conducted within the context of a workplace learning environment which limits the generalisability of the findings. Although such a context was suggested by others (Gedrimiene et al., 2020; Ruiz-Calleja et al., 2017). Another limitation is that the reference frame preferences were yielded from mock LAD designs with fictious data and therefore, it is not clear if preferences of learners would differ when learners are faced with authentic LADs in real time. Perhaps preferences based on real time data might yield different findings. Future research based on real time actual performance scores could provide valuable insight into learner preferences for LADs, how and why LADs are used (e.g., interaction, feedback processing, uptake feedback) and their effect on SRL and performance behaviour.

An additional limitation is the sample size of our study. While our sample was representative of a specific group of apprentices, it was relatively moderate in size, which may constrain the broader generalisability of our results and limit the extent to which our findings can be applied to a wider population. It's important to note that this limitation is, in part, a consequence of our intent to enhance the ecologically validity of the study. On one hand, this approach provides us with context-specific insights that hold substantial real-world value. On the other hand, the very nature of such studies place boundaries on the sample size. Despite this, our findings align with existing research in the field on learner reference frame preferences. This consistency, even within our specific context, underscores the value of our study and boosts our confidence that our work meaningfully contributes to the broader understanding of learner preferences in LAD designs.

Another limitation is that the feedback in the LADs always showed downward comparisons, except for one indicator in the during task phase LAD, meaning the fictious level of performance was greater than the point of comparison. It may be the case that when learners are required to make lateral or upward comparisons, preferences differ. The decision to primarily test downward comparisons was made to not overburden participants with visually similar comparative judgements. Finally, the number of participants in the study was rather small for the multinomial logistic regressions, which means we are unable to draw any strong conclusions regarding the relationship between SRL skill level and reference frame preference. Nevertheless, the results do indicate that further investigation into this line of inquiry would be fruitful.

2.4.2 Conclusion

We investigated learner preferences for LADs designed with different reference frames for before, during and after task performance. We used mock LADs which were designed to stimulate SRL in the context of an immersive training environment for the chemical process industry. For the before and after task phase, we found that the progress reference frame was most preferred. The results of the study can be used by LAD designers who wish to design LADs with reference frames that are appreciated by learners. Further research is needed to determine why learners appreciate different types of reference frames and how they affect learner performance and SRL processes.

3

CHAPTER 3

Reference Frames for Learning Analytics Dashboards: The Progress and Social Reference Frame and Occupational Self-Efficacy

This chapter is based on:
Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2023). Reference frames for learning analytics dashboards: The progress and social reference frame and occupational self-efficacy. Journal of Computer Assisted Learning, 1–19. https://doi.org/10.1111/jcal.12912.

Acknowledgement of author contributions
All authors designed the study. Timothy Gallagher contributed to the study's conceptualization, methodology, formal analysis, investigation, data curation, drafting the original manuscript, manuscript review and editing, visualization, and project administration. Bert Slof, Marieke van der Schaaf, and Liesbeth Kester contributed to the conceptualization, methodology, manuscript review and editing, supervision, and funding acquisition. Ryo Toyoda, Yusra Tehreem, and Sofia Garcia Fracaro contributed to the methodology, investigation, and resources. Additionally, Sofia Garcia Fracaro contributed to project administration.

Acknowledgement:
The authors would like to thank Prof. Herbert Hoijtink for sharing his statistical expertise.

This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement No 812716. This publication reflects only the authors' view exempting the community from any liability. Project website: https://charming-etn.eu/

Abstract

The potential of learning analytics dashboards in virtual reality simulation-based training environments to influence occupational self-efficacy via self-reflection phase processes in the Chemical industry is still not fully understood. Learning analytics dashboards provide feedback on learner performance and offer points of comparison (i.e., comparison with one's own past performance or comparison with peer performance) to help learners make sense of their feedback. We present a theoretical framework for describing learning analytics reference frames and investigate the impact of feedback delivered through dashboards with different reference frames on occupational self-efficacy, while controlling for workplace self-reflection. This experimental study engaged 42 chemical operator employees, aged between 18 and 55 years, each with at least one year of experience. We utilised a two-group design to ask two research question each with three competing hypotheses related to changes in occupational self-efficacy, employing Bayesian informative hypothesis evaluation. Results for the primary research question suggest that dashboards with progress reference frames do not elicit greater change to self-efficacy than those with social reference frames, however, they may elicit equal change. Furthermore, dashboards with social reference frames may elicit greater change to self-efficacy than those with progress reference frames. Exploratory results found that dashboards with progress reference frames may elicit greater positive directional change than those with social reference frames and that they may elicit equal directional change.

These findings contribute to the understanding of self-efficacy beliefs within the Chemical industry, with potential impacts on skill development. The research may inform the design of targeted interventions and training programs to influence self-efficacy. From a practical perspective, this research suggests that careful consideration is needed when

choosing reference frames in learning analytics dashboards due to their potential consequences on the formation of learner self-efficacy.

Key words:
Learning analytics dashboards, Self-efficacy, Self-regulated learning, Self-reflection, Social comparison, Reference frames.

3.1 Introduction

Virtual reality (VR) simulation-based training environments are gaining popularity for workplace training. They offer immersive and realistic experiences, enabling learners to practice tasks in a safe and controlled setting, eliminating risks and costs associated with physical training (Garcia Fracaro et al., 2021). These environments can incorporate learning analytics, which track progress and provide data-driven insights into learner performance (Siemens & Gasevic, 2012). These insights can be shared with learners as feedback (Gasevic, Dawson, & Siemens, 2015). One common method for sharing learning analytics insights is through learning analytics dashboards (LADs) (Matcha et al., 2020). For instance, an LAD can provide feedback on a learner's performance on a procedural training task, such as task completion time and number of mistakes made. To help learners make sense of such feedback, LADs can also incorporate reference frames, offering points of comparison learners can judge themselves against (Wise, 2014). Reference frames can include information related to changes in a learner's performance over time and how it compares to others (Jivet, Scheffel, Drachsler, & Specht, 2017). This information can potentially affect aspects of a workplace learner's self-regulated learning (SRL), such as self-reflection phase processes, and occupational self-efficacy, as they become more aware of their own capabilities and areas for improvement (Bandura, 1997). Self-reflection phase processes include the cognitive and metacognitive activities individuals engage in during the self-reflection phase of the SRL cycle, including self-evaluation and the self-assessment of one's own learning (Zimmerman, 2002). Occupational self-efficacy, on the other hand, refers to an individual's belief in their ability to successfully perform specific tasks related to their job (Rigotti et al., 2008).

Little is known about how LADs with different reference frames influence the self-reflection phase and self-efficacy in the context of VR simulation-based training environments for the workplace. This knowledge is important because digital learning environments utilising LADs to stimulate SRL, with the self-reflection phase being a core component, are growing in popularity (Valle, Antonenko, Dawson, et al., 2021) and self-efficacy has been shown to be a powerful determining factor of job performance (Bandura & Locke, 2003; Carter et al., 2018; Song et al., 2018). Therefore, this study investigates the effects of two LADs designed with different reference frames in a VR simulation-based training environment. We first discuss the potential of LADs and learning analytics reference frames, followed by an introduction to Zimmerman's Cyclical Phases Model of SRL (Zimmerman, 2002; Zimmerman & Moylan, 2009) and the self-reflection phase's role within this cycle. We then discuss Bandura's theory of self-efficacy (Bandura, 1997) and present competing hypotheses for two research questions proposing an effect of learning analytics reference frames on occupational self-efficacy, mediated by the self-reflection phase of the Cyclical Phases Model of SRL.

3.1.1 Learning Analytics Dashboards and Reference Frames

Learning analytics is "the measurement, collection, analysis and reporting of data about learners and their contexts, for purposes of understanding and optimising learning and the environments in which it occurs" (Long, Siemens, Conole, & Gasevic, 2011, p. 3). Effective use of learning analytics tools can support SRL in the workplace (Ruiz-Calleja et al., 2017). This is crucial as employees must continually self-regulate their learning processes and adapt to changing workplace demands, such as acquiring new knowledge and skills for tasks (Fontana et al., 2015). For example, in the chemical process industry, operators must learn new procedures when producing a new compound. Learning analytics can provide operators automatic feedback on the procedural steps they know and still need to learn (Garcia Fracaro et al., 2021), which has the potential to help them improve task proficiency and reduce the time taken to complete them. Additionally, it can potentially help them identify areas where they may need further instruction or assistance. Furthermore, because learning analytics feedback is an automated process, workplace learners become less reliant on their trainers to support them in reaching their workplace goals. One tool for delivering such feedback is through LADs with reference frames.

LADs are a “single display that aggregates different indicators about learner(s), learning process(es) and/or learning context(s) into one or multiple visualisations” (Schwendimann et al., 2017, p. 8). When designed specifically for learners, LADs aim to provide useful feedback and support for SRL. By doing so, they enhance both domain-generic and domain-specific skills, thereby aiming to improve overall learning performance (Lim, Joksimović, Dawson, & Gasevic, 2019; Valle et al., 2021). A key aspect of LAD design is incorporating reference frames to help learners better make sense of their feedback (Wise & Vytasek, 2017). Reference frames are “the comparison point to which students orient when they examine their learning analytics” (Wise, 2014, p. 6). We extend this definition and argue that learning analytics reference frames consist of three main components which assist learners in making sense of their learning analytics data: (1) ‘performance outcome’, which represents the learners' achievement on a task, (2) ‘point of comparison’, which provides a context to the performance outcome by comparing it against another relevant data point including but not limited to historical performance, peer performance, or set goals, and (3) ‘score delta’, which refers to the difference between the performance outcome and the point of comparison. Furthermore, the nature of the 'point of comparison', whether it is based on self-comparison over time, comparison with peers, alignment with curriculum expectations, or proximity to set goals, determines the type of the reference frame.

Transitioning to the practical application of these components, it is important to examine each one more closely. The performance outcome component indicates how well a learner has performed a specific task for which they receive feedback. For example, if a learner completes a simulation-based training task and receives feedback indicating that they completed 75% of the operational steps correctly, their performance outcome will be the same. However, the performance outcome alone is limited in usefulness when evaluating one's own performance as it lacks context.

The point of comparison component helps learners evaluate their performance outcome by providing context (Wise & Vytasek, 2017). When historical performance data is used as a point of comparison, learners receive a progress reference frame. For example, a learner's performance outcome on a previous attempt at the same or a related simulation-based training task can be used to contextualise their most recent performance outcome. The progress reference frame encourages learners to engage in temporal comparisons.

Temporal comparison theory (Albert, 1977; Wilson & Shanahan, 2020) posits that learners find comparing current performance with past results useful (Möller & Marsh, 2013; Wilson & Shanahan, 2020). This theory is strongly tied to the progress reference frame via two propositions. First, during periods of rapid change, learners evaluate change in their ability via temporal comparisons. When feedback encapsulates multiple time points, learners can more readily discern whether the desired shift in ability has taken place. Second, temporal comparison theory proposes that individuals have a desire to predict future outcomes. By examining past performance and comparing it with current performance, learners can gauge their proximity to their learning goals.

When peer performance is used as a point of comparison, learners receive a social reference frame (Jivet, Scheffel, Specht, & Drachsler, 2018). Take, for instance, the mean correct steps taken by trainees in a similar simulation-based training task. This serves to place a learner's own performance in perspective. LADs designed with a social reference frame can activate social comparison, a theory suggesting individuals assess their capabilities by comparing themselves with others (Festinger, 1954; Gerber, Wheeler, & Suls, 2018; Mcintyre & Eisenstadt, 2011; Suls, Martin, & Wheeler, 2002; Wood, 1996). This proves particularly helpful for learners who lack reliable, objective feedback, as gauging one's task-specific ability becomes feasible through peer comparison. In workplace learning environments, where performance feedback may be scarce, this proves particularly advantageous. Employees, by drawing comparisons with peers, can pinpoint their strengths and identify areas where they need improvement.

The score delta, the third learning analytics reference frame component, further assists learners in making sense of their feedback. This component represents a variable linked to the LAD reference frame design, mirroring the difference between the performance outcome and the point of comparison. Consequently, even with identical performance outcomes, score deltas can differ. For example, suppose a learner receives feedback indicating a performance outcome of 75% and a point of comparison derived from historical performance data indicating 60%. In this case, the score delta is a positive 15%, which the learner may perceive as substantial progress. However, if the point of comparison, drawn from the average performance of peers, stands at 85%, the score delta becomes a negative 10%. Although the performance outcome remains at 75%, the learner may interpret this as underperformance. This illustrates that a meaningful portion of a learner's interpretation of feedback is influenced by factors beyond their control. Factors such as the LAD design and the applied reference frame have the potential to significantly influence interpretation, apart from the task performance itself. Therefore, if our goal is to optimise the design of VR simulation-based training environments, it is worthwhile investigating how different learning analytics reference frames influence learner interpretation of feedback, which depends largely on the self-reflection phase of the SRL cycle (Zimmerman & Moylan, 2009).

3.1.2 Self-regulated learning and the Self-reflection phase

SRL refers to the process of actively controlling one’s thoughts, feelings, emotions and behaviour in relation to learning (Panadero, 2017; Zimmerman, 2002). SRL is vital in the workplace because there is a continuous need for employees to keep up to date with changing workplace demands, such as the acquisition of new knowledge and skills required to complete workplace tasks (Fontana et al., 2015).

SRL consists of three phases: forethought, performance, and self-reflection (Zimmerman, 2002). The forethought phase occurs before task performance and involves learners engaging in strategic planning and goal setting. The performance phase occurs during task performance and includes learning processes such as self-monitoring and help-seeking. The self-reflection phase involves learners judging their performance and formulating reasons for their results. They may do this by comparing their performance outcomes with available points of comparison, such as a standard or goal (Panadero, 2017).

As Zimmerman and Campillo (2003) suggest, during the self-reflection phase, learners judge their performance against different criteria of success including mastery, previous performance, and/or normative criteria. Mastery criteria offer absolute measures of success and help learners assess whether a task has been executed correctly or incorrectly, and to what degree. Previous performance criteria offer a relative measure of success and lead learners to compare their current performance to past performance levels (temporal comparisons). Normative criteria also offer a relative measure of success and lead learners to evaluate themselves based on the performance of others (social comparisons).

In the context of learning analytics reference frames, both the progress and social reference frame offer mastery criteria through their performance outcome component. The progress reference frame offers previous performance criteria through its point of comparison component and the social reference frame delivers normative criteria through its respective comparison point. Consequently, learners who receive a progress reference frame have access to both mastery criteria and previous performance criteria, shaping their self-reflection phase processes. On the other hand, learners receiving a social reference frame gain access to mastery criteria and normative criteria. As a result, learners in the self-reflection phase of the SRL cycle may interpret their feedback differently, depending on whether their LAD incorporates a progress or social reference frame.

Exploring the implications of self-reflection phase processes driven by either a progress or social reference frame is important, particularly in the context of workplace-based training. This is because the way in which learners encounter LADs with different reference frames can influence their perception of their own ability to perform workplace tasks, also known as occupational self-efficacy.

3.1.3 Occupational Self-efficacy

Self-efficacy is the belief an individual has in their capacity to successfully complete a task (Bandura, 1977, 1997, 2012). Longitudinal field-based research (Carter et al., 2018) and multiple meta-analyses (Cherian & Jacob, 2013; Sadri & Robertson, 1993; Stajkovic & Luthans, 1998) have shown self-efficacy to be positively related (correlations around r = .4) to: work performance, (e.g., meeting the formal performance requirements of a job), work behaviour choice (e.g., propensity to change rather than accept organisational processes) and work-related performance (e.g., achieving predetermined task goals within an organisation). Each of which is associated with our central variable of interest: occupational self-efficacy, which refers to “the competence that a person feels concerning the ability to successfully fulfil the tasks involved in his or her job.” (Rigotti et al., 2008, p. 2). However, while self-efficacy can drive performance, an inflated sense of one's capabilities can present challenges. As noted by Talsma et al (2019), an overestimation of self-efficacy in the workplace may give rise to a discrepancy between perceived and actual performance. Such a mismatch can result in inadequate task and effort allocation, unrealistic goal setting, and ultimately, subpar workplace performance.

Self-efficacy affects learning and performance through its influence on task selection and self-set goals (Bandura, 1997), the effort one puts into workplace tasks (Fejoh et al., 2018; Leon-Perez et al., 2011; Pan et al., 2011) and the persistence with which learners attempt new and difficult tasks (Lunenburg, 2011). Individuals develop their self-efficacy beliefs by analysing information about their capabilities (Bandura, 1997). Two major sources of information which inform this analysis are mastery experiences information and social modelling information (Bandura, 1997). Panadero et al. (2017) conducted a meta-analysis from 19 studies on the effect of self-assessment, a self-reflection phase process, on self-efficacy beliefs and found a substantial (d=.73) combined effect size. This evidence suggests that interventions prompting learners to engage in self-reflection can impact self-efficacy. However, a recent study by Lishinski and Yadav (2021), not included in the meta-analysis, yielded more modest results. Participants were asked to reflect on feedback, identifying what they did well and areas for improvement, while exposing them to mastery criteria and mastery experience information. The study found that self-efficacy beliefs were only marginally affected by self-reflection phase processes.

Mastery experience information is considered the most influential source of self-efficacy information because it offers direct evidence of one's ability to perform a particular task (Bandura, 1997). This evidence is gathered through the direct experience of performing a task. Success can build a robust sense of self-efficacy and failure can undermine it. Conversely, social modelling information shapes these beliefs through the act of social comparison (Bandura, 1997; Gerber et al., 2018; Mcintyre & Eisenstadt, 2011; Suls et al., 2002; Wood, 1996). This process is especially relevant in the workplace, where objective performance measures may be elusive (Bandura, 1997). Employees, therefore, often gauge their proficiency by comparing their performance to their peers during the self-reflection phase. The recognition of their ability to perform workplace tasks more proficiently than their peers may enhance their self-efficacy beliefs, while inferior levels of performance potentially diminish them. Bandura (1997) argues that social modelling is less a dependable source of information compared with mastery experiences, because unlike mastery experiences information, it is not informed by direct evidence of personal achievement.

On the topic of social modelling information and its influence on self-efficacy beliefs, Adams (2004) investigated the impact of an intervention on self-efficacy among postgraduate students. The intervention involved self-reflection phase processes, which exposed the students to normative criteria and social modelling information. This exposure occurred through the observation of peer performances specifically related to public speaking. Evidence from this study suggests that these interventions can influence self-efficacy beliefs. In other research, Karaoglan Yilmaz (2022) reported greater changes in self-efficacy among university students who received feedback via an LAD than their counterparts without such feedback.

Viewed through the lens of learning analytics reference frames, the progress reference frame offers mastery experience information through mastery criteria and past performance criteria. Conversely, the social reference frame provides mastery experiences information through mastery criteria but supplements it with social modelling information via normative criteria. Therefore, it stands to reason that learners receiving feedback via an LAD with different reference frames will interpret that feedback differently, thereby influencing their occupational self-efficacy in distinct ways.

In the context of online university education, Rets et al. (2021) investigated the perceived usefulness of learner facing LADs, particularly focusing on the influence of social comparison features. Their results indicated that participants found these features less useful than those offering temporal comparison, but many maintained a positive outlook on the social comparison functionality. This underscores the importance of investigating the effects of LADs providing different comparison types.

Furthermore, substantial theoretical and empirical evidence suggests that self-reflection phase processes can influence self-efficacy, with mastery experiences information and social modelling information playing key roles (Bandura, 1997). Therefore, there is value in investigating these claims in the context of a VR simulation-based training environment for the chemical process industry. Given the growing popularity of VR simulation-based training environments in this industry, and the autonomy that LADs provide to learners (Garcia Fracaro et al., 2021)., the potential benefits of exploring these claims in this specific setting are evident.

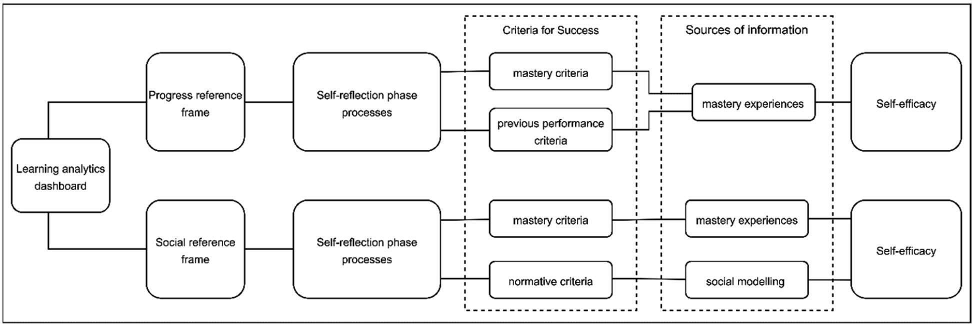
Although to our knowledge there is no direct evidence that different reference frame designs affect occupational self-efficacy in the context of a VR simulation-based training environment, previous research into the effect of LADs with different reference frames offers valuable insights. Davis et al. (2017) found that LADs with social comparison features delivering feedback for a Massive Open Online Course caused desirable changes in learner engagement. Furthermore, a study by Brusilovsky et al. (2016), showed that learners responded more positively—in terms of attitude, engagement, and performance—when LADs offered both progress and social reference frames. However, Jonathan et al. (2022) found that an LAD with a social reference frame and an LAD with a progress reference frame similarly affected critical reading self-efficacy among high school students.

Considering these varied findings and acknowledging the lack of studies on the effects of different LAD designs in the context of workplace VR simulation-based training, our study aims to investigate the impact of progress and social reference frames on occupational self-efficacy.

Figure 1 provides a visual representation of two presumed feedback pathways from an LAD with either a progress or social reference frame, each leading to different self-reflection phase processes and sources of self-efficacy information. The Progress Reference Frame pathway offers mastery experience information based on mastery and previous performance criteria, while the Social Reference Frame pathway provides social modelling and mastery experience information, based on mastery and normative criteria. Both pathways converge on influencing learners' self-efficacy.

Figure 1

Learning Analytics Dashboard Feedback Pathways and Their Influence on Occupational Self-Efficacy



3.1.4 Study Overview, Research Question and Hypotheses

Investigating the influence of progress and social reference frames on occupational self-efficacy, we designed a two-group experimental study. Participants undertake a VR simulation-based training task and receive feedback via an LAD with either progress or social reference frame. Workplace SRL self-reflection phase processes are controlled using a pre-test.

This research draws on established theories, namely Social and Temporal Comparison Theory, Self-Regulated Learning Theory, and Self-Efficacy Theory, to inform our experimental design and guide our hypotheses. As such, the study aims to extend our understanding of the differential effects of two distinct LAD designs on occupational self-efficacy, thereby making contributions to both theoretical (e.g., educational psychology and learning analytics) and practical domains (e.g., industrial training programs). Notably, to our knowledge, the differential impact of these reference frames on occupational self-efficacy has yet to be explored, marking a clear research gap.

We aim to determine if the reference frames differentially impact occupational self-efficacy and, if an impact is observed, to identify which specific reference frame exerts the most substantial influence. Our primary measure is the absolute change in occupational self-efficacy, assessed by calculating the absolute differences between pre-test and post-test scores. We follow this with an exploratory analysis of the direction of these changes, assessed by calculating the difference between pre-test and post-test scores.

Knowing both the absolute change and direction of change in self-efficacy is important for LAD designers and stakeholders aiming to influence self-efficacy levels. The absolute change shows the magnitude of impact a particular reference frame can have, while the direction of change (positive or negative) provides insight into the nature of impact, aiding in the effective design of LAD tools to manage self-efficacy dynamics.

While a positive change in self-efficacy can often foster enhanced performance and engagement, it may also lead to overconfidence. This is particularly relevant in the chemical industry, where overconfidence can have severe consequences. For instance, an overconfident chemical plant operator, emboldened by past successes, could potentially begin to disregard safety protocols, leading to an increase in hazardous behaviour. Previous research has highlighted the dangers of overconfidence, such as reduced task effort and persistence with failing strategies (Audia et al., 2000; Nilsen & Campbell, 1993). Talsma et al. (2019) further substantiate this, demonstrating that self-efficacy is not a self-fulfilling prophecy; rather, inflated self-efficacy can negatively influence aspects of SRL and performance.

Consequently, our dual focus on both absolute and directional changes in self-efficacy allows for a more comprehensive understanding of the impact of LAD reference frames on occupational self-efficacy, enriching dialogues across both academic and industry contexts.

To answer our primary research question, we apply Bayesian informative hypothesis evaluation with three competing hypotheses related to type of learning analytics reference frame and absolute change in occupational self-efficacy.

Research Question: When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in total change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?

The following three competing hypotheses are informed by the theoretical framework and empirical results presented above:

Hypothesis 1 (H1): The progress reference frame will elicit greater changes in occupational self-efficacy than the social reference frame.

During the self-reflection phase, the progress reference frame provides learners with more mastery experiences information, facilitated by both mastery and previous performance criteria. On the other hand, the social reference frame offers learners mastery experiences information facilitated only by mastery criteria, supplemented by social modelling information via normative criteria. Given the greater influence of mastery experiences information over social modelling information on self-efficacy, we predict a more pronounced impact from the progress reference frame.

Hypothesis 2 (H2): The social reference frame will elicit greater changes in occupational self-efficacy than the progress reference frame.

During the self-reflection phase, learners with the social reference frame derive both mastery experiences information and social modelling information from mastery and normative criteria. In contrast, the progress reference frame provides learners only with mastery experiences information, facilitated by mastery and previous performance criteria. Given the additional source of social modelling information in the social reference frame, we predict it will exert a greater impact on occupational self-efficacy change.

Hypothesis 3 (H3): The effect of the progress and social reference frames on changes to occupational self-efficacy is equal, which may include no impact at all.

Although the progress and social reference frames differ, neither offer a substantially strong enough influence on occupational self-efficacy in the context of this research. This could be due to the overwhelming strength of the effect of the VR simulation-based training environment on occupational self-efficacy beliefs. Alternatively, it could be the case that the VR simulation-based training environment in its current design fails to influence occupational self-efficacy at all.

Building on our primary focus on absolute changes, we conduct an exploratory analysis to examine the direction of change in occupational self-efficacy between the two groups. Note that in our context a positive direction of change may not always be desirable due to the risks attached to overconfidence in contexts like the chemical industry.

Exploratory Research Question: When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in direction of change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?

Informed by the same theoretical framework as our primary hypotheses, our exploratory hypotheses focus on the directional impact of progress and social reference frames on occupational self-efficacy.

Exploratory Hypothesis 1 (eH1): The progress reference frame will elicit a more positive direction of change in occupational self-efficacy than the social reference frame.

Exploratory Hypothesis 2 (eH2): The progress reference frame will elicit a less positive direction of change in occupational self-efficacy than the social reference frame.

Exploratory Hypothesis 3 (eH3): The direction of change in occupational self-efficacy elicited by the progress and social reference frames is equal, which may include no directional change at all.

3.2 Methods

3.2.1 Preregistration

This study's hypotheses, design plan, sampling plan, variables and analysis plan were preregistered on the Open Science Framework prior to any data being analysed (Gallagher, 2021).

3.2.2 Participants

Participants (N=42) were employees of an international science and technology company located in Germany. This company was an industry beneficiary of the Horizon 2020 funded European Training Network for Chemistry Engineering Immersive Learning project (Grant Agreement 812716).

The participants had at least one year experience working as chemical process operators and were aged between 18 and 55 years. See Table 1 and Table 2 for their age and experience level, respectively. Participants had no prior experience manufacturing n-Butyllithium. A total of 40 participants were male and two were female.

Table 1

Age of Participants

Age	Total	Progress Group	Social Group
20 - 30 years	4	2	2
31 - 40 years	15	10	5
41 - 50 years	9	4	5
51 - 60 years	12	4	8
60 +	2	1	1
Total	42	21	21

Table 2

Experience Level of Participants

Years of Experience	Total	Progress Group	Social Group
1 - 5 years	1	1	0
6 – 10 years	9	4	5
11 - 15 years	9	5	4
16 - 20 years	11	7	4
More than 20 years	12	4	8
Total	42	21	21

Of these participants, prior experience with VR varied: Five members of the Progress group and two members of the Social group reported previous exposure to VR. The remainder of the participants had no prior experience with VR.

Convenience sampling was relied upon to select the participants who volunteered and signed informed consent after reviewing an information letter explaining the research. The employees' workers council approved the study instruments. Participation in the training was optional and participants could exit the study without consequence. The participants' working language was German, and a German language version of the prototype was used.

3.2.3 Experimental design

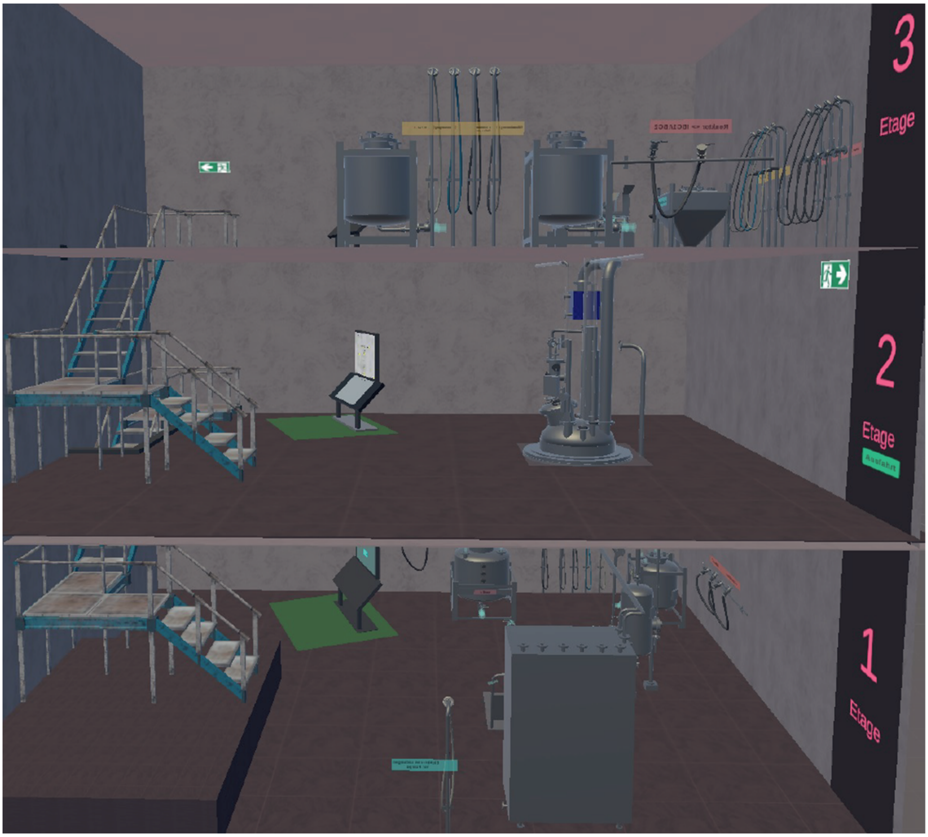
The study followed a two-group design with participants (N=42) randomly allocated to either the progress (n=21) or social (n=21) reference frame group. The effect of LADs designed with a progress reference frame was tested against the effect of LADs designed with a social reference frame on occupational self-efficacy. Participant SRL self-reflection at work was analysed as a potential confounding variable.

3.2.4 Training with 'Operate your own reactor'.

The VR training simulator, 'Operate your own reactor', is a design prototype developed for the Oculus Quest VR head mounted display with Oculus Touch controllers (Tehreem et al., 2022). It aims to train chemical process operators in manufacturing n-Butyllithium within a chemical process plant setting with commercial chemical manufacturing equipment (Figure 2). The training procedure takes approximately 60 minutes to complete (Tehreem et al., 2022). This simulated learning environment provides a risk-minimised platform for training due to the highly volatile nature of n-Butyllithium production.

Figure 2

Operate your own reactor VR environment.



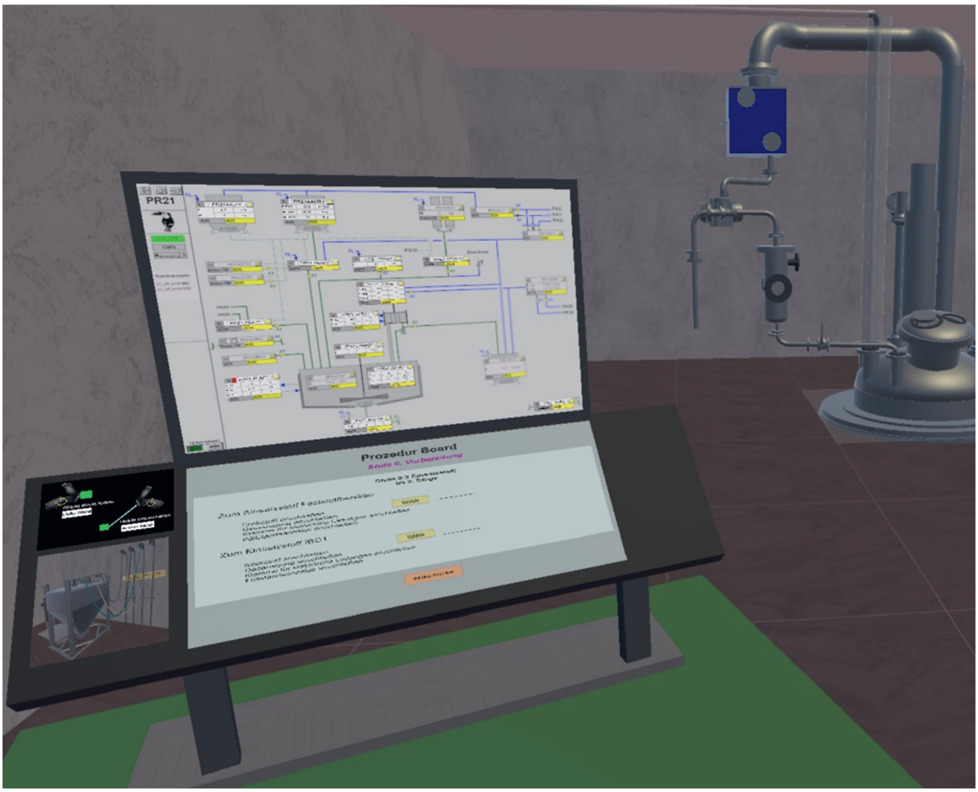
The training simulator and procedure underwent validity and efficiency checks in a previous pilot study with German-speaking employees from the same company where this research took place (Tehreem et al., 2022), aligning the training with participants' expectations. User comfort was assessed, with minimal symptoms of VR sickness reported (Tehreem et al., 2022).

The procedure consists of four steps representing different stages of n-Butyllithium chemical production. The first three steps constitute the training phase, offering built-in support in the form of automatically provided hints to guide learners. The last step, classified as the evaluation phase, also provides built-in support but only offers hints upon learner request. In the evaluation phase, the instructions provided by the system are less detailed compared to the training phase.

In the VR environment, learners navigate three interconnected levels (Figure 2) equipped for n-Butyllithium preparation. Two types of interactions occur: physical tasks with digital instruments for producing the desired chemical reaction, and digital tasks with a digital monitor that controls the instruments and displays a procedure board outlining the manufacturing steps. The procedure board provides sufficient detail to correctly guide skilled operators. Less skilled operators may misinterpret the steps and subsequently make mistakes. The design reflects the layout of the industrial partner's chemical plants. Instructions for VR controllers and tasks are displayed on the monitor's left side (Figure 3).

Figure 3

Operate your own reactor digital instruments and digital monitor with procedure board.



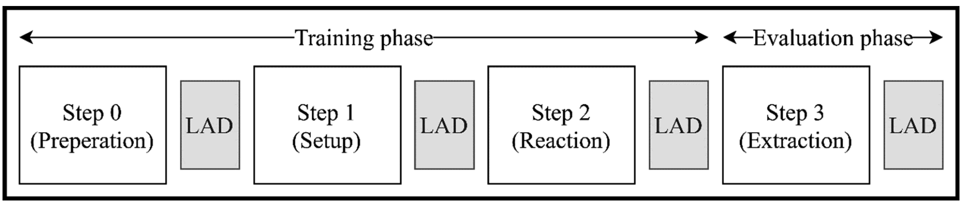
A learning analytics system automatically collects and analyses task feedback data on the learners' actions within the simulated environment. This system was developed in collaboration with experts in the field of assessment and chemical process operator training (Garcia Fracaro et al., 2021; Tehreem et al., 2022). The data collected and analysed during the training phase includes number of mistakes and time spent to complete the task. Data during the evaluation phase similarly accounts for number of mistakes, time spent and the number of requested hints. Hints were automatically provided during the training phase, which is why hint data were not collected during that phase. It is worth noting that we opted not to control for the number of hints requested during the evaluation phase. This decision was informed by two considerations. First, the number of hints could potentially be influenced by the design of the reference frames in the LADs, making it unnecessary to control for this variable in the context of our research aims. Second, the random assignment of participants to experimental groups and the controls for SRL self-reflection phase process were designed to minimise potential confounding variables.

3.2.5 Timing and Features of the LADs

Learners receive task feedback on their performance after each step of the task in the form of an LAD (Figure 4). This LAD is presented within the simulated environment. Therefore, each participant receives task feedback via LADs on four occasions by the time they have completed the training.

Figure 4

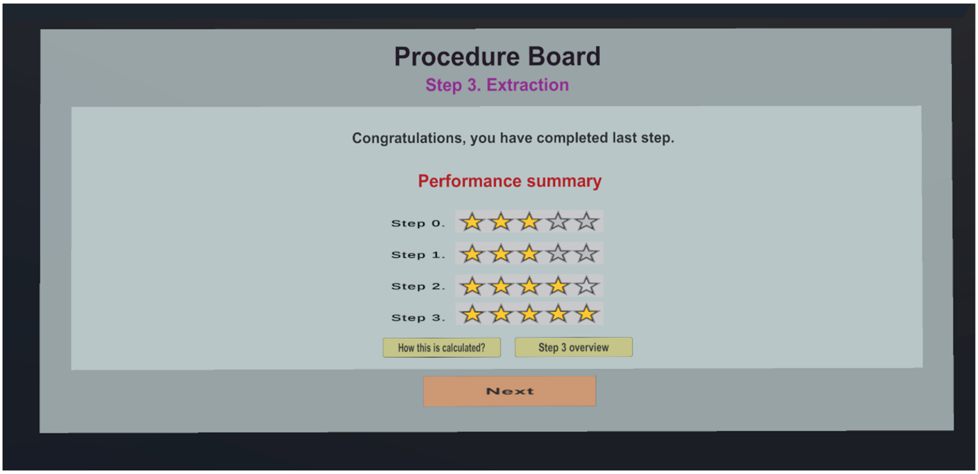
The timing of Learning Analytics Dashboard presentation with steps and phases



Each LAD is made up of six main features (Figure 5): (1) name of the step which has just been completed, (2) message congratulating the participant on completing the step, (3) summary of the participants performance with an accompanying reference frame (progress or social, depending on the group), (4) 'How this is calculated?' button, (5) 'Step overview' button and (6) 'Next' button.

Figure 5

Screenshot of the Progress Reference Frame presented to a participant after step 3 (Extraction).



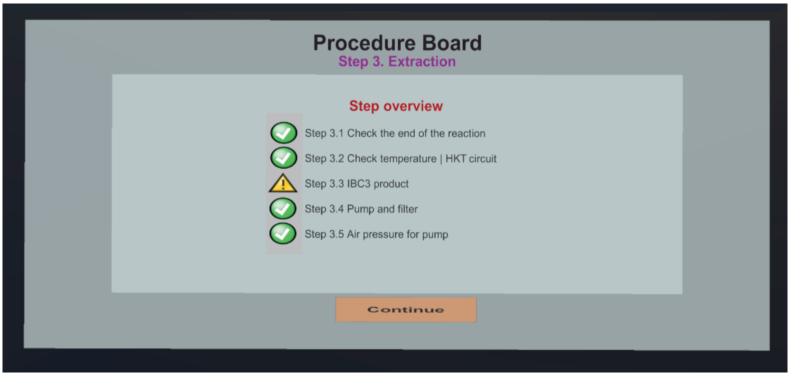
Note. Learners are presented with each step's performance outcome as they progress through the task. In this example, an LAD with a progress reference frame, learners can compare how they performed on step 3 with each previous step.

The performance summary feature is a means of communicating how well the learner performed. It uses a star system designed to be quickly and easily understood by the learner reviewing it. The greater number of stars awarded, the better the performance, ranging from one to five stars. Only whole stars can be awarded. Additional information regarding the star system can be found below.

The 'Step overview' button (Figure 5) and the 'How this is calculated?' button (Figure 5) can be selected by the participants if they want additional information regarding their performance outcome. If either of these buttons are selected, additional context is provided about the feedback they received. Figure 6 provides an example of the resulting 'Step overview' visualisation. Figure 7 shows the 'How this is calculated?' visualisation (blue text).

Figure 6

Screenshot of the 'Step overview' visualisation after Step 3.

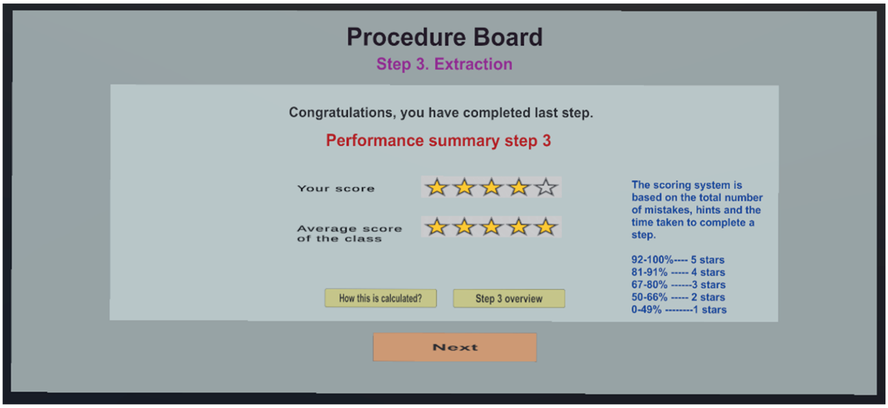


Note. Green ticks indicate correctly executed sub-tasks and yellow exclamation marks indicate incorrectly executed sub-tasks. A description of each sub-task is also visible (e.g., Check the end of the reaction).

When selected, the 'Step overview' button displays an additional page of the LAD which outlines the performance outcome of the participant on the task at a sub-task level. The assessment framework is described in more detail below to provide additional information pertaining to the sub-task level. When selected, the 'How this is calculated?' button displays a basic formula for calculating the star system (i.e., 92-100% awards five stars).

Figure 7

Screenshot of the social reference frame with the 'How this is calculated?' indicator displayed alongside Step 3 (Extraction).



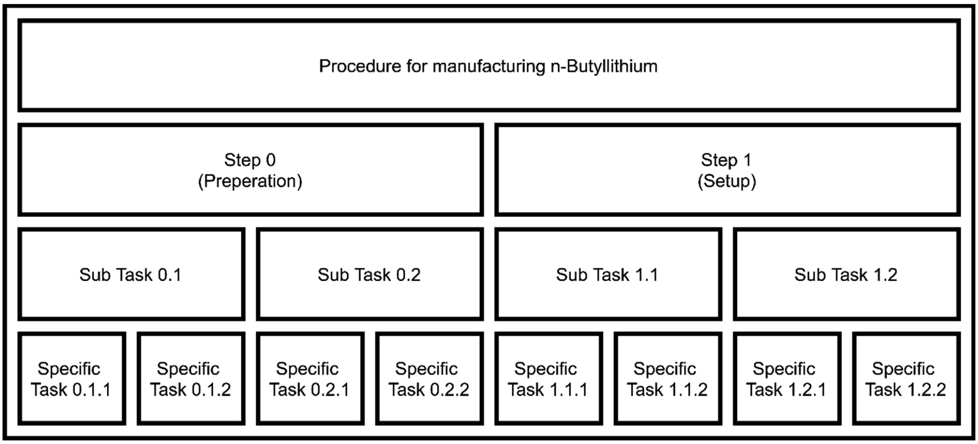
The LADs implement a star-based rating system, awarding participants between 1 and 5 stars according to their performance outcome. The assessment formula, developed collaboratively with chemical engineers, educational researchers and chemical process plant operator trainers from the industry partner, differs between the training and evaluation phases. In the training phase, the number of stars is determined based on the number of mistakes and the time taken to complete each step. In the evaluation phase, the number of stars is determined by considering the number of mistakes, hints used, and time to complete the task. Hints (e.g., "Click on ALAC1. Set internal temperature to 20 °C and press start.") were automatically provided during the training phase but are not included in its assessment formula. During the evaluation phase, hints were available upon request. Hints, mistakes, and time carry equal weight in the assessment.

3.2.6 Assessment framework and task levels

The assessment framework was designed in collaboration with trainers from the company in which this research was conducted and researchers developing the software (Garcia Fracaro et al., 2021; Tehreem et al., 2022) and includes descriptions of different levels of tasks ranging from general to specific tasks. The four levels of tasks are: Overall tasks (e.g., complete reaction phase), main tasks (e.g., disconnect hose connection), sub-tasks (e.g., close valve and disconnect hose from reactor 1 to reactor pipe), and specific tasks (e.g., turn off ball valves) (Figure 8). Based on the proficiency of the user in completing the tasks, the system determines a performance outcome. This performance outcome is then communicated to the learner using the star system. More detailed task feedback is presented within the LAD to those learners who wish to receive it via the 'Step overview' button (Figure 6). It is here where the learner can review the sub-tasks they completed correctly or incorrectly. Sub-tasks, which consist of multiple specific tasks, are marked as correct when no errors at the specific task level are made.

Figure 8

Simplified illustration of procedure and task level hierarchy for steps 0 and 1 (preparation and setup).



Note. Naming convention specifies that the first step in this process (preparation) is Step 0.

3.2.7 The progress and social reference frame LADs

The LADs provide task feedback to learners upon completing each step. The progress reference frame group received LADs with a progress reference frame (Figure 5), while the social reference frame group received LADs with a social reference frame (Figure 7). For the progress reference frame, prior step performance data was processed locally on the VR device. However, the LAD presented after the first step in the progress reference frame group lacks a point of comparison due to it being the initial task attempt without historical data. From step 1 (Setup) onward, all progress reference frame components are presented. In contrast, the LADs for the social reference frame group include all components after each step. The progress reference frame group completed the training before the social reference frame group, allowing us to use their performance data to calculate the social reference frame point of comparison.

3.2.8 Measures

n-Butyllithium Procedure Competence Levels

Descriptive data on participants' performance in each step of the procedural task were collected through VR device log-files. These data include the time taken to complete each step, including breaks, the number of mistakes, and the number of hints requested in step three. Additionally, the number of breaks taken by each participant during training was manually recorded.

Occupational Self-efficacy Questionnaire

Self-efficacy is operationalised as occupational self-efficacy because it is better suited to measure participant self-efficacy in the context of workplace learning. The German version of the Short Occupational Self-Efficacy Scale (Rigotti et al., 2008) was used to measure participant self-efficacy before and after the task. The scale was developed for use in the workplace to measure employee occupational self-efficacy. It was validated in five different countries (N =1,535), one of which being Germany (n =200), in German the reliability was found to be good with Cronbach’s alpha (α) = .86 and the validity measures met the threshold for good model fit (Rigotti et al., 2008). All six items from the Short Occupational Self-Efficacy Scale questionnaire (e.g., I feel prepared for most of the demands in my job) were used in which participants responded on a six-point Likert scale ranging from 1 (not at all true) to 6 (completely true). High values reflect high levels of occupational self-efficacy. The

Occupational Self-Efficacy Questionnaire was administered as part of the pre-test and post-test. We will report in the results section a confirmatory factor analysis for validity and the estimated reliability of the instrument using Cronbach's alpha. Similar studies have shown that VR simulation-based training environments on self-efficacy have been influential on self-efficacy beliefs within a training single session (Hough et al., 2019).

Our analytical approach addresses both the primary and exploratory research questions by examining the absolute and directional changes in occupational self-efficacy, respectively. This is in line with our central objective of determining whether the reference frames differentially impact occupational self-efficacy. By focusing on both absolute and directional changes, we underscore the importance of assessing the overall effectiveness of LADs in inducing shifts in learners' self-efficacy, irrespective of whether these shifts represent an increase or a decrease.

Self-regulated Learning at Work – Self-reflection Questionnaire

To control for group differences in self-reflection phase processes at work, the self-reflection scale of the validated Self-regulated Learning at Work Questionnaire ($\alpha = .86$) (Fontana et al., 2015) was used. This measure examines self-reflection phase processes that can influence self-efficacy change. The self-reflection phase items (e.g., I try to understand how new information I‘ve learned impacts my work.) from the Self-regulated Learning at Work questionnaire were administered as part of the pre-test.

3.2.9 Procedure

Participants were invited to a training session with the 'Operate your own reactor' prototype during working hours. Upon acceptance, collaborating researchers introduced them to the research and provided an information sheet. Those agreeing to participate were then asked to sign an informed consent form.

Upon obtaining informed consent and just before entering the VR simulation-based training, all participants (N=42) completed several questionnaires and a prior knowledge test on n-Butyllithium production. The questionnaires included demographic information, prior VR experience, the Self-regulated Learning at Work - Self-reflection questionnaire, and the Short Occupational Self-Efficacy Scale questionnaire. For this research, only the last two questionnaires were used. Next, participants were made familiar with the VR hardware required to interact with the Operate your own reactor prototype to ensure they could use it competently. They were informed of the option to take breaks after each step throughout the training.

Upon completing the training, participants re-took the Short Occupational Self-Efficacy Scale questionnaire, followed by a knowledge post-test and a VR presence and cybersickness questionnaire. Only the data from the Short Occupational Self-Efficacy Scale was used for this research. The entire procedure, with a 55-minute average training task and up to an hour for the surrounding questionnaires and tests, took less than two hours.

3.2.10 Data analysis

Bayesian informative hypothesis evaluation

Our data analysis used a method of Bayesian informative hypothesis evaluation. Informative hypotheses are theoretically and empirically informed statements about a phenomenon. In the case of this research, the informative hypotheses pertain to both our primary research question (H1, H2 and H3) and exploratory research question (eH1, eH2, eH3). Informative hypotheses are formulated using terms of equality (=) and inequality (<, >) (Van Lissa et al., 2021). One benefit of this approach over classical null hypothesis testing with p values is that it enables the comparison of multiple hypotheses (Hoijtink et al., 2019; Van Lissa et al., 2021).

Manipulated variables

We manipulated the type of reference frame used when presenting feedback within the LAD. The two reference frame types which act as variables are the progress reference frame and the social reference frame.

Statistical models and Analysis plan

JASP version 0.16 is used for all statistical analyses. For both the Self-regulated Learning at Work – Self-reflection questionnaire and the Occupational self-efficacy questionnaire, unidimensional reliability tests were used to calculate Cronbach's α and confirmatory factor analyses were used to determine validity. We set the threshold for good reliability for α at > 0.70 (Cortina, 1993; Taber, 2018). Four statistical tests are used to measure validity: the Root Mean Square Error of Approximation (RMSEA) with values below .06 indicating good fit, the Standardised Root Mean Residual (SRMR) with values below .05 indicating good fit, the Comparative Fit Index (CFI) with values above .90 as indicating good fit and the Tucker Lewis Index (TLI) with values above .95 as indicating good fit (Hu & Bentler, 1999).

For the primary research question, we compare the mean of the progress reference frame group change in occupational self-efficacy with the mean of the social reference frame group change in occupational self-efficacy using the Bayes factor as implemented in 'bain' (Van Lissa et al., 2021). To do so we conducted an ANOVA with change to occupational self-efficacy as the dependent variable and reference frame groups as the fixed factors. Change to occupational self-efficacy is equal to the post-test occupational self-efficacy score minus the pre-test occupational self-efficacy score, represented as an absolute value. We will do a sensitivity analysis using fraction 1, 2 and 3 and will report each result (the posterior model probabilities (PMPs)) and interpret them at once. The higher the PMPs the more support for the associated hypothesis. The Bayesian error associated with preferring the best hypothesis in terms of PMPs will be reported. This is the sum of the PMPs of the other hypotheses. For example, the PMPa of H1 plus the PMPa of H2 is equal to the error probability of H3, that is the likelihood H3 is not the best hypothesis compared to H1 and H2.

For the exploratory analysis concerning the direction of change in occupational self-efficacy, we conducted an ANOVA with the direction of change as the dependent variable and reference frame groups as the fixed factors. The direction of change is calculated as the post-test score minus the pre-test score. The exploratory hypotheses (eH1, eH2, and eH3) will be evaluated using the Bayesian informative hypothesis evaluation method as with the primary hypotheses.

3.3 Results

3.3.1 n-Butyllithium Procedure Performance Outcomes and Breaks

Table 3 presents performance data of each of the groups on each step of the procedural task. It reports the mean number of minutes each group spent completing each step including breaks, the mean number of mistakes made by each group on each step and the mean number of hints each group requested during Step 3.

Table 3

Group Performance Outcomes

Mean Group Performance Outcomes for n-Butyllithium Procedure									
	Time in minutes				Mistakes				Hints
	Step 0	Step 1	Step 2	Step 3	Step 0	Step 1	Step 2	Step 3	Step 3
n	21	21	21	21	21	21	21	21	21
Progress	10.41	16.08	15.43	11.41	0.67	1.38	1.52	0.91	0.95
	(4.60)	(4.04)	(3.84)	(3.20)	(1.96)	(2.29)	(1.81)	(1.45)	(3.28)
Social	11.60	18.78	17.22	13.40	0.91	2.05	2.43	2.05	1.24
	(4.82)	(7.23)	(5.69)	(4.97)	(1.73)	(2.87)	(3.28)	(2.85)	(2.36)

Note. Number in brackets equals standard deviation from the mean.

Participants were able to request breaks. The time spent taking a break was included in the total time of the following step. Once a break was requested, they were asked to finish the step they were working on before removing the head mounted display. In total 10 participants took one break, five in each group, and one participant in each group took two breaks. The total time spent on breaks was not recorded.

3.3.2 Reliability and Validity check of the Self-regulated Learning at Work – Self-reflection and Occupational Self-efficacy questionnaires

For the Self-regulated Learning at Work – Self-reflection questionnaire, the confirmatory factor analysis thresholds for validity were met (RMSEA = 0.000, SRMR = 0.030, CFI = 1.0, TLI = 1.067), as was the reliability threshold (α = .86). Furthermore, the confirmatory factor analysis good fit thresholds for both the pre and post-test Occupational self-efficacy questionnaire were met (RMSEA pre-test = 0.078, RMSEA post-test = 0.000, SRMR pre-test = 0.048, SRMR post-test = 0.034, CFI Pre-test = 1.0, CFI post-test = 1.0, TLI pre-test = 0.969, TLI post-test = 1.007, as were the reliability thresholds for the pre-test and post-test (α pre-test = .88, α post-test = .92). The only exception was the RMSEA for the pre-test, which was slightly above the conventional threshold. However, considering the sample size and the other fit measures being within acceptable ranges, the results can still be deemed satisfactory.

3.3.3 Baseline Self-regulated Learning at Work – Self-reflection Analysis

The Self-regulated Learning at Work – Self-reflection scale was tested for normality and outlier assumptions, both of which were met. Therefore, we conducted an independent sample Student's t-test to determine if there was a statistically significant difference in the results between the progress and social reference frame groups, which had the potential to act as a confounding variable on self-efficacy change. The results from this test found that there was no significant difference (p = .49, d = -0.216) between the progress reference frame group (n = 21, M = 4.06, SD = 1.15) and the social reference frame group (n = 21, M = 4.27, SD = 0.74). Therefore, we could continue with our analysis without controlling further for baseline self-reflection phase processes at work.

3

3.3.4 Absolute Changes in Occupational Self-Efficacy: Comparing Progress vs. Social Reference Frames

Table 4 reports descriptive statistics of the self-efficacy mean and standard deviations for the pre-test, post-test, absolute change and direction of change. The posterior model probabilities (PMP) are presented in Table 5 for the primary research question and Table 6 for the exploratory research question and indicate how much evidence there is in support of each hypothesis. The higher the PMP the more evidence there is that that hypothesis is correct. The results are based on a sensitivity analysis, that is, using decreasing variances in the prior distribution. This is achieved using fraction = 1, 2, 3 in bain (Hoijtink et al., 2019).

Table 4

Self-efficacy means and standard deviations for pre-test, post-test, absolute change, and direction of change.

Reference frame	Mean Pre-test	Mean Post-test	Mean Absolute Change	Mean Direction of Change
Progress	4.476 (1.108)	4.611 (1.154)	0.230 (0.233)	0.135 (0.301)
Social	4.667 (0.730)	4.627 (0.677)	0.325 (0.261)	-0.040 (0.421)

Note. Numbers in brackets are standard deviations from the mean.

Table 5

Bayesian informative hypothesis evaluation ANOVA

Hypotheses	PMP a*	PMP a**	PMP a***
H1 Progress > Social	0.043	0.052	0.057
H2 Progress < Social	0.360	0.436	0.481
H3 Progress = Social	0.598	0.512	0.462

Note. Progress reference frame group n=21, Social reference frame n = 21. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probabilities.

As we can see in Table 5, both H2 and H3 are about equally supported by the PMPs (for each fraction). This means that neither H2 nor H3 is a superior hypothesis. However, the small PMP values of H1 means that this hypothesis can be discarded.

Therefore, these results show that it is possible that the progress and social reference frame equally affect change in occupational self-efficacy (PMP a = 0.598, 0.512, 0.462), which includes the possibility that they both have a similar or no effect. It is also possible that the social reference frame elicits greater change to occupational self-efficacy than the progress reference frame (PMP a = 0.360, 0.436, 0.481). Finally, it is unlikely that the progress reference frame elicits greater change to occupational self-efficacy than the social reference frame (PMP a = 0.043, 0.052, 0.057).

3.3.5 Directional Changes in Occupational Self-Efficacy: Comparing Progress vs. Social Reference Frames

Transitioning to the exploratory aspect of our study, we delved into analysing the direction of change in occupational self-efficacy between the two reference frame groups. The results from the exploratory Bayesian analysis are presented in Table 6 below, which follows a similar structure to Table 5 but focuses on the directional change.

Table 6

Bayesian informative hypothesis evaluation ANOVA

Hypotheses	PMP a*	PMP a**	PMP a***
eH1 Progress > Social	0.474	0.554	0.599
eH2 Progress < Social	0.031	0.036	0.039
eH3 Progress = Social	0.495	0.410	0.362

Note. Progress reference frame group n=21, Social reference frame n = 21. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probabilities.

As we can see in Table 6, both eH1 and eH3 are about equally supported by the PMPs (for each fraction). This means that neither eH1 nor eH3 is a superior hypothesis. However, the small PMP values of eH2 means that this hypothesis can be discarded.

Therefore, these results suggest that there is evidence that the progress reference frame leads to greater positive change in occupational self-efficacy than the social reference frame (eH1) (PMP a = 0.474, 0.554, 0.599), and that both frames might induce equivalent directions of change in self-efficacy (eH3) (PMP a = 0.495, 0.410, 0.362), which includes the possibility that they both have a similar or no effect. Finally, it is unlikely that the social reference frame results in a more positive change than the progress frame (eH2) (PMP a = 0.031, 0.036, 0.039).

3.4 Discussion

Within the context of a VR simulation-based training environment for the chemical industry, this study investigated the impact of LADs designed with progress versus social reference frames on occupational self-efficacy, while controlling for workplace self-reflection. The LADs provided task feedback in an n-Butyllithium manufacturing process. Both the progress and social reference frames offer distinct points of comparison: the former uses prior performance data, and the latter uses peer performance data to support self-reflection phase processes. These reference frames introduce distinct self-evaluation criteria (Zimmerman & Campillo, 2003). Mastery criteria are common in both and serve as an absolute measure for assessing task correctness. Past performance criteria in the progress reference frame and normative criteria in the social reference frame act as relative measures, guiding learners in comparing their current performance to previous performance levels or peer performance levels. These points of comparison serve as criteria for the self-reflection phase, and consequently, as sources of information for shaping self-efficacy beliefs. Previous research indicates that such reference frames are key LAD features for aiding learning analytics data interpretation (Jivet et al., 2020).

The analysis of our primary research question suggests that H2 (i.e., progress < social) and H3 (i.e., progress = social) are about equally supported by the evidence and that H1(i.e., progress > social) can be confidently discarded as a viable. This means it is plausible that either the social reference frame elicits greater change in occupational self-efficacy than the progress reference frame (H2) or the effect of the progress and social reference frames on change to occupational self-efficacy is equal, including the possibility that they have no effect at all (H3). Furthermore, the findings underscore that the progress reference frame does not elicit greater changes in occupational self-efficacy compared to the social reference frame, as hypothesised in H1. One explanation for the potential viability of H2 lies in the types of self-evaluation criteria each frame offers. While the progress reference frame provides mastery and past performance criteria, the social reference frame provides mastery and normative criteria. The latter introduces social modelling information as an additional information source for self-efficacy beliefs, compensating for the absence of past performance criteria (Bandura, 1997). This wider range of information which influences self-efficacy can lead to greater self-efficacy change.

Another possibility is that the progress reference frame's emphasis on previous performance criteria may be redundant if learners can recall their past performances. Such recall can serve as a substitute for explicit previous performance criteria in shaping self-efficacy beliefs. For example, a learner reviewing their LAD on step 3 could recall scores from steps 0, 1, and 2, and in turn, make judgements about their progress without needing it displayed on the LAD. This redundancy could vary if performance steps were spaced days or weeks apart, making it harder to remember past outcomes. This explanation is consistent with the finding that the progress reference frame fails to affect self-efficacy change more than the social reference frame (H1). Previous studies support this showing that LADs incorporating both mastery experiences and social modelling information have a differential impact on learning (Brusilovsky et al., 2016).

Results for our primary research question also support H3, suggesting that both progress and social reference frames equally affect self-efficacy change. This points to the performance outcome component in both LADs as the key influencer, rather than the distinct points of comparison. For example, a learner may evaluate their own performance based on how many stars they received out of five and ignore the point of comparison. However, this would be surprising because we know social comparison information can play a significant role in self-reflection phase processes (Wilson & Shanahan, 2020). For example, Davis et al. (2017) showed that LADs with social comparison features can drive learning behaviour change, likely through self-reflection processes.

Another reason for the evidential support of H3 could be that the LADs' feedback fails to impact self-efficacy due to the overshadowing effect of the simulation-based training. Supporting this, Hough et al. (2019) found significant self-efficacy changes in a simulation training without LADs among physiotherapy students. While LADs may stimulate self-reflection, they may not be impactful enough to alter self-efficacy in some contexts, including ours. However, it is essential to consider the evolving nature and potential of LADs. Research by Susnjak et al. (2022) propose advancing LADs from descriptive to predictive and prescriptive analytics to enhance self-reflection and guide learners more effectively.

The results of the exploratory research question suggest that eH1 (i.e., progress > social) and eH3 (i.e., progress = social) are about equally supported by the evidence and that eH2 (i.e., progress < social) can be confidently discarded as a viable hypothesis. This means it is plausible that either the progress reference frame elicits a more positive direction of change in occupational self-efficacy than the social reference frame (eH1) or the effect of the progress and social reference frames on direction of change to occupational self-efficacy is equal, including the possibility that they have no effect at all (eH3). One explanation for the potential viability of eH1 is that the access to mastery experiences information and in turn past performance criteria leads to more positive self-evaluations and therefore, their self-efficacy beliefs are more positively or less negatively affected. As described earlier, Bandura's theory of self-efficacy states that mastery experiences are one of the most potent sources of self-efficacy (Bandura, 1997). The past performance criteria, acting as tangible evidence of mastery, can bolster individuals' self-evaluations. Seeing clear evidence of their progress or achievements could enhance their beliefs in their capabilities, thereby positively influencing their self-efficacy.

The results supporting the rejection of eH2, that it is unlikely the social reference frame results in a more positive change to occupational self-efficacy compared to the progress frame, can potentially be explained through the effect of favourable and unfavourable evaluations (Gerber et al., 2018). When unfavourable evaluations occur via the social reference frame, the negative impact on learners could be more pronounced as it may highlight a learners own inadequacies compared to the superiority of others. This comparison with peers could potentially exacerbate the negative effect on self-efficacy beliefs.

Conversely, the progress reference frame draws from the learner's own past performance for evaluations (Zell & Alicke, 2010). While unfavourable evaluations in this reference frame might still be disheartening, they may be less influential on self-efficacy beliefs as they can be accepted as valuable setbacks necessary in a workplace learner's professional development.

Alternatively, favourable evaluations within a social reference frame, though potentially boosting self-efficacy, might offer a less impactful boost as they hinge on the variable performance of others, which is outside the learner's control (Gerber, 2020). This could potentially make the social reference frame less powerful in eliciting positive change in self-efficacy. In contrast, favourable evaluations within a progress reference frame signify personal growth and achievement, detached from others' performance. This scenario potentially fosters a more substantial boost in self-efficacy as it underscores the individual's capabilities alongside the efficacy of their effort and task strategies, all of which are within their control (Zimmerman, 2012). Hence, the progress reference frame possibly provides a stronger basis for positive changes in occupational self-efficacy.

3.4.1 Limitations

Immersive learning environments using VR technology are relatively new to the chemical process industry. Therefore, one limitation of the study may be that it did not control for the novelty effect (Clark, 1983; Makransky & Petersen, 2021), which relates to the positive effect new technology can have on learner task performance, which decreases once the novelty of the technology wanes. Longitudinal research design can control for such an effect; however, this was infeasible due to resource constraints. Nevertheless, due to the context of the research and its focus on self-efficacy change and not task performance outcomes, it is unclear what role the novelty effect could play in influencing results, if any.

Further, this study did not account for technology acceptance factors that could influence self-efficacy (Udeozor et al., 2021). These unmeasured variables introduce potential confounders, and their absence could skew the participants' interpretation of feedback, given their varying comfort and familiarity with VR technology. Future research should include measures of technology acceptance for a richer view of its impact on learning outcomes.

Methodologically, the absence of a manipulation check is another limitation. While learners were required to interact with their LAD to advance to the subsequent task step, this interaction did not confirm the extent of their engagement with the LAD for self-reflection phase processes. Future studies could employ eye-tracking tools to assess this (Clay et al., 2019).

The study did not account for potential variations in attention and motivation during LAD interaction, which could influence occupational self-efficacy changes. Incorporating measures based on Bandura's reciprocal determinism theory (Bandura, 1997) in future investigations could provide deeper insights into how these factors affect the use of LAD feedback and the subsequent change in occupational self-efficacy.

Another limitation to consider is the breadth of analysis regarding the directions of comparison that learners engaged in throughout the task. The effects resulting from upward and downward comparisons (Guyer & Vaughan-Johnston, 2018; Suls et al., 2002), corresponding to favourable and unfavourable evaluations respectively, were not fully explored. For instance, learners likely encountered a mix of upward and downward comparisons during task performance. Knowing the differential effects of these directions of comparison in combination with the types of reference frames would be of additional value.

3.4.2 Implications and Future Research

This study advances learning analytics design theory, particularly research examining reference frames. It highlights how Temporal and Social Comparison Theory can help us understand how learners engage in self-reflection phase processes when confronted with certain learning analytics designs, such as LADs with different reference frames. Future research could focus on how other types of reference frame designs affect the self-reflection phase, such as reference frames which incorporate self-set and assigned goals (Hollenbeck & Brief, 1987; Seo et al., 2018). Another fruitful line of future research could be to focus on exploring the potential impact of different reference frames on performance outcomes, providing valuable insights into the effectiveness of different instructional approaches and the role of reference frames in optimising performance.

In addition, exploring the impact of years of work experience on self-efficacy change within the context of different reference frames could be insightful. Understanding the interplay between experience levels and different reference frames may shed light on the dynamics of self-efficacy beliefs in relation to expertise development, crucial in industries like the chemical sector (Garcia Fracaro et al., 2021). This research could inform the design of targeted interventions and training programs to target self-efficacy and promote optimal performance among individuals at different stages of experience.

Additionally, this study contributes to learning analytics design theory by introducing the concept of learning analytics reference frame components. The study argues that reference frames consist of three components which play a role in helping learners make sense of their feedback and include the performance outcome, point of comparison, and score delta. The point of comparison component was the primary focus of this study; however, future research is needed on how score delta, which aligns theoretically with upward, lateral, and downward directions of temporal and social comparison, influence self-efficacy change and direction of change (Suls et al., 2002).

The exploratory section of this study unveiled the potential differential impacts of reference frames on self-efficacy change direction. However, a deeper understanding requires further examination of score delta, aligning with various directions of temporal and social comparison. This study initiated this exploration, yet a more detailed examination of the interplay between these constructs is needed to clarify the findings and address the earlier mentioned limitations.

A final theoretical implication of this study is that it provides early evidence that immersive learning environments for the workplace which incorporate LADs with reference frames can affect learner occupational self-efficacy. Further research into how to best design LADs for immersive learning environments to optimise self-efficacy change is needed.

From a practical perspective, the implications of this research are that LAD designers should carefully consider which reference frames they use when designing LADs because this decision likely has consequences on the formation of learner self-efficacy (Matcha et al., 2020).

CHAPTER 4

4

Impact of Reference Frames on Learning Analytics Dashboard Use

This chapter is based on:
Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2023). Impact of Reference Frames on Learning Analytics Dashboard Use. *Manuscript submitted for publication.* Department of Education, Utrecht University.

Acknowledgement of author contributions
All authors designed the study. Timothy Gallagher contributed to the study's conceptualization, methodology, formal analysis, investigation, data curation, drafting the original manuscript, manuscript review and editing, visualization, and project administration. Bert Slof, Marieke van der Schaaf, and Liesbeth Kester contributed to the conceptualisation, methodology, manuscript review and editing, supervision, and funding acquisition. Ryo Toyoda and Yusra Tehreem contributed to the methodology, investigation, and resources. Sofia Garcia Fracaro contributed to the methodology, investigation, resources, and project administration.

Acknowledgement:
The authors would like to thank Prof. Herbert Hoijtink for sharing his statistical expertise.

This project has received funding from the European Union's EU Framework Programme for Research and Innovation Horizon 2020 under Grant Agreement No 812716. This publication reflects only the authors' view exempting the community from any liability. Project website: https://charming-etn.eu/

Abstract

This study uses log-file data to investigate how chemical plant employees interact and engage with two distinct learning analytics dashboard designs, framed by achievement goal orientation theory and temporal and social comparison theories. The learning analytics dashboards are implemented in a virtual reality simulation-based training environment. The designs differ by reference frame: the progress reference frame offers historical performance data as a point of comparison, and the social reference frame offers average peer group performance data as a point of comparison. We analysed participants' time spent reviewing the dashboards, time spent reviewing detailed task feedback, and frequency of engagement with the LADs, measured by the number of times specific features related to detailed task feedback and assessment formula were selected. Results suggest that participants who receive a progress reference frame are likely to spend less time reviewing their main LAD which contains a reference frame than those who receive a social reference frame. However, those who receive a progress reference frame are likely to spend more time reviewing detailed task feedback LADs. This may indicate that progress reference frames can encourage mastery goal orientation behaviours, while social reference frames may promote performance goal orientation behaviours.

4.1 Introduction

Virtual reality (VR) training environments are becoming popular tools for training employees because they offer advantages over other forms of training (Makransky & Petersen, 2021). For example, these environments can take advantage of log-file data, which can be used with learning analytics tools such as learning analytics dashboards (LAD) (Ruiz-Calleja et al., 2017). While learning analytics refer to the collection and analysis of data to optimize learning (Siemens & Baker, 2012), LADs aggregate data collected during the learning analytics process and display it within one or multiple visualizations to help stakeholders make sense of the learning analytics data (Matcha et al., 2019). LADs are often designed to provide feedback on task performance to learner stakeholders (Schwendimann et al., 2017). LAD designers can help learners make sense of their feedback by including reference frames, which contextualize a learner's performance against a particular point of comparison (Wise & Vytasek, 2017). Two types of reference frames are the progress and social reference frame (Jivet et al., 2017). The progress reference frame uses historical performance data as a point of comparison, while the social reference frame uses aggregated peer performance data as a point of comparison.

Despite the growing use of LADs in workplace training environments (Poquet et al., 2022; Ruiz-Calleja et al., 2017), there remains a significant gap in understanding how varying reference frame designs impact learner interaction with LADs that provide feedback, particularly within the chemical industry. This research is of high relevance, as operator training is crucial due to the potentially hazardous nature of procedures and the severe consequences of operator errors on process operation and safety (Garcia Fracaro et al., 2021). Addressing challenges in operator training, such as high costs, safety limitations, time constraints, and employee engagement, is essential for the industry, and immersive technologies such as VR training environments with LADs can offer innovative solutions to overcome these obstacles (Garcia Fracaro et al., 2021).

In this study, we explore how workplace learners in the chemical industry interact with feedback presented by two LADs, one designed with a progress reference frame and informed by temporal comparison theory (Albert, 1977; Wilson & Shanahan, 2020), and one designed with a social reference frame and informed by social comparison theory (Festinger, 1954; Gerber, 2020). Underpinned by achievement goal orientation theory, which defines two primary goal orientations: mastery and performance, our hypotheses posit that the reference frames will differentially impact learners' engagement with LADs (Pintrich, 2000a). Because these orientations influence learner behaviour, motivation, and cognition (Pintrich, 2000a), it is important to examine the influence of progress and social reference frames on learners' interactions with LADs and their subsequent learning behaviours. By understanding the impact of different reference frames on learning, this research aims to provide insights for designing more effective LADs that optimize the learning experience in workplace training environments.

4.1.1 Literature Review

The rising adoption of VR simulation-based training environments is partly due to their immersive nature and the safe, controlled setting they offer learners, particularly in high-risk sectors such as the chemical industry (Srinivasan et al., 2022). A recent study illustrated that VR-based safety training in the chemical sector matched traditional methods in learning outcomes, while enhancing trainees' perceptions of learning, suggesting VR's potential for better knowledge assessment and retention (Poyade et al., 2021). Additionally, research has demonstrated the potential of VR to enhance training methodologies in the chemical industry by safely simulating hazardous emergency scenarios, providing real-time feedback, and facilitating in-training evaluation (Garcia Fracaro, Bernaerts, et al., 2022).

Concurrently, empirical investigations into LADs have highlighted their potential to foster environments to support learning experiences and outcomes (Verbert et al., 2020; Wise & Vytasek, 2017). For example, facilitating social comparison in MOOCs through LADs can significantly increase course completion rates (Davis et al., 2017). Moreover, bachelor students interacting with an LAD design informed by principles of goal orientation and social comparison showed greater motivation compared to those without access, outperforming peers as the course progressed and ultimately achieving higher final grades (Valle et al., 2023). Such designs aim to positively influence motivation and enhance academic performance. Supporting this, research found that when compared to a control group, implementing an LAD, similarly informed by goal orientation and social comparison, promoted learner extrinsic motivation, which subsequently led to improved academic performance as the course progressed (Fleur et al., 2023).

Initial investigations have begun to explore the different impacts of reference frames in LADs, specifically focusing on progress and social reference frames. Early evidence indicates that LADs designed with either a progress or a social reference frame, particularly when aligned with learners' own preference for reference frames, can positively influence learning persistence and academic performance (Janson et al., 2022). Complementing this, research in the context of a higher-education database management course, found that LADs incorporating both the progress and social reference frame enhanced learning, as well as user attitude and engagement, when compared to LADs with only a progress reference frame (Brusilovsky et al., 2016). The study also found that the different LAD designs differentially influenced student usage patterns. However, the research into LAD design and reference frames within VR training environments, especially in the context of the chemical industry, remains underexplored.

4.1.2 Employing Temporal and Social Comparison for Enhanced Learning Analytics Dashboard Design

When designing workplace LADs for feedback, instructional designers must consider how reference frames, such as progress and social reference frames, can help learners make sense of their feedback and influence their interaction with the available LADs .

When presented with a progress reference frame, learners are stimulated to engage in temporal comparisons, which take place when one compares their own performance at different points in time (Jivet et al., 2017). Temporal comparisons are rooted in the idea that people are motivated to evaluate their progress towards goals and adjust their efforts accordingly (Wilson & Shanahan, 2020; Zell & Alicke, 2010). In the context of learning, this can be based on their ability to perform learning tasks at different points in time. By highlighting progress in task performance over time, temporal comparisons can help learners determine if they have been improving. Therefore, temporal comparisons may influence learner interaction with LADs. For example, if given the opportunity, learners may wish to review detailed task feedback to find out what they can to do to improve (Wilson & Shanahan, 2020), which is representative of a mastery goal orientation (Pintrich et al., 2003). The advantage of temporal comparisons in the context of LADs is that they promote self-improvement and deeper understanding of one's performance. However, the potential drawback is that, in the absence of external benchmarks, learners may have difficulty assessing the adequacy of their progress (Wilson & Shanahan, 2020). This desire for an external benchmark may in part explain why research has shown that learners in certain contexts consider the ability to access their own grades as the most relevant feature of an LAD (Jivet et al., 2020), because grades are typically considered to be an objective and quantifiable measure of their performance and progress in a course.

When presented with a social reference frame, learners are stimulated to engage in social comparison, which takes place when one compares their own performance with that of their peers and do so to gauge how effective they are at tasks (Cleary, 2009). Social comparison theory is based on the premise that individuals are driven to evaluate their abilities by comparing themselves to others, particularly when objective standards are absent (Festinger, 1954; Gerber, 2020). Therefore, social comparison may influence learner interactions with LADs because it may impact their motivation (Corrin & de Barba, 2015a; Jivet et al., 2017) and encourage them to focus on performing better than their peers instead of self-improvement, which is representative of a performance goal orientation (Pintrich et al., 2003). The potential advantages of social comparisons in the context of LADs are that they can enhance motivation and foster a sense of competition. However, the potential drawback is that they may lead to negative emotions, such as frustration (Fukubayashi & Fuji, 2021) or anxiety (Erdoğan et al., 2011), and promote surface-level learning strategies rather than deeper understanding (Pintrich et al., 2003). To better understand how learners' goals influence their LAD use, achievement goal theory could provide interesting insights.

4.1.3 Achievement Goal Orientation

Achievement goal orientation theory provides a framework for understanding how learners' goals influence their motivation, behaviour, and cognition (Pintrich, 2000a). The theory distinguishes between two primary goal orientations: mastery and performance (Pintrich, 2000a). Each of these can further be categorized along approach or avoidance dimensions (Elliot & McGregor, 2001), although this aspect is not the focus of the present study.

Mastery goals focus on skill development, self-improvement, and a deep understanding of subject matter or tasks (Pintrich et al., 2003). In contrast, performance goals emphasize outperforming others (Pintrich et al., 2003). Notably, the understanding of how reference frames can influence learners' interactions with LADs is underpinned by these goal orientations. Specifically, the theory highlights how a progress reference frame may align with mastery goals, due to the induced temporal comparisons, while a social reference frame may align with performance goals, due to its induced social comparisons, thereby shaping learners' engagement and experiences with LADs.

Specifically, exposure to a progress reference frame within an LAD is likely to foster a mastery goal orientation, encouraging learners to concentrate on self-improvement and understanding their progress. Conversely, a social reference frame is likely to induce a performance goal orientation, leading learners to compare themselves with others and aim for superiority (Pintrich et al., 2003). Consequently, the type of reference frame presented in LADs may dictate the orientation of goals learners adopt, thereby affecting their engagement and subsequent learning behaviours.

4.1.4 Context of the Study and Study Overview

This study explores the influence of two LAD designs implemented into a VR simulation-based training environment for employees in the chemical industry. VR simulation-based training is suitable for this industry due to its potential for providing realistic and safe training in complex and potentially hazardous environments (Garcia Fracaro et al., 2021; Garcia Fracaro, Tehreem, et al., 2022).

The focus of this research is the comparison of two distinct LAD designs based on different reference frames. The first, termed the Progress LAD, incorporates a progress reference frame, while the second, the Social LAD, employs a social reference frame. Despite this difference, both designs share common features such as 'Step Overview' and 'How is this calculated?' buttons. The 'Step Overview' button leads users to a secondary dashboard offering detailed task feedback, whereas the 'How is this calculated?' button provides insights into the computation of step scores.

To better understand the impact of progress and social reference frames on learner interaction with LADs, we designed a two-group experimental study. Participants took part in

VR simulation-based training and received feedback via either the Progress or Social LAD. We operationalized LAD interaction by analysing log-file data, which included metrics such as the amount of time participants spent on each LAD, the time dedicated to reviewing detailed task feedback through the 'Step Overview' button, and the frequency of their interactions with the LAD, as indicated by the number of times they selected the 'Step Overview' and 'How is this calculated?' buttons.

4.1.5 Research Questions

This study investigates the impact of reference frame designs in LADs on user interaction. To further explore this subject, we have formulated three research questions, each accompanied by three competing hypotheses. These are designed to address the overarching research question: How do reference frames influence LAD interaction? The specific research questions are as follows:

RQ1: Are there between group differences in total time spent reviewing Learning Analytics Dashboards (LADs) with a reference frame?

RQ2: Are there between group differences in total time spent reviewing detailed task feedback?

RQ3: Are there between group differences in engagement with LADs?

4.1.6 Bayesian Informative Hypotheses

To answer these research questions, we use Bayesian Informative Hypothesis Evaluation, enabling a direct comparison of competing hypotheses and quantification of their relative support through data analysis (Hoijtink et al., 2019; Van Lissa et al., 2021).

An informative hypothesis is defined as a "theoretically derived statement about directional differences and equality constraints between model parameters of interest" (Van Lissa et al, 2020, p. 2). Hypotheses are formulated using equality (=) and inequality (<, >) terms, with parameter groups constrained using parentheses.

4

4.1.7 Hypothesis formulation

The hypotheses below are grounded in the theoretical frameworks and empirical literature explored in earlier sections. Specifically, as learners are exposed to either the progress or social reference frame, and in turn, temporal, and social comparison respectively, they are anticipated to exhibit interaction behaviours representative of either mastery or performance goal orientation. Consequently, this implies that the level of engagement and interaction with LADs may vary based on the reference frames presented. However, while differences in LAD interaction are expected, the current theoretical and empirical evidence is insufficient to precisely detail how these differences will appear. Given this uncertainty, the proposed hypotheses for each research question aim to account for expected variations in interaction due to the progress and social reference frames.

4.1.8 Research Question and Informative Hypotheses

In this section, we present three research questions accompanied by three competing hypotheses each, aiming to explore the influence of reference frames on LAD interaction.

RQ1: Are there between group differences in total time spent reviewing LADs with a reference frame?

H1.1 The Progress LAD group mean time spent reviewing the LAD with a reference frame will be greater than the Social LAD group.

H1.2 The Progress LAD group mean time spent reviewing the LAD with a reference frame will be less than the Social LAD group.

H1.3 The Progress LAD group mean time spent reviewing the LAD with a reference frame will be equal to the Social LAD group.

RQ2: Are there between group differences in total time spent reviewing detailed task feedback?

H2.1: The Progress LAD group mean time spent reviewing the detailed task feedback dashboard will be greater than the Social LAD group.

H2.2: The Progress LAD group mean time spent reviewing the detailed task feedback dashboard will be less than the Social LAD group.

H2.3: The Progress LAD group mean time spent reviewing the detailed task feedback dashboard will be equal to the Social LAD group.

RQ3: Are there between group differences in engagement with LADs?

H3.1: The Progress LAD group mean LAD engagement frequency will be greater than the Social LAD group.

H3.2: The Progress LAD group mean LAD engagement frequency will be less than the Social LAD group.

H3.3: The Progress LAD group mean LAD engagement frequency will be equal to the Social LAD group.

4.2 Materials and Methods

4.2.1 Participants

Table 1 shows study participants (N=42) were experienced chemical plant operators aged between 18 and 55 years, consisting of 40 males and 2 females. There were no participants who said they would prefer not to say. Convenience sampling was used to select participants. Informed consent was obtained voluntarily from each participant after they reviewed an explanatory letter detailing the research objectives. The workers' council, representing the employees, granted approval for the study and the use of the instruments. Participation in the training program was elective, and subjects could withdraw from the study at any point without repercussions. The working language of the participants was German, and a German-language prototype was employed accordingly.

Table 1

Participants' years of experience

Years of Experience	Total	Progress Group	Social Group
1 - 10 years	10	5	5
11 - 20 years	20	12	8
More than 20 years	12	4	8
Total	42	21	21

4.2.2 Experimental Design

The study was a between group design in which the effect of LADs designed with a progress reference frame were tested against the effect of LADs designed with a social reference frame on three dependent variables associated with LAD interaction: time spent reviewing LADs with a reference frame, time spent reviewing detailed task feedback, engagement frequency with LAD. Participants were randomly assigned to the progress LAD group (n=21) and the social LAD group (n=21).

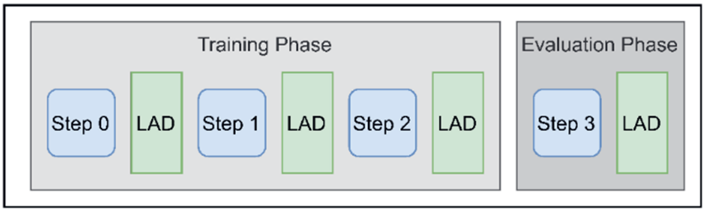
The sample size of 42 participants, equally divided between the two groups, was deemed adequate for this study, as the utilization of Bayesian informative hypothesis evaluation allowed for robust competing hypothesis testing even with a relatively smaller sample size. However, additional research is required with a larger sample size for these results to be generalizable.

4.2.3 Description of 'Operate your own reactor' training environment

The Operate your own reactor (Tehreem et al., 2022) VR training simulator, designed for Oculus Quest, aims to train employees in the n-Butyllithium manufacturing process. Trainees complete a series of procedural tasks, involving the operation of commercial chemical reactor equipment. As described in Figure 1, the procedure comprises four steps, each containing specific tasks (i.e., Search for the ADKE1 panel in the control monitor.), sub-tasks (i.e., Complete the task on flushing ADKE1.) and main tasks (i.e., Complete all tasks on the Inertization process.), with the first three steps considered as a training phase, and the final step as an evaluation phase.

Figure 1

Procedure Steps



Note. Design of training and evaluation phase and the timing of the LAD presentation.

Trainees navigate and perform tasks within the VR chemical plant, which contains the necessary commercial equipment for n-Butyllithium preparation. They interact with digital instruments (e.g., valves, hoses, batch reactors) and digital monitors that control these instruments and provide procedural information via the procedure board. As described in Figure 2, the procedure board offers sufficient guidance for skilled operators, while less skilled operators may misinterpret these instructions and subsequently make mistakes and require corrective feedback. In the first three steps, corrective feedback is automatically given. During the final step, trainees must request a "hint" to receive corrective feedback. The LADs are displayed on the digital monitors as are details of the button mapping for the VR controllers.

Figure 2

Procedure Board

Note. The procedure board guides the participants through the training and enables them to control the chemical reactor.

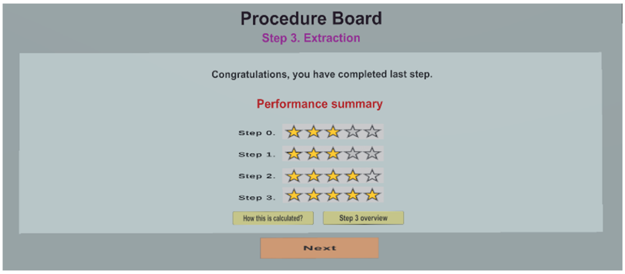
4.2.4 Learning Analytics System and Dashboard Features

The learning analytics system collects and analyses log-file data associated with performance criteria, such as correct and incorrect actions, hints requested, and time taken for each step. Learners receive a score out of five, represented by stars. Figure 3 and Figure 4 show screenshots of the Progress LAD and Social LAD, respectively.

Both the Progress LAD and Social LAD comprise of six primary components: (1) the name of the recently completed step, (2) a congratulatory message to acknowledge the participant's achievement, (3) the participant's performance summary accompanied by a reference frame (progress or social, contingent upon the group), (4) a button labelled 'How is this calculated?', (5) a button labelled 'Step overview', and (6) a button labelled 'Next'.

Figure 3

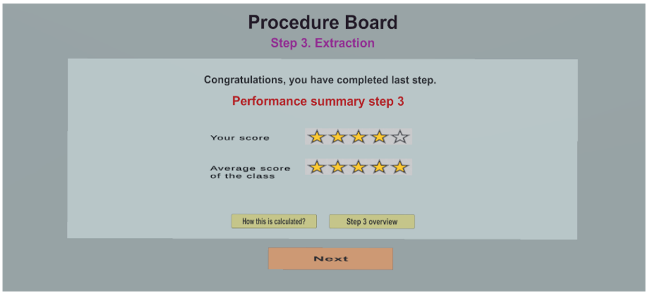
Progress LAD After Step 3



Note. Progress LAD after step 3 (English version shown). Learners can compare how they performed on step 3 with previous steps.

Figure 4

Social LAD After Step 3

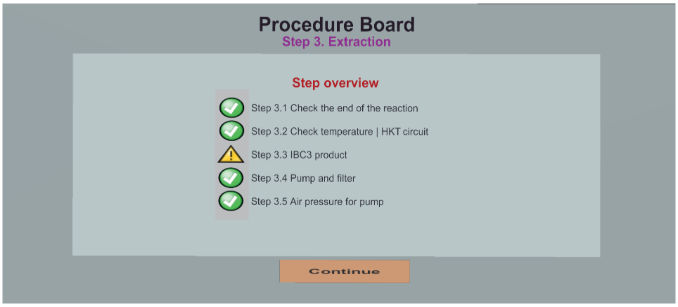


Note. Social LAD after step 3 (English version shown). Learners can compare how they performed on step 3 with the average score of their peers on step 3.

The performance summary communicates the learner's performance level using a star system that is intended to be quick and easy to understand. It ranges from one to five stars and allocates a greater number of stars to denote better performance. Participants may choose to select either the 'Step overview' or 'How is this calculated?' (Fig. 3) buttons to gain additional insights into their performance. Once selected, additional context related to the feedback received is provided. Figure 5 illustrates an instance of the 'Step overview' button being selected. When selected, it displays an additional page of the LAD that outlines the participant's performance outcome on the task at a sub-task level. Figure 6 portrays the 'How is this calculated?' indicator (blue text). When selected, the 'How is this calculated?' button, presents an indicator on the progress or social LAD depending on condition describing the basic formula for calculating the star system (e.g., 92-100% awards five stars)

Figure 5

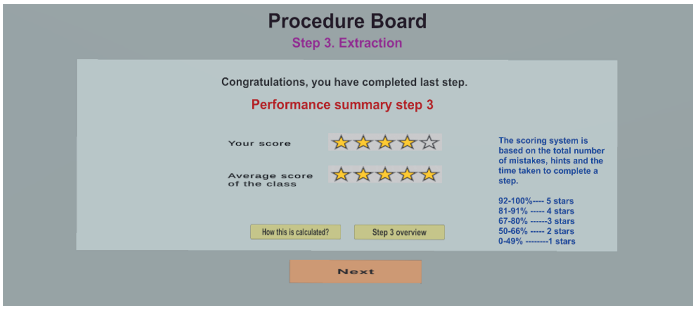
Detailed Task Feedback Dashboard



Note. Detailed task feedback dashboard indicating performance on sub-steps (English version shown).

Figure 6

Social LAD with Calculation Indicator



Note. Social LAD when the ‘How is this calculated?’ button is selected (English version shown).

As shown in Figure 5, learners automatically receive their LADs upon the completion of each step of the task. The detailed task feedback dashboard is presented only when the ‘Step Overview’ button is selected. The assessment formula indicator is only presented when the ‘How is this calculated?’ button is selected. Figure 3 is a screenshot of the Progress LAD after completing step 3. It can be described as incorporating a progress reference frame because the learner's most recent performance outcome (step 3) is compared with their previous performances (step 0 – step 2). Figure 4 is a screenshot of the Social LAD after completing step 3. It can be said to incorporate a social reference frame because the learner's most recent performance outcome (step 3) is compared with the average of their peers.

4.2.5 Procedure

Upon arrival, participants were welcomed and introduced to the 'Operate Your Own Reactor' prototype training session, which took place during their working hours. Researchers provided each participant with an information sheet and an informed consent form to sign if they agreed to participate. Before entering the VR simulation-based training environment, all participants completed a demographic questionnaire and other instruments which were used by collaborating researchers. Next, they were shown how to navigate and interact with the virtual environment using the required VR hardware, ensuring they were familiar with the controls and equipment. Once comfortable with the controls and environment, participants were instructed to begin the training. They were informed that they could take breaks after completing a step if needed. Participants completed the training with either the Progress LAD or Social LAD. They were not aware of the two different LAD designs.

Throughout the training, log-files related to LAD interactions were automatically generated and stored on the VR device which were then exported for data analysis. Upon completion of the training, which typically lasted around an hour, participants completed additional instruments used by collaborating researchers.

4.2.6 Data analysis and statistical models

We conducted three separate Bayesian ANOVA's in JASP (version 0.18.1), with LAD groups set as fixed factors. The aim was to compare learner LAD interaction scores between the Progress and the Social LAD conditions.

The variables associated with LAD interaction included: (1) time spent reviewing LAD with a reference frame, (2) time spent reviewing the detailed task feedback dashboard, and (3) engagement frequency with LAD, as measured by the frequency with which the 'Step Overview' and 'How is this calculated?' buttons were selected. Hypotheses for evaluation for each RQ are:

RQ1: H1.1: Progress > Social, H1.2: Progress < Social, H1.3: Progress = Social.

RQ2: H2.1: Progress > Social, H2.2: Progress < Social, H2.3: Progress = Social.

RQ3: H3.1: Progress > Social, H3.2: Progress < Social, H3.3: Progress = Social.

A sensitivity analysis, as outlined by (Hoijtink et al., 2019), was conducted using fractions of 1, 2, and 3, with each result reported as Posterior Model Probabilities (PMPs). This analysis highlights the impacts of altering the prior distribution on the computation of the PMPs by offering a collective interpretation of these three sets of results. Following this, the Bayesian error associated with preferring the best hypothesis in terms of PMPs will be reported as the sum of the PMPs of the other hypotheses.

4.3 Results

In this study, we aimed to examine evidence in support of three competing hypotheses for each research question. First, we present the descriptive statistics in Table 2. Next, Tables 3, 4, and 5 report the PMPs for RQ1, RQ2, and RQ3, respectively, indicating the level of support for each hypothesis. A higher PMP suggests stronger evidence for the corresponding hypothesis.

Tabel 2

Descriptive statistics for interaction with LAD

	Seconds reviewing LADs with reference frame		Seconds reviewing detailed task feedback		LAD engagement frequency	
	Progress (n=21)	Social (n=21)	Progress (n=21)	Social (n=21)	Progress (n=21)	Social (n=21)
Mean	27	33.6	2.1	0.3	0.9	0.4
StdD	9.4	16.3	3.8	1.2	1	0.7
Min	14	9	0	0	0	0
Max	54	64	15	5	3	2

Table 3

Bain ANOVA RQ1 Time spent reviewing LAD with reference frame

	PMP a*	PMP a**	PMP a***
H1: Progress > Social	0.028	0.033	0.035
H2: Progress < Social	0.503	0.583	0.627
H3: Progress = Social	0.469	0.384	0.338

Note. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probabilities.

Table 4

Bain ANOVA RQ2 Time spent reviewing detailed task feedback

	PMP a*	PMP a**	PMP a***
H1: Progress > Social	0.698	0.763	0.795
H2: Progress < Social	0.015	0.016	0.017
H3: Progress = Social	0.287	0.221	0.188

Note. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probabilities.

Table 5

Bain ANOVA RQ3 Learning analytics dashboard engagement

	PMP a*	PMP a**	PMP a***
H1: Progress > Social	0.480	0.560	0.605
H2: Progress < Social	0.030	0.035	0.038
H3: Progress = Social	0.490	0.404	0.357

Note. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probabilities.

Table 4 shows that the hypothesis stating less time is spent reviewing the Progress LAD (H1.2) has the most support, while the hypothesis suggesting more time is spent reviewing the Progress LAD (H1.1) is substantially unsupported. This seems supportive for the claim that learners with a Progress LAD spend less time reviewing their LAD than those with a Social LAD. However, due to the error probability (0.497, 0.417, 0.373), we cannot entirely rule out the hypothesis that the two LAD groups spend an equal amount of time reviewing their LADs with a reference frame (H1.3).

Table 4 reveals that the hypothesis stating the Progress LAD group spends more time reviewing the detailed task feedback dashboard than the Social LAD group (H2.1) has the most support. Conversely, the hypothesis stating the Progress LAD group spends less time reviewing the detailed task feedback dashboard than the Social LAD group (H2.2) is substantially unsupported. This suggests that the Progress LAD leads to more time spent reviewing detailed task feedback. However, due to the error probability for H2.1 (0.302, 0.237, 0.205), we cannot entirely rule out the hypothesis that the two groups spent an equal amount of time reviewing the detailed task feedback dashboard (H2.3).

Table 5 indicates that we do not have evidence to rule out either the hypothesis claiming the Progress LAD group engages more with the LADs than the Social LAD group (H3.1) or the hypothesis stating the two groups engage equally with the LADs (H3.3). On the other hand, the hypothesis stating the Progress LAD group engages less with the LADs than the Social LAD group (H3.2) is substantially unsupported. This suggests that the Progress LAD may lead to greater or equal engagement with the LADs compared to the Social LAD.

4.4 Discussion

The findings from RQ1 demonstrate that the total time learners spent reviewing their LADs with a reference frame varies depending on the reference frame provided. Learners receiving LADs with a social reference frame were found to have spent more time reviewing their LADs compared to those receiving a progress reference frame. This could be attributed to the fact that the Social LAD provides a more surface-level indication of performance, potentially fostering a performance orientation as learners may have spent additional time trying to interpret and contextualize their standing relative to their peers. Additionally, the increased time spent on the Social LAD may be attributed to the negative emotions associated with social comparison information, such as frustration and anxiety (Wortha et al., 2019), which were discussed as potential drawbacks in the theoretical framework. Conversely, learners in the Progress LAD group may have spent less time on their LADs as they are more inclined to seek feedback or move on to the next step of the task, thus exhibiting a stronger mastery orientation. The temporal comparison allows them to focus on their own progress and improvement, leading them to select either the "Step Overview" button, "How is this calculated?" button, or move to the next step more quickly. This suggests that the reference frame of the LADs can meaningfully influence learners' interaction and goal orientation.

While the results for RQ1 suggest differences in learners' time spent reviewing their LADs with a reference frame, it is important to note that due to the error probability, we cannot entirely rule out the hypothesis that the two LAD groups spent an equal amount of time reviewing their LADs with a reference frame. This caveat highlights the need for further research to examine the potential impact of individual differences or contextual factors on learners' interaction with different reference frames.

The results from RQ2 and RQ3 offer insights into learners' subsequent actions. RQ2 findings suggest that the Progress LAD group is more inclined to spend additional time reviewing the detailed task feedback dashboard. This result aligns with the theoretical framework presented in the introduction, discussing the relationship between the progress reference frame and mastery goal orientation. Specifically, learners engaging with the progress reference frame spent more time reviewing the detailed task feedback dashboard, reflecting their desire for a deeper understanding of their task performance (Pintrich et al., 2003). This evidence underscores the relevance of achievement goal orientation theory in interpreting the impact of reference frames on learner interactions with LADs. It suggests that temporal comparisons may stimulate learners to consider self-improvement, prompting them to seek detailed task feedback. This is consistent with a mastery goal orientation, as it concerns learners desiring a deeper understanding of their performance (Pintrich et al., 2003). In contrast, the detailed task feedback LAD was reviewed for a shorter amount of time by the Social LAD group. While social comparison might align with performance goal orientation, potentially leading to superficial learning, this study did not investigate that specific relationship.

It is important to note that the feedback provided by the detailed task feedback dashboard is a form of corrective feedback (i.e., highlighting specific incorrectly performed sub-tasks). As a result, learners who receive a perfect score of five stars in both the progress and social reference frame conditions may feel that they have no reason to review their detailed task feedback, as they already know they performed well. Despite this, differences in time spent reviewing detailed task feedback between the two reference frame conditions were still present. Our study did not account for this effect, which should be a focus for future research. Subsequent studies could either control for this effect or investigate how to make the detailed task feedback screen valuable to learners, irrespective of their scores.

The results for RQ3 suggest that the Progress LAD group may be more open to mastery goal orientations, while the Social LAD group may be more open to performance goal orientations. However, the potential inclinations of each LAD group towards each respective goal orientation require further investigation. The results do allow us to more confidently rule out higher engagement with LADs in the Social LAD group, but additional research is necessary to definitively determine whether the Progress LAD group engages more or the same as the Social LAD group.

While the results of this study provide evidence of variances in learners' interactions with LADs based on the reference frames provided for each research question, it is important to consider the possibility that there may be no differential effect between the two reference frames considering the error probability towards the equal effect. One reason for the lack of difference between groups (or equal effect) could be that learners' pre-existing goal orientations, such as mastery or performance goal orientation, may override the influence of the reference frames in LADs. In such cases, learners may maintain their established goal orientation regardless of the reference frame presented.

Additionally, in reflecting on the depth and effectiveness of learners' interactions with the LADs within the observed time frames, we must consider what learners could realistically engage with and comprehend in these brief periods. Despite observing differences in the average time spent on the LADs between groups, the limited duration raises questions about the actual depth of these interactions. Were learners able to meaningfully interpret and contextualize their performance, or were these interactions more superficial due to time constraints? This question is important for understanding the quantitative aspect of time spent but also the qualitative nature of these interactions. Optimizing LAD designs to ensure informative and effective brief interactions becomes vital, and further qualitative investigations, such as user feedback or observational studies, could shed light on how learners navigate and prioritize information in LADs.

Furthermore, individual differences in learners' preferences or tendencies for making comparisons may also contribute to an absence of differential effects. Some learners may naturally engage in both temporal and social comparisons simultaneously (Pintrich et al., 2003), making it difficult to observe the distinct impacts of the reference frames. This may be particularly true in the case of those participants who receive the Social LAD because they can quite easily remember the score they received on prior steps and thus easily engage in temporal comparisons while also having access to the social comparison information not available to those with the Progress LAD. Therefore, we may expect different results if the training of each step was separated days or weeks apart. Investigating this effect under varying training schedules could therefore be a valuable area for future research.

Moreover, issues with the usability of the dashboard may have affected the learners' engagement, potentially diluting the differential effects between the two reference frames (Rets et al., 2021; Venkatesh et al., 2003). If the dashboard design did not suit the users' needs or expectations, it might have hindered their ability to effectively interpret and utilize the provided feedback. Implementing a human-centered learning analytics design process may help address usability concerns and enhance the effectiveness of the LADs (Shum et al., 2019). By refining the dashboard design based on users' preferences and needs, future research could uncover clearer distinctions in the impact of different reference frames on learners' interactions with LADs.

These findings have important implications for designing LADs to support different learning goals. For instance, if the objective is to foster a mastery goal orientation among learners, incorporating a progress reference frame in the LAD design may be more effective. Conversely, if the goal is to motivate learners through social comparison, a social reference frame may be more suitable. For example, (Fleur et al., 2023) found that LADs designed with social comparison features promoted extrinsic motivation among higher education students. However, it is important to consider the potential drawbacks of emphasizing performance goal orientation, as it may lead to a focus on surface-level learning rather than deeper understanding. For example, (Corrin & de Barba, 2015) found that providing students with feedback about their performance alongside their class average had a positive effect on motivation, half of the students in their study found it distracting from their overall learning goal.

As an alternative approach, incorporating both progress and social reference frames within a single dashboard could be beneficial in various contexts (Brusilovsky et al., 2016). Further exploration of the effects of other types of reference frames, such as those offering comparisons to assigned and self-set goals, is a promising area for future research. Furthermore, the results highlight the importance of understanding individual differences among learners. Some learners may respond more positively to temporal comparisons, while others may be more motivated by social comparisons. This relates to fit effects highlighted by (Janson et al., 2022) and is consistent with findings from (Roberts et al., 2017), which found that a majority of students interviewed agreed that being able to modify the features shown on an LAD would be valuable and that just under 70% of these students saw value in a feature which enabled them to compare their own grades with grades of their peers. Therefore, offering customizable LADs that allow learners to choose between different reference frames may help cater to diverse design preferences and optimize their learning experience.

4.4.1 Conclusions

This study examined workplace learners' interactions with two distinct LADs in a VR simulation-based training environment and compared the effects of a progress reference frame and a social reference frame on learners' interaction with LADs. The results provide early evidence that learners may interact differently with LADs depending on the reference frame used in the design. While further research is needed to clarify these findings, they have important implications for designing LADs to support different learning goals and individual learner differences.

This study contributes to the field by providing empirical evidence on how different reference frames in LAD design impact learner interactions. This research paves the way for future studies aimed at developing more adaptive and personalized LADs, and it sheds light on the potential benefits and drawbacks of using progress and social reference frames. Further investigation is needed to explore additional factors, such as corrective feedback and individual differences, that could optimize learners' experiences in simulation-based training environments.

CHAPTER 5

5

Learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students

This chapter is based on:
Gallagher, T., Slof, B., Van der Schaaf, M., Hoijtink, H., & Kester, L. (2023). Learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students. *Manuscript submitted for publication.* Department of Education, Utrecht University.

Acknowledgement of author contributions
All authors designed the study. Timothy Gallagher contributed to the study's conceptualization, methodology, formal analysis, investigation, data curation, drafting the original manuscript, manuscript review and editing, visualization, and project administration. Bert Slof, Marieke van der Schaaf, and Liesbeth Kester contributed to the conceptualization, methodology, manuscript review and editing, supervision, and funding acquisition. Herbert Hoijtink contributed to the conceptualization, statistics and methodology, and reviewing and editing of the manuscript.

Acknowledgement:
The authors would like to thank Suze Hodzelmans and the teachers of the course Designing of Learning Situations – advanced at Utrecht University for their support of the research.

Abstract

With increasing digital learning environment adoption in higher education, learning analytics dashboards (LADs) used for delivering feedback have garnered attention. LADs use various reference frames, including progress and social reference frames, to provide context for learners' self-evaluations. This study investigates how these reference frames impact academic self-efficacy, aiming to improve LAD design. We posited that comparison type—temporal or social—and comparison direction—downward, lateral, or upward—would distinctly influence learners' academic self-efficacy. A 2x3 mixed factorial switching replications design was employed, with dashboard design (progress and social) and comparison direction (downward, lateral, and upward) as factors. Bayesian Informative Hypothesis Evaluation was used to compare academic self-efficacy changes across six conditions. One hundred and forty seven university students aged between 18-54 participated. Findings suggest academic self-efficacy change varies based on comparison type and direction. Evidence shows that temporal downward comparisons may result in greater positive self-efficacy changes than social downward comparisons, while temporal upward comparisons led to greater negative changes. In upward comparison scenarios, participants seemingly discounted social comparison information when forming self-efficacy beliefs. Further research is required to confirm these results. These results underscore the importance of strategic LAD design, considering how different types and directions of comparisons can influence self-efficacy. By highlighting the potential impact of comparison elements on academic self-efficacy, our study provides valuable insights for creating more effective LADs, contributing to the discourse on the influence of digital learning environment design on academic self-efficacy.

5.1 Introduction

Amid the rising adoption of digital learning environments in higher education, enhancing academic self-efficacy through the strategic use of learning analytics has emerged as a significant objective (Jonathan et al., 2022; Karaoglan Yilmaz, 2022). Elevated academic self-efficacy, which refers to a learner's judgment about their ability to successfully attain educational goals, can lead learners to maintain effort, persevere through challenges, and perform at a higher level (Elias & MacDonald, 2007; Schunk & Pajares, 2002). Learning analytics dashboards (LADs), as common learning analytics tools, gather various indicators into visualisations to inform both teaching and learning decisions (Schwendimann et al., 2017). This includes providing performance feedback to learners, which aids in self-evaluation processes which are directly related to self-efficacy (Matcha et al., 2020; Zimmerman, 2000b, 2002). Therefore, prioritising academic self-efficacy is crucial for those involved in the design and deployment of educational tools like LADs. Additionally, the design of LADs must take into account the impact of diverse reference frames on students' self-efficacy, as these can subsequently affect learning behaviours and outcomes (van Leeuwen et al., 2023).

Specifically, LADs employ reference frames to contextualise feedback, offering various comparison points like performance data . These frames help learners understand their performance by facilitating self-evaluation. For example, a progress reference frame assesses a learner's past and present performances, encouraging temporal comparison (Albert, 1977; Wilson & Shanahan, 2020). Meanwhile, a social reference frame compares a learner's achievements with the average performance of peers, enabling social comparison (Festinger, 1954; Gerber et al., 2018).

The use of reference frames in LADs not only shapes learners' interpretation of their performance but also potentially influences their academic self-efficacy. Understanding the long-term effects of different reference frames on academic self-efficacy can offer critical insights for LAD design aimed at supporting students' academic success (van Leeuwen et al., 2023). Given the pivotal role of academic self-efficacy in influencing educational outcomes, it becomes important to understand how LADs could shape this key learning related variable. Lack of understanding could result in LAD designs that inadvertently undermine self-efficacy, thereby reducing their potential educational impact. There is a growing body of evidence linking academic self-efficacy to key academic outcomes like performance and grade point average (Bartimote-Aufflick et al., 2016; Ferla et al., 2009; Luszczynska et al., 2005; Richardson et al., 2012; Zimmerman et al., 1992). This evidence forms a strong basis for exploring the impact of various LAD designs on academic self-efficacy, an essential consideration for ensuring that LADs are genuinely beneficial to learners (Gasevic et al., 2015).

Another factor affecting how learners contextualise feedback is the "direction of comparison," which is their achievement level relative to a specific comparison point (Collins, 1996; Suls et al., 2002). Central to establishing this direction is the 'score delta,' a measurable gap representing the difference between a learner's performance and their chosen point of comparison. This score delta not only sets the direction of comparison but also affects self-evaluation outcomes. Depending on both the reference frame and the score delta value, learners may view their performance as superior, equivalent, or inferior, which results in downward, lateral, or upward comparisons, respectively (Guyer & Vaughan-Johnston, 2018; Suls et al., 2002).

Each type of comparison (i.e., social and temporal) and direction (i.e., downward, lateral, or upward) can yield different self-evaluation outcomes (Collins, 1996). For example, a downward social comparison could lead to a positive self-evaluation if learners perceive themselves as outperforming their peers (Wheeler & Suls, 2020). Conversely, an upward social comparison might positively influence self-evaluation by motivating learners to aspire to their peers' level of success. However, an upward temporal comparison could either undermine self-evaluation if learners feel they are not making adequate progress (Wilson & Shanahan, 2020) or be dismissed if they think their subpar or declining performance doesn't accurately represent their capabilities (Zell & Alicke, 2010).

The objective of this study is to investigate how the nature of comparisons—comprising both types (i.e., social and temporal) and directions (i.e., downward, lateral, upward)—offered by LADs affects academic self-efficacy. We hypothesise that these specific combinations will have distinct impacts on learners' academic self-efficacy. To provide background for our study, we will first review relevant LAD research. Then, we will outline a framework linking progress and social reference frames to their corresponding theories of comparison. Our discussion will conclude by examining self-efficacy theory and two key sources of information—mastery experiences and social modelling—that influence academic self-efficacy beliefs.

This study was pre-registered on the Open Science Framework. All data and supplementary materials can be accessed in the associated repository at https://osf.io/xzr8f.

5.1.1 Theoretical Framework

Learning Analytics Dashboards and Reference Frames

Although educational research and learning theory are called on to inform LAD design (Gasevic et al., 2015; Matcha et al., 2020) evidence suggests that these dashboards frequently lack a robust theoretical foundation (Jivet et al., 2017). Such oversight can lead to ineffective or even detrimental learning experiences, as choices in design elements like comparison points significantly influence learner interpretation (Matcha et al., 2020). For example, Corrin and de Barba (2015) noted that some students were distracted from their educational objectives when they accessed peer performance data through an LAD intervention.

To address the gap in theory-informed LAD design, researchers are building a solid evidence base for designers. For instance, Jivet et al. (2018) offer five recommendations for LAD design: consider social and emotional learning factors, grant students control over their data, ensure data transparency and interpretability, integrate LADs with other educational tools, and engage both students and educators in the design process. Similarly, Bennett & Folley (2020) propose four design principles: allow students to customise their LADs, incorporate design elements that facilitate sense-making, enable the identification of actionable insights, and seamlessly integrate into the educational workflow.

Other research has explored learners' preferences for different LAD reference frames (Jivet et al., 2020), a critical factor as these perceptions can affect tool adoption and use (Buckingham Shum et al., 2019). For example, a study by Guerra et al. (2016) used a usability survey to demonstrate that both social and progress frames are valued by highly motivated learners. However, understanding learner preferences is not enough; studies investigating the impact of design elements on learning-related constructs are also necessary. To illustrate, Jonathan et al. (2022) revealed that LADs featuring both types of reference frames equally improved critical reading self-efficacy. Moreover, Guerra et al. (2016) assessed an LAD built on theories of self-regulated learning and goal orientation, finding that the social reference frame positively affected student engagement and performance.

While some studies have examined both progress and social reference frames, the emphasis has largely been on the impact of social comparison on learning outcomes. Research by Fleur et al., (2023), Beheshitha et al. (2016), and Davis et al., (2017) highlights the importance of understanding this relationship for LAD design, particularly in influencing motivation, participation, and self-regulation. A significant advancement in aiding LAD designers has been identifying the specific points of comparison used within LADs.

Operational Framework for Learning Analytics Reference Frames

In synthesising existing research, we introduce an operational framework for learning analytics reference frames. This framework integrates theories of self-regulated learning (Zimmerman, 2002; Zimmerman & Moylan, 2009), temporal comparison theory (Albert, 1977; Wilson & Shanahan, 2020) and social comparison theory (Festinger, 1954; Gerber, 2019; Wheeler & Suls, 2020; Zell & Alicke, 2010). The framework identifies three primary components: performance outcome, point of comparison, and score delta. The performance outcome provides learners feedback, via an LAD, on task accomplishment. The point of comparison, influenced by temporal and social comparison theories, aids in learner self-evaluation. This may encompass a comparison with personal performance or peers' performance. The score delta, based on theories of comparison direction, signifies the score disparity between the performance outcome and the point of comparison, leading to downward, lateral, or upward comparisons. Table 1 offers a detailed overview of these components within the framework.

Table 1
Operational Framework for Learning Analytics Reference Frames

Component	Related Theories	Description	Example
Performance Outcome	Sources of Self-efficacy: Mastery experiences information	Feedback provided to a learner, by an LAD, about how well they have performed a particular task.	Sarah's LAD displays that she has received a score of 85% on her recent exam.
Point of Comparison	Sources of Self-efficacy: Mastery experiences information and Social modelling information	Defines the reference frame type, informed by temporal and social comparison theory, and refers to the point of comparison being offered to the learner to aid self-evaluation.	For Sarah, points of comparison include her previous exam score of 80% and her peers' average score of 82%.
Score Delta	Theories of Directions of Comparison: Upward, Lateral and Downward comparison	The difference in score between the performance outcome and point of comparison, linked to theories related to the direction of comparison, which can lead to downward, lateral, or upward comparison.	Sarah's score delta leads to a +5% downward temporal comparison (compared to her previous score (80%)) and a +3% upward social comparison (compared to her peers' average (82%)).
Reference Frame Types			
Progress Reference Frame	Temporal Comparison Theory	Offers a point of comparison in the form of performance of an earlier self.	Sarah's LAD shows her previous exam score of 80%, allowing her to compare her current results (85%) with her past performance and showing her that her performance increased with 5%.
Social Reference Frame	Social Comparison Theory	Offers the performance of peers as the point comparison.	Sarah's LAD presents her score of 85% alongside the average performance of her peers, which is 82%, allowing to compare her current results (85%) to those of her peers and showing her that her current score is 3% higher than that of her peers.

Temporal comparison theory posits that individuals assess their capabilities by comparing their current and past performances (Albert, 1977; Wilson & Shanahan, 2020). In the context of LADs, learners are often presented with historical performance data on similar tasks, facilitated by a progress reference frame (Jivet et al., 2017). When recent performance surpasses previous benchmarks, a more favourable self-evaluation generally follows (Wilson & Ross, 2000; Wilson & Shanahan, 2020). Therefore, we expect positive self-evaluations during downward temporal comparison and negative self-evaluations during upward temporal comparison. However, studies suggest that learners may not consistently employ temporal comparison in self-evaluation processes (Zell & Alicke, 2010). Investigating the influence of temporal comparison across varying directions of comparison within LADs is crucial for establishing a robust evidence base for dashboard designers.

Social comparison theory posits that individuals derive their self-evaluations from comparisons with one or more peers (Suls et al., 2002; Zell & Alicke, 2010). Within the LAD context, learners often receive aggregated peer performance data, presented through a social reference frame. Generally, downward social comparison has a positive impact on self-evaluations, while upward social comparison has a negative effect (Gerber et al., 2018). However, studies have shown that upward comparison, even when it reveals poorer performance, does not necessarily undermine self-evaluations (Mussweiler, 2003). Similar to temporal comparison, examining the effects of social comparison across varied comparison directions is crucial for optimising LAD design.

Having outlined the framework for learning analytics reference frames, we now turn to its relationship with academic self-efficacy, underscoring its significance for LAD design considerations.

Academic Self-Efficacy and Learning Analytics Dashboards

Academic self-efficacy refers to the belief in one's ability to use motivational, cognitive, emotional, behavioural, and social resources to achieve educational goals (Nielsen et al., 2017; Richardson et al., 2012). Students with high academic self-efficacy are more likely to invest effort, overcome obstacles, and achieve better academic outcomes (Schunk & Pajares, 2002). Understanding the impact of academic self-efficacy on student performance is essential for LAD designers who aim to improve it.

Therefore, academic self-efficacy should be a priority for those designing educational tools like LADs. Given that LADs offer various forms of performance feedback, it is crucial to understand how different feedback types can influence academic self-efficacy. The LAD design must consider the effects of diverse reference frames on students' self-efficacy, as this can subsequently impact their learning behaviours and outcomes (van Leeuwen et al., 2023). LAD designers can address this by considering the information sources learners use to form their self-efficacy beliefs.

Bandura (1997) identifies mastery experiences and social modelling as key influences on self-efficacy beliefs. In a learning context, mastery experiences refer to a learner's task performance and the evidence it provides about their capability to complete the task (Bandura, 1997). When learners review this information via an LAD employing a progress reference frame, they engage in temporal comparisons. Consequently, we anticipate that temporal comparisons, facilitated by progress reference frames, will serve as a form of mastery experiences information that shape learner self-efficacy.

In the learning context, social modelling refers to how the performance of others, such as peers, can influence one's self-efficacy (Bandura, 1997). When learners review such information through an LAD using a social reference frame, they engage in social comparisons. Hence, we anticipate that social comparisons, facilitated by social reference frames, will act as a form of social modelling information that shapes self-efficacy beliefs.

Mastery experiences are often considered the most influential source of self-efficacy information, providing "the most authentic evidence of whether one can muster whatever it takes to succeed" (Bandura, 1997, p. 80). However, the extent to which this is true for LADs offering either mastery experiences or social modelling information is ambiguous. Previous research indicates that social comparison can have a potent effect on self-evaluations (Van Yperen & Leander, 2014).

Previous research on LADs has largely centred on the effects of social comparison on learning related variables, leaving a gap in understanding the role of mastery experiences within progress reference frames on academic self-efficacy. Moreover, the comparative impact of mastery experiences and social modelling on academic self-efficacy in LAD contexts remains underexplored.

Research Question

How does learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students?

Bayesian Informative Hypotheses

To address the research question, we employ Bayesian Informative Hypothesis Evaluation, which allows for direct comparison between competing hypotheses and quantifies their relative support using the data (Hoijtink et al., 2019; Van Lissa et al., 2021). This method moves beyond traditional dichotomous decisions based on p-values and significance thresholds, using posterior model probabilities to measure hypothesis support instead. While we offer our interpretation of the findings, the ultimate judgment on the sufficiency of evidence to reject hypotheses rests with the scientific community.

Each hypothesis represents a distinct scenario, informed by both theory and empirical evidence, allowing for a thorough examination of our research question. This approach investigates the interplay between types of comparison (temporal, social) and directions (downward, lateral, upward) and their impact on academic self-efficacy.

An informative hypothesis is a "theoretically derived statement about directional differences and equality constraints between model parameters of interest" (Van Lissa et al, 2020, p. 2). We formulate each hypothesis using terms of equality (=) and inequality (<, >) and groups of parameters can be constrained using parentheses. The theoretical components which support our hypotheses are outlined above and elaborated upon below.

Competing Informative Hypotheses

The following hypotheses include those that were part of our original pre-registration as well as additional hypotheses introduced to provide a more comprehensive analysis. These additional hypotheses (H1, H4, H5, H8) were not part of the original pre-registration but are consistent with our initial theoretical framework and are supported by empirical evidence.

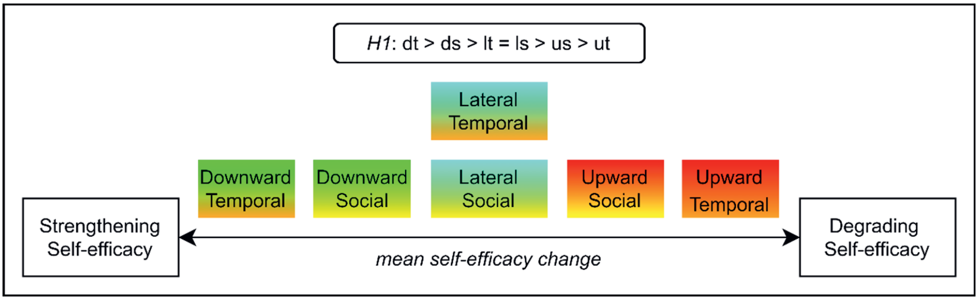
Hypothesis 1 (H1):

The mean self-efficacy change will decrease as the comparison shifts from downward temporal to downward social, then to lateral temporal/social, and finally to upward social and upward temporal conditions (Figure 1).

This is attributed to the overriding strength of mastery experiences information in the downward and upward conditions, while both sources of information, mastery experiences and social modelling, are expected to have an equal effect in the lateral conditions.

Figure 1

Hypothesis 1 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

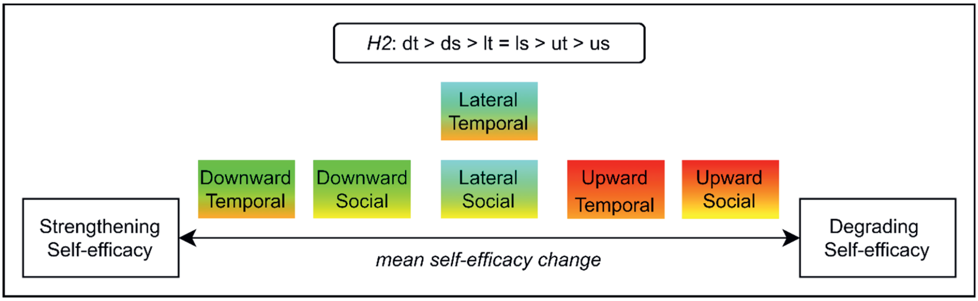
Hypothesis 2 (H2):

The mean self-efficacy change will decrease as the comparison shifts from downward temporal to downward social, then to lateral temporal/social, and finally to upward temporal and upward social conditions (Figure 2).

This is attributed to the overriding strength of mastery experiences information in the downward condition and the overriding strength of social modelling information in the upward condition. Both sources of information are expected to have an equal effect in the lateral conditions.

Figure 2

Hypothesis 2 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

5

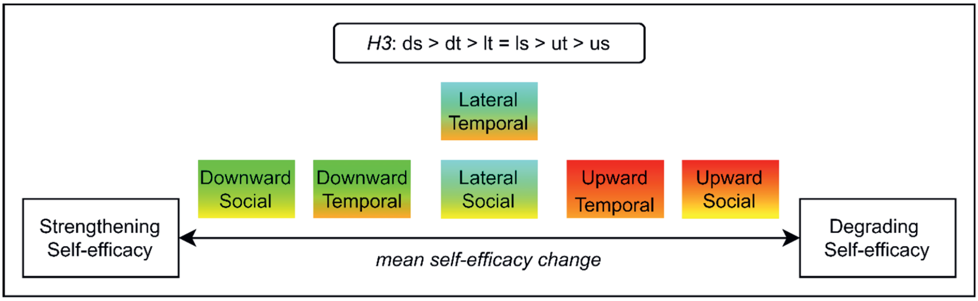
Hypothesis 3 (H3):

The mean self-efficacy change will decrease as the comparison shifts from downward social to downward temporal, lateral temporal/social, upward temporal, and upward social conditions (Figure 3).

This is attributed to the overriding strength of social modelling information in the downward and upward conditions, while both mastery experiences and social modelling information are expected to have an equal effect in the lateral conditions.

Figure 3

Hypothesis 3 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

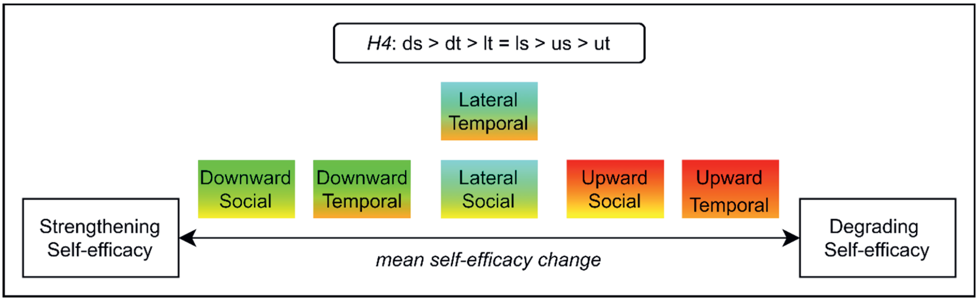
Hypothesis 4 (H4):

The mean self-efficacy change will decrease as the comparison shifts from downward social to downward temporal, lateral temporal/social, upward social, and upward temporal conditions (Figure 4).

This is attributed to the overriding strength of social modelling information in the downward condition, the overriding strength of mastery experiences information in the upward condition, and the expectation that both types of information will have an equal effect in the lateral conditions.

Figure 4

Hypothesis 4 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

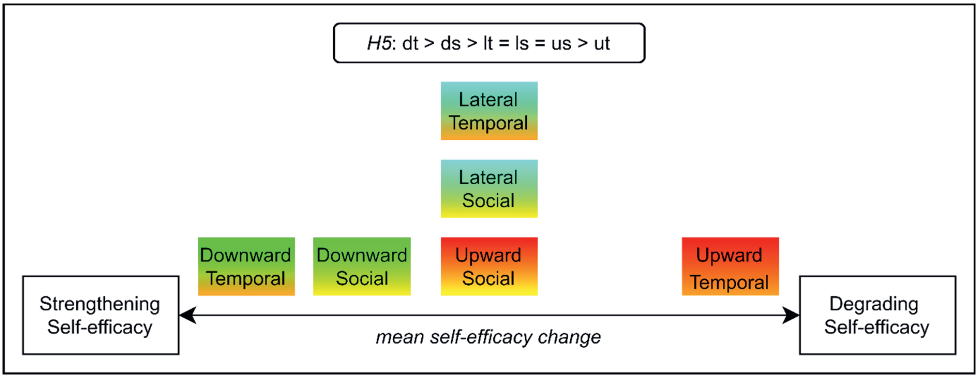
Hypothesis 5 (H5):

The mean self-efficacy change will decrease as the comparison shifts from downward temporal to downward social to lateral temporal/social and upward social to upward temporal conditions (Figure 5).

This is attributed to the overriding strength of mastery experiences information in the downward and upward conditions. In the upward social condition, learners may perceive their relatively poorer performance compared to their peers as having no bearing on their own capacity, resulting in social modelling information having an equal effect to mastery experiences information in the lateral conditions.

Figure 5

Hypothesis 5 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

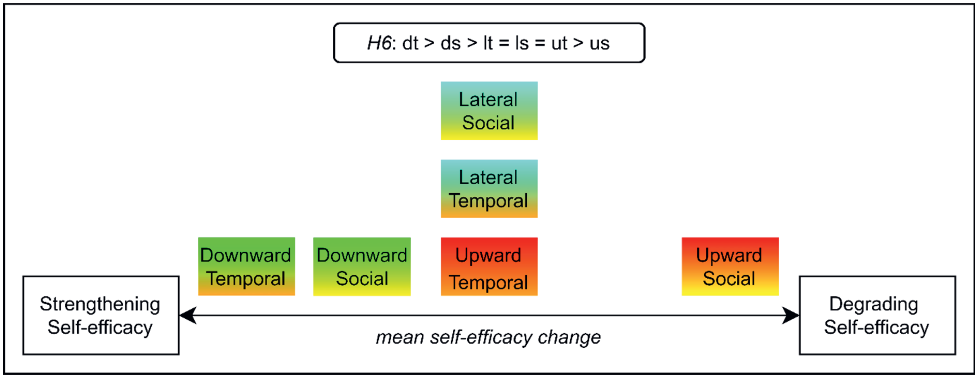
Hypothesis 6 (H6):

The mean self-efficacy change will decrease as the comparison shifts from downward temporal to downward social to lateral temporal/social and upward temporal to upward social conditions (Figure 6).

This is attributed to the overriding strength of mastery experiences information in the downward condition and the overriding strength of social modelling information in the upward condition. In the upward temporal condition, learners may perceive their declining performance over time as not reflecting their true ability, leading them to discount it, resulting in upward condition mastery experiences information having an equal effect to social modelling and mastery experiences information in the lateral conditions.

Figure 6

Hypothesis 6 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

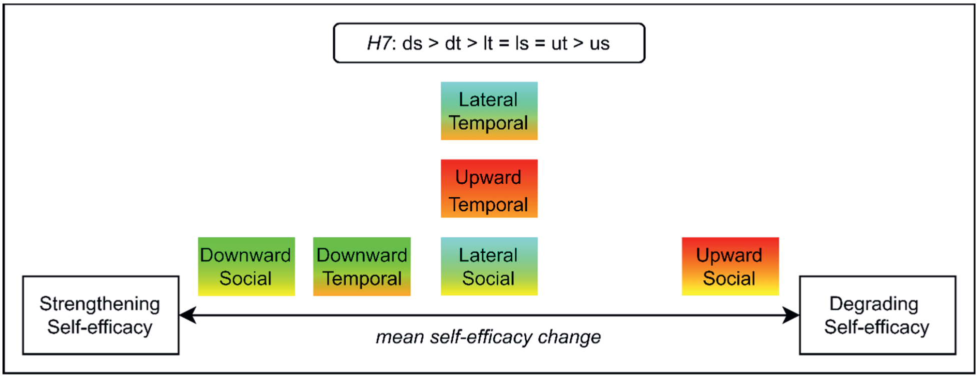
Hypothesis 7 (H7):

The mean self-efficacy change will decrease as the comparison shifts from downward social to downward temporal to lateral temporal/social and upward temporal to upward social conditions (Figure 7).

This is attributed to the overriding strength of social modelling information in the downward and upward conditions. In the upward temporal condition, learners may perceive their declining performance as not reflecting their true ability, leading them to discount it, resulting in upward condition mastery experiences information having an equal effect to social modelling and mastery experiences information in the lateral conditions.

Figure 7

Hypothesis 7 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

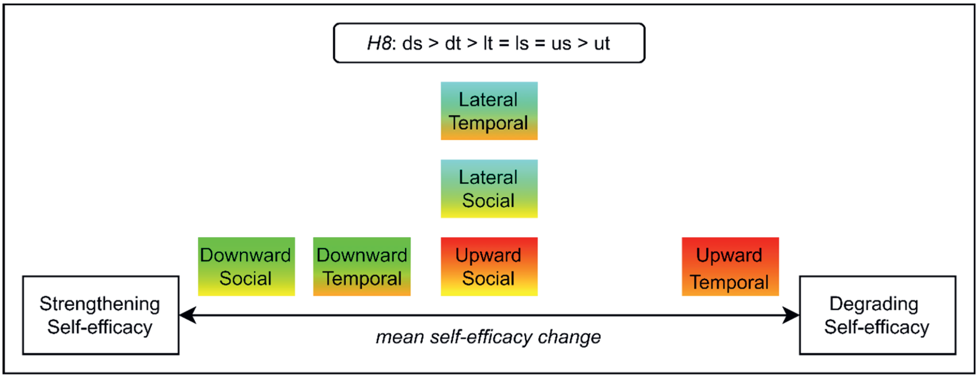
Hypothesis 8 (H8):

The mean self-efficacy change will decrease as the comparison shifts from downward social to downward temporal to lateral temporal/social and upward social to upward temporal conditions (Figure 8).

This is attributed to the overriding strength of social modelling information in the downward condition and mastery experiences information in the upward condition. In the upward social condition, learners may perceive their relatively poorer performance compared to their peers as having no bearing on their own capacity, resulting in upward social modelling information having an equal effect to social modelling and mastery experiences information in the lateral conditions.

Figure 8

Hypothesis 8 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions.

Hypothesis 9 (H9):

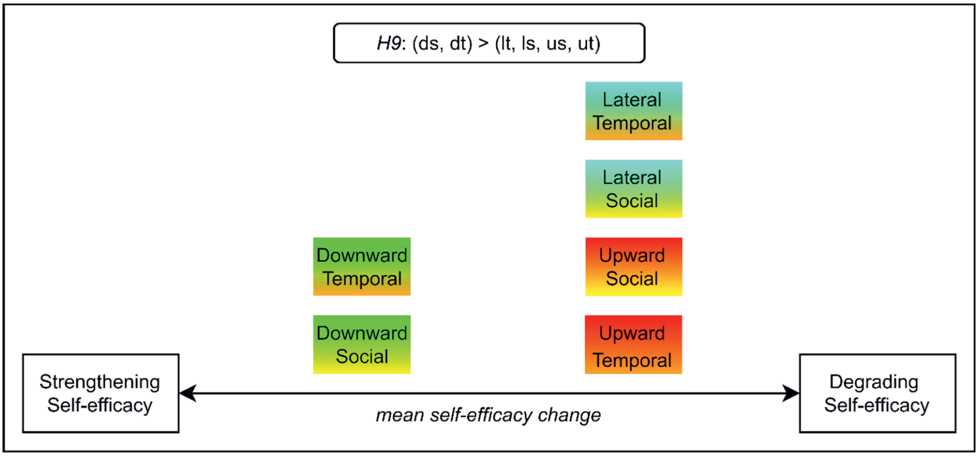
The mean self-efficacy change will decrease as the comparison shifts from the downward condition to the lateral and upward conditions, with no notable influence from mastery experiences or social modelling information (Figure 9).

This is attributed to the overriding effect of downward comparison on lateral and upward comparison, and no effect of upward comparison compared with lateral comparison due to mastery experiences and social modelling information being ignored or discounted in the upward condition.

5

Figure 9

Hypothesis 9 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions. Note that, where stacked groups in the previous figure represented means constrained to be equal, now they represent means that are mutually unconstrained.

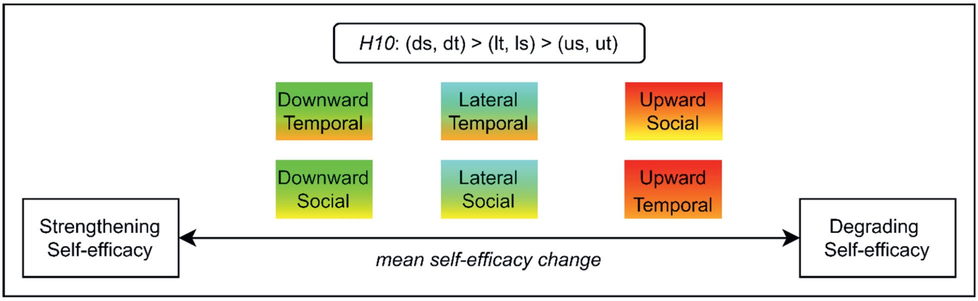
Hypothesis 10 (H10):

The mean self-efficacy change will decrease as the comparison shifts from downward comparison to lateral comparison to upward comparison, with no significant influence from mastery experiences or social modelling information (Figure 10).

This is attributed to the overriding effects of downward, lateral, and upward comparisons, which override the influence of mastery experiences and social modelling information.

Figure 10

Hypothesis 10 graphical representation



Note. Abbreviations and symbols used in the hypotheses, in reference to mean self-efficacy change, include t = temporal, s = social, d = downward, l = lateral, u = upward, > = greater than, < = less than, = = equal to. Combined abbreviations represent conditions. Note that, stacked groups represent means that are mutually unconstrained.

5.2 Methods

We explore the differential effects of two LAD designs on university students' academic self-efficacy, specifically examining the role of downward, lateral, and upward comparisons in self-efficacy change.

5.2.1 Participants

The study was conducted during an eight-week university course called ‘Designing of Learning Situations – advanced.”. Of the 167 enrolled students, 147 consented to participate: 110 were female, 35 male, and 2 opted not to disclose their gender. Ages ranged from 18 to 54, with a majority (n=120) between 18-24 years, followed by 18 in the 25-34 bracket, 8 in the 35-44 bracket, and 1 in the 45-54 bracket. Informed consent was obtained from all participants, who were free to withdraw at any time. The study followed ethical guidelines and was approved by an ethical committee.

5.2.2 Study Design

The study employed a 2 x 3 mixed factorial switching replications design. This design enabled the examination of multiple dashboard designs on the same participants, thus mitigating the influence of individual differences. The first factor, "Dashboard Design," had two levels: Progress and Social dashboards. The second factor, "Direction of Comparison," included three levels: Downward, Lateral, and Upward comparisons. This resulted in six conditions characterised by different combinations of score deltas on either dashboard.

Exclusion Criteria and Group Assignment

Conditions were excluded if they occurred fewer than 10 times. All participants interacted with both dashboard types. They were divided into two groups: Group 1 used the Progress dashboard during weeks two to five and switched to the Social dashboard for weeks six to eight. Conversely, Group 2 used the Social dashboard first and then switched to the Progress.

Data Collection and Self-Efficacy Assessment

Academic self-efficacy was measured weekly following each LAD review. Changes were calculated based on the difference between consecutive weeks' scores. A total of 471 instances met the analysis criteria.

Nature of Comparison

The independent variables can be described as natures of comparison of which there are six: (1) downward temporal comparison, (2) downward social comparison, (3) lateral temporal comparison, (4) lateral social comparison, (5) upward temporal comparison, (6) upward social comparison.

5.2.3 Materials

Course Description: Designing of Learning Situations – Advanced

The eight week course "Designing of Learning Situations - advanced" course, aimed to provide students with knowledge and skills related to the instructional design of training complex skills. It comprised four main elements: knowledge clips, tutorials, Q&A sessions, and assessments that included an instructional design assignment and a take-home exam. Students were required to engage with weekly subject-specific knowledge clips, facilitated by the Feedback Fruits platform, which allows interactive quizzes and discussions to be integrated directly into video content. Clips ranged from 5-30 minutes and incorporated multiple-choice or true/false questions. For instance, one question posed was, "What 3 factors are necessary for complex learning?" with answer choices of "Integration, Coordination, Transfer," "Compartmentalisation, Fragmentation, Transfer paradox," and "Segmentation, simplification, and fractionation.

For the purpose of this study, we made a slight modification to the original course design. Students were asked to use an LAD and complete an academic self-efficacy questionnaire at the beginning of each tutorial. This procedure took approximately ten minutes and was the same for all participating students. Non-participating students did not complete the questionnaire and their data was not included in the analysis. The LADs were populated with data from students' responses to the knowledge clip questions and aimed to assist in monitoring their course literature mastery. Feedback based on their answers was provided through links to the Feedback Fruits platform.

The Learning Analytics Dashboards

LADs were used to provide students with performance metrics, both individual and comparative. These dashboards were developed using Microsoft Power BI and shared with students via a link from their respective tutorial group teachers.

Two types of dashboards were used: the Progress and the Social dashboards. The Progress dashboard used a progress reference frame, presenting the student's weekly score alongside their own historical data and the temporal score delta—the week-to-week difference in individual performance. Conversely, the Social dashboard provided a social reference frame, displaying the student's weekly score, the weekly average score of all students, and the social score delta—the difference between the individual's and the group's average scores.

The score delta served as a metric of relative performance: a positive delta signified scoring above the average, a negative delta below, and a zero delta indicated parity. All scores were rounded to the nearest 5%. Figure 11 illustrates the Progress dashboard, and Figure 12 displays the Social dashboard.

Figure 11

Progress dashboard

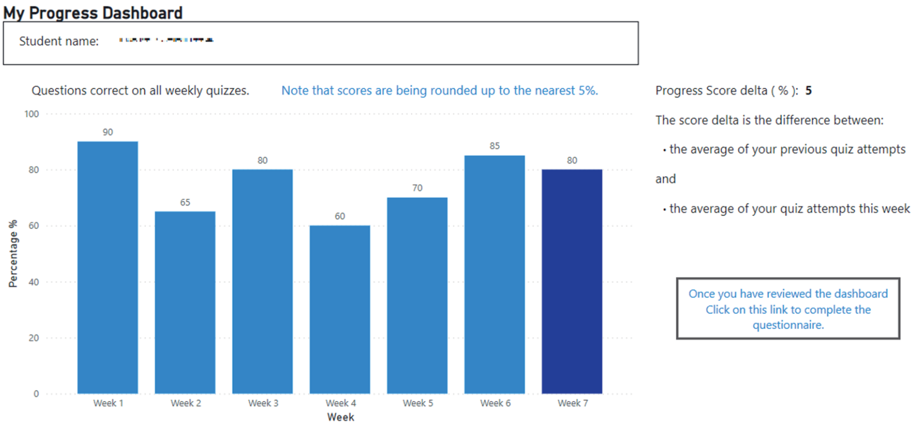
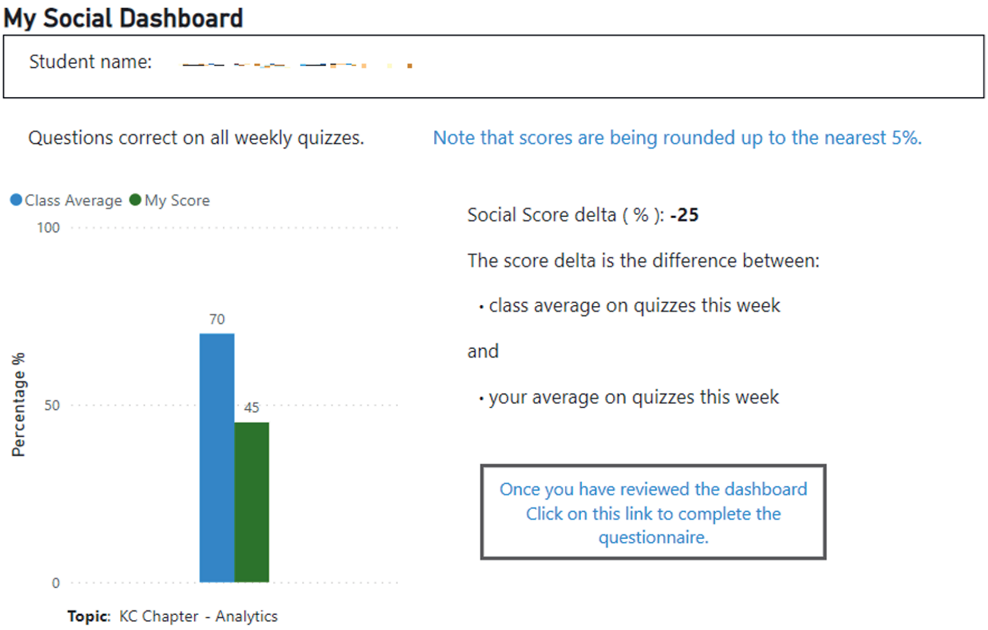


Figure 12

Social dashboard



Note. ‘KC Chapter – Analytics’ denotes the title of the chapter under study.

Each dashboard, titled either “My Progress Dashboard” or “My Social Dashboard” based on the experimental group, displayed the student’s name below the title. The core feature was a labelled bar chart. In the Progress dashboard, the x-axis denoted the study weeks, and the y-axis represented quiz performance as a percentage of correct answers. Conversely, the Social dashboard’s x-axis related to study topics while maintaining the y-axis for quiz performance. Both dashboards clarified that scores were rounded to the nearest 5% and included the score delta along with its calculation methodology. A box containing a link for a follow-up questionnaire completed the layout.

5.2.4 Measures

Demographic Questionnaire

A demographic questionnaire obtained information about participants age and gender for descriptive purposes.

Academic Self-efficacy

Originally, our study intended to use the Specific Academic Learning Self-efficacy (SAL-SE) and the Specific Academic Exam Self-efficacy (SAE-SE) scales as suggested by Nielsen et al., (Nielsen et al., 2017). However, in a deviation from our pre-registered plan, we opted for the more established Motivated Strategies for Learning Questionnaire Self-efficacy Sub-Scale (MSLQ-SE) developed by Pintrich (1991). This decision was made to take advantage of the scale's broader acceptance and empirical validation (Taylor, 2012). The MSLQ-SE consists of a total of eight items, which is a combination of the eight items originally split between the SAL-SE and SAE-SE scales. Each item assesses students' beliefs about their capabilities to perform specific academic tasks (e.g., 'I expect to do well in this class.'). Participants rated items on a five-point Likert scale, ranging from 1= Strongly disagree to 5 = Strongly agree. The scale thus provides a measure of the extent to which students agree or disagree with each statement, serving as an indicator of their academic self-efficacy.

Procedure

Participants were initially informed about the study via pre-course email and formally invited during the first teaching session. Upon agreeing to participate, they received an information letter that detailed the study design, data confidentiality, and the voluntary nature of participation, including the option to withdraw at any time without repercussions. After signing the informed consent, they completed demographic and academic self-efficacy questionnaires.

All participants belonged to one of nine tutorial groups, which were randomly assigned to either experimental Group 1 or Group 2. Group 1 accessed the Progress dashboard from weeks two to five and switched to the Social dashboard for weeks six to eight, while the schedule was reversed for Group 2. Both groups interacted with the LADs identically.

At each tutorial's outset, students received an email link to their respective dashboard, requiring their student email and password for access. Teachers prompted them to review the LADs and reflect on their performance before completing an embedded academic self-efficacy questionnaire. After completing the questionnaire, students were directed to Feedback Fruits for quiz feedback. To control access, the link to the LADs was active only during the designated tutorial sessions.

Scoring: Calculation of Academic Self-efficacy Change

Changes in academic self-efficacy were assessed by comparing consecutive weeks' self-efficacy scores. Conditions resulting in less than 10 observations were excluded. If a participant's data lacked consecutive weeks of self-efficacy scores, that data set was omitted.

We then assigned:

- progress dashboards with score deltas > 0 to the downward temporal comparison condition.
- social dashboards with score deltas > 0 to the downward social comparison condition
- progress dashboards with score deltas = 0 to the later temporal comparison condition
- social dashboards with score deltas = 0 to the lateral social comparison condition
- progress dashboards with score deltas < 0 to the upward temporal comparison condition

- social dashboards with score deltas < 0 to the upward social comparison condition

To illustrate this process, consider a participant named Sarah. In the first week, Sarah scores 75% on the quiz, and 80% in the second week, giving her a score delta of +5%. According to our design, this indicates a downward temporal comparison due to her improved performance. Sarah's weekly self-efficacy scores are also recorded to track its fluctuation over time. Any changes, or lack thereof, are matched to the relevant dashboard condition. In Sarah's case, her self-efficacy changes would be attributed to the downward temporal comparison category.

In summary, each participant's eligible academic self-efficacy change score is collated based on their exposure to a particular dashboard condition. For instance, Sarah's scores would be aggregated with those from participants experiencing the same downward temporal comparison induced by a positive score delta on the Progress dashboard.

5.2.5 Analysis

Reliability and Validity Check for Academic Self-efficacy

The statistical analysis of the reliability of the academic self-efficacy questionnaires was carried out using JASP version 0.16. Due to constraints on sample size, a full confirmatory factor analysis of our study sample was infeasible. Nonetheless, the MSLQ has been widely used in educational research and has exhibited strong reliability and validity evidence in previous studies. The MSLQ's reliability was initially reported as 0.93 (Pintrich, 1991) and a meta-analysis further confirmed a mean reliability distribution of 0.91 (SD = 0.02) and prospective and concurrent validity coefficients ranging from 0.31 to 0.58 (Credé & Phillips, 2011). Moreover, the MSLQ has undergone formal validity and reliability assessment in multiple languages including Spanish and Chinese, and was utilised in more than 50 research studies within a period of just five years (2000 and 2005) (Duncan & McKeachie, 2005).

In the current study, even though we were unable to perform a confirmatory factor analysis due to our sample size, we conducted a unidimensional reliability test on our data using the MSLQ-SE. We calculate Cronbach's α for our data to estimate the reliability.

Statistical Models and Bayesian Hypothesis Evaluation

In our analysis, we adopted a two-level model featuring change in self-efficacy as the dependent variable, and a factor encompassing conditions with 10 or more observations as the predictor. The model originally included a random intercept, as per our pre-registration. However, during the analysis phase, we discovered zero variance in the random intercept. Consequently, we deviated from our pre-registered plan to employ a two-level random intercept model using 'lmer.' Instead, we used the 'lm' function from the base R package, thus eliminating the random intercept. This adjustment has been updated in the Open Science Framework repository linked to our pre-registration.

Two factors, namely 'Dashboard Design' and 'Direction of Comparison', were defined. 'Dashboard Design' consisted of two levels (Progress Dashboard, Social Dashboard), while 'Direction of Comparison' had three (Downward Comparison, Lateral Comparison, Upward Comparison), resulting in six distinct conditions. We compared the mean change in academic self-efficacy across these conditions using the Bayes factor, as implemented in 'bain' (Version 0.2.8, [R package] https://CRAN.R-project.org/package=bain) (Hoijtink et al., 2019; Van Lissa et al., 2021).

Hypotheses 1 through 10 will be evaluated using so-called posterior model probabilities (Hoijtink et al., 2019). These probabilities indicate the likelihood that each hypothesis is the best fit for the collected data, with the sum of all PMP values equalling 1.0. If one hypothesis is preferred, the sum of the posterior model probabilities of the other hypotheses constitutes the Bayesian error. For example, if H1 is preferred with a PMP value of .8, the Bayesian error would be .2.

In line with Hoijtink et al. (Hoijtink et al., 2019) we conducted a sensitivity analysis using fractions of 1, 2, and 3 and reported each result (the posterior model probabilities (PMP)), providing a joint interpretation of these three sets of results. This analysis highlights the effects of changing the prior distribution underlying the computation of the PMPs.

5.3 Results

5.3.1 Reliability estimation of the Academic Self-efficacy Questionnaire

To assess the reliability of the MSLQ-SE, we calculated Cronbach's alpha for each week the questionnaire was administered, and Table 2 reports the results which show good reliability for each week.

Table 2

Summary of Estimated Reliability Over an 8-Week Period for MSLQ-SE

Time Point	α
Week 1 (n= 147)	.86
Week 2 (n= 132)	.91
Week 3 (n= 114)	.92
Week 4 (n= 132)	.92
Week 5 (n= 99)	.95
Week 6 (n= 45)	.94
Week 7 (n= 74)	.96

Note. α = Cronbach's Alpha, n = number of participants

5.3.2 Descriptive Statistics of Academic Self-Efficacy Change

We performed a descriptive analysis to examine changes in academic self-efficacy, segmented by dashboard type and comparison direction (Table 3). Academic self-efficacy was gauged on a 5-point Likert scale, allowing for a theoretical change range of -4 to 4. The observed mean changes in self-efficacy were minimal across all comparison types, ranging from -0.042 under the 'Social - downward' condition to 0.038 in the 'Social - lateral' scenario (Cohen’s d = -0.207). These findings suggest that while dashboard type and comparison direction do influence academic self-efficacy, the real-world impact of these effects may be modest.

Table 3

Descriptive Statistics of Academic Self-Efficacy Change by Dashboard Type and Direction of Comparison

Nature of comparison	Frequency	Mean Self-efficacy Change (SD)	Min	Max
Temporal - downward	84	−0.013 (0.353)	-1.25	0.88
Temporal - lateral	11	0.023 (0.200)	-0.25	0.38
Temporal - upward	186	−0.041 (0.367)	-1.50	1.13
Social - downward	99	−0.038 (0.384)	-2.00	1.00
Social - lateral	30	0.037 (0.286)	-0.63	1.00
Social - upward	61	0.004 (0.376)	-1.00	1.13

Note. Numbers in brackets are standard deviations from the mean.

5.3.3 Analysing the Impact of Learning Analytics Reference Frames on Students' Academic Self-Efficacy.

To answer the research question, 'How does learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students?' we evaluated ten hypotheses using Bayesian Informative Hypothesis Testing. Each hypothesis proposed a unique relationship between two types of comparison—social and temporal—and three directions of comparison—downward, lateral, and upward.

The results of our Bayesian informative hypothesis testing are reported in Table 4. Here, the PMPa values for each hypothesis at the three different fractions are displayed. As indicated by the PMPa values, H5 received the highest support across all fractions (0.357, 0.324, 0.304), followed by H8 with PMPa values of 0.250, 0.238, and 0.210, respectively. The remaining hypotheses received lower support, with H6 (0.121, 0.111, 0.109) and H1

(0.081, 0.106, 0.115) showing moderate support and H2, H3, H4, H7, H9 and H10 showing minimal support.

Table 4

Posterior Model Probabilities for Different Hypotheses at Varying Fractions in order of support.

Hypothesis	PMP a*	PMP a**	PMP a***
H5: td > sd > tl = sl = su > tu	0.357	0.324	0.304
H8: sd > td > tl = sl = su > tu	0.250	0.238	0.210
H6: td > sd > tl = sl = tu > su	0.121	0.111	0.109
H1: td > sd > tl = sl > su > tu	0.081	0.106	0.115
H7: sd > td > tl = sl = tu > su	0.074	0.070	0.067
H4: sd > td > tl = sl > su > tu	0.058	0.070	0.088
H2: td > sd > tl = sl > tu > su	0.032	0.037	0.045
H3: sd > td > tl = sl > tu > su	0.019	0.024	0.030
H10: (td, sd) > (tl, sl) > (tu, su)	0.006	0.014	0.023
H9: (td, sd) > (tl, sl, tu, su)	0.003	0.006	0.010

Note. t = temporal, s = social, d = downward, l = lateral, u = upward. * denotes Fraction set to 1, ** denotes Fraction set to 2, *** denotes Fraction set to 3. Posterior model probabilities (PMP) (a: excludes the unconstrained hypothesis) is based on equal prior model probability.

The findings suggest that self-efficacy changes vary by both comparison type and direction. For instance, H5 (td > sd > tl = sl = su > tu; Figure 5), which received the highest support, suggests that downward comparisons driven by mastery experiences information contribute more to positive changes in academic self-efficacy than do social modelling information. Conversely, for upward comparisons, mastery experiences result in greater negative changes. This hypothesis further suggests that participants tend to discount social modelling information in upward comparison scenarios. This idea is supported by the fact that the mean change in self-efficacy for upward social conditions mirrors those in lateral conditions, hereby supporting the claims made by H5.

There is also notable support for H8 (sd > td > tl = sl = su > tu; Figure 8), although less than that for H5. H8 posits that in the downward comparison condition, social modelling information prompts a greater positive change in academic self-efficacy than mastery experiences information. In alignment with H5, H8 suggests that in the upward condition, social modelling information’s effect equals that of both types of information in the lateral condition, while mastery experiences information in the upward condition leads to the most pronounced relative negative change.

While H5 and H8 garnered the most support, modest evidence also backs H6 and H1. H6 posits that in upward comparisons, social modelling exerts a greater influence than mastery experiences, contrasting with H5 and H8. This hypothesis otherwise aligns closely with them, except for its differing stance on social modelling's role in downward comparisons. H1, on the other hand, suggests that upward comparisons induce a more significant negative change in academic self-efficacy due to social modelling, also diverging from H5 and H8. These latter hypotheses contend that in upward comparisons, social modelling's impact is so minimal that it mimics the changes observed in lateral scenarios. Except for these distinctions, both H6 and H1 largely conform to H5.

Given the robust support for H5 and H8 and considering the lesser backing for H1 and H6, which largely appear as variations of the former, we might infer that the temporal downward condition could be equal to the social downward condition. H9 and H10, which received minimal support, can be safely disregarded. They suggest that neither mastery experiences nor social modelling significantly affect academic self-efficacy in any comparison conditions, a claim contradicted by the evidence.

These findings highlight the significant role of the relationship between comparison types and directions, as well as information sources, in shaping students' academic self-efficacy. However, it's essential to recognise that the best-supported hypothesis, H5, isn't without uncertainty. In fact, its error probabilities equal the sum of the PMPa values for all other hypotheses. Given this, continued investigation is needed to develop a comprehensive understanding of how academic self-efficacy is influenced by various types of comparisons.

5.4 Discussion

Our study investigated how different types of reference frames and comparison directions in LADs affect academic self-efficacy in higher education students. A key finding is that mastery experiences and social modelling information differentially affect academic self-efficacy depending on the direction of comparison.

Specifically, mastery experiences seem to have a more pronounced positive and negative impact on academic self-efficacy during downward and upward temporal comparisons, respectively, compared to the effects of social modelling information in these conditions. This supports Bandura’s (1997) position that mastery experiences have a greater influence on self-efficacy than social modelling. The evidence suggests that students rely more on actual performance as a gauge for determining self-efficacy beliefs than modelled performance (Schunk & Pajares, 2002). One reason for this could be that mastery experiences offer direct evidence of one’s capabilities (Bandura, 1997). In downward comparisons, students who see an improvement in their current versus past performance typically experience a boost in self-efficacy (Honicke et al., 2023; Wilson & Ross, 2000; Wilson & Shanahan, 2020). Supporting this observation, recent research has indicated that when a learner's current performance surpasses their prior achievements, their belief in their own abilities tends to increase (Honicke et al., 2023). This suggests that using mastery experiences in downward comparisons can be an effective strategy to enhance academic self-efficacy. Conversely, in upward comparisons where performance is declining, feedback from mastery experiences can erode self-efficacy.

There is also some evidence supporting the hypothesis that social modelling has a more overpowering effect in upward comparisons. Several interpretations of this result are possible. One explanation draws on research indicating that social comparisons significantly impact self-evaluations, particularly when individuals perceive superior performance by their peers (Van Yperen & Leander, 2014). This aligns with studies suggesting that individuals often underestimate negative temporal trends in their abilities (Zell & Alicke, 2010). In our study, this could account for partial support of the observed lesser negative impact of upward temporal comparisons on academic self-efficacy compared to upward social comparisons. Individuals may better manage declining personal performance over time than confronting evidence of peer superiority. This contrast between the effects of mastery experiences and social modelling highlights the complex relationship between comparison direction and type, and their collective impact on self-efficacy.

Our findings provide initial evidence on how different types and directions of comparisons may influence academic self-efficacy, particularly in the context of LADs. A recent study by van Leeuwen et al. (2023) highlighted a tension in participatory design between stakeholder input and educational theory, suggesting that teachers' perspectives on design might not always align with optimal design to support student learning outcomes. This consideration is crucial when designing LADs. The insights gathered related to the relative strength of mastery experiences information versus social modelling information could serve as a guide for educators and dashboard designers aiming to enhance learners' academic self-efficacy through more effective LADs.

Complementing our results is a recent study on the influence of LADs on self-regulated learning which found that a well-designed dashboard encouraging slight upward comparison can enhance extrinsic motivation and academic achievement (Fleur et al., 2023).While this stands in contrast to our findings, which suggest that upward social comparisons may either reduce or produce similar effects on academic self-efficacy as lateral comparisons. The findings highlight the complex nature of comparison dynamics in educational settings. While some dashboards can foster increased motivation and performance, they may also adversely affect academic self-efficacy. Such contrasting results underscore the need for LADs that are both carefully designed and empirically validated.

The utilisation of mastery experiences in a downward comparison setting appears to be a robust strategy for strengthening academic self-efficacy. In contrast, in the context of upward temporal comparisons, where students are faced with a declining performance trajectory, the feedback from mastery experiences can undermine their perceived capabilities, leading to a decrease in self-efficacy beliefs. Understanding how these different types and directions of comparison influence academic self-efficacy can inform dashboard designers' decisions. For instance, highlighting score delta may be beneficial in certain circumstances.

Nevertheless, it is important to consider that the level of self-efficacy influenced by these comparative settings is not always beneficial. While enhanced self-efficacy generally correlates with improved academic performance, an imbalance can have unintended consequences. While high levels of self-efficacy are associated with higher performance, among other beneficial factors for learning, it is crucial to note that inflated self-efficacy may not always be advantageous. For example, Talsma et al. (2019) showed that over-efficacious students, those whose self-efficacy beliefs exceed their actual performance, may experience negative impacts on academic self-regulation and subsequent performance. In contrast, students with underestimated self-efficacy performed better on similar tasks. This highlights the importance of aligning self-efficacy beliefs with actual capabilities.

While our study offers valuable insights into the effects of LADs on academic self-efficacy, it is essential to address its limitations. First, the observed shifts in self-efficacy were relatively minor, suggesting that while the type and direction of comparison do influence self-efficacy, their immediate practical impact may be limited. Nonetheless, these factors could have a more pronounced effect when combined with other elements of the learning environment, such as teaching methods, course material, and individual differences. It is therefore worth considering moving away from a one-size-fits-all approach to LAD design and towards personalised, adaptive learning strategies that consider individual differences (Park et al., 2022).

Additionally, our research has methodological limitations. The inherent nature of self-reported questionnaires raises concerns about the potential for bias in participants' responses, even though these tools are widely used and accepted in the research community. Further, our study had a somewhat limited sample size, which may restrict the generalisability of our findings. Future research should aim to include a larger and more diverse sample to strengthen the external validity. Also, while our study design allowed for examination of the impact of reference frames on academic self-efficacy, it failed to account for other potential influencing factors, such as the students' past academic performance. Moreover, the research was conducted over a relatively short period, which may limit our understanding of long-term changes in academic self-efficacy. Future research should consider extending the study period, as changes in self-efficacy may unfold gradually over time.

5.4.1 Conclusion

This study makes a significant contribution to the current dialogue surrounding the design of LADs and their potential to influence academic self-efficacy. Our research reveals that different types and directions of comparison integrated within LADs can likely influence learners' self-efficacy beliefs. Given these results, particular care must be exercised in the design and implementation of LADs. We found that comparisons in upward contexts can undermine academic self-efficacy. In light of this, LAD designers must proceed with caution, integrating psychological and educational theories along with individual user needs and experiences in their processes.

CHAPTER 6

General Discussion

Chapter 6 General Discussion

6.1 Introduction

Learning analytics systems, notably learning analytics dashboards (LADs), are increasingly influential in educational and training settings. Despite their potential, a distinct knowledge gap persists, specifically concerning the impact of LAD design on key variables associated with learning. To address this knowledge gap, this dissertation sets out to answer the following main research question: What is the impact of LAD design on learner preferences, interaction, and self-efficacy in educational and training settings?

At the heart of this research lies a comprehensive theoretical framework connecting technology and pedagogy. This framework not only serves as a conceptual lens for the four empirical studies but can also offer guidance for learning analytics researchers, enabling them to leverage existing educational and psychological theories in their learning analytics research efforts. Central to this is understanding how existing LAD design interacts with various educational and psychological theories including Self-Regulated Learning (SRL) Theory, Self-Efficacy Theory, Social Comparison Theory, Temporal Comparison Theory, Goal Origin Theory, and Achievement Goal Orientation Theory. Alongside the development of this theoretical framework, the goal of this dissertation is to present empirical evidence that will guide the development of LAD designs that align with established pedagogical and psychological principles, particularly as they apply to digital learning environments in workplace and higher educational settings.

This general discussion provides a summary and synthesis of the main findings from the four empirical studies conducted. Following that, it discusses the theoretical and practical implications of this research and then raises future issues related to both theory and practice.

6.2 Main Findings

Chapter 2 of this dissertation, which focuses on learner preferences for various LAD designs in a workplace setting is guided by two main research questions:

RQ1: In the context of an immersive learning environment, what are workplace learner preferences for learning analytics reference frames in LADs designed for before, during, and after task performance?

RQ2: In the context of an immersive learning environment, how are workplace learner preferences for learning analytics reference frames in LADs related to their perceived SRL skills?

The empirical results presented in this chapter show that there is a marked preference for the progress reference frame, which offers a form of temporal comparison which allows learners to evaluate their own performance over time. This preference for the progress reference frame was particularly notable during the 'before' and 'after' task performance stages. These stages align with the forethought and self-reflection phases of the SRL cycle (Panadero & Alonso-Tapia, 2014; Zimmerman, 2002). At these stages, the progress reference frame was favoured over all other types, including the external achievement reference frame (trainer assigned goal comparison), internal achievement reference frame (self-set goal comparison), and social reference frame (peer comparison) (Albert, 1977; Festinger, 1954; Seo et al., 2018; Suls et al., 2002; Zell & Alicke, 2010).

Notably, while previous research suggested that LAD designers most commonly use the social reference frame in their designs (Jivet et al., 2017), Chapter 2 reveals a discrepancy between prevalent design choices, which favour the social reference frame, and the actual least preferred status of this frame among learners. This counterintuitive finding that learner preferences are misaligned with the prevalent LAD design choices underscored the need for a deeper exploration.

Beyond these findings on reference frame preferences, Chapter 2 further explored the potential relationship between learner preferences and their perceived SRL skills for the 'before, during, and after task performance' stages. In analysing the before task performance stage, a significant association between learners' SRL skills and their preference for the Progress reference frame was uncovered. This underscores the potential impact of SRL skills on learners’ reference frame preferences in the initial stage of task performance.

The primary research question addressed in **Chapter 3** was:

‘When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in total change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?’

This question was pivotal in guiding the study to explore the effects of the progress and social reference frames on changes in occupational self-efficacy, and to understand the underlying mechanisms that influence these changes. Three competing hypotheses were proposed. Hypothesis 1 (H1) predicted greater occupational self-efficacy changes through the progress reference frame, citing its provision of additional mastery experience information. Hypothesis 2 (H2) posited that the social reference frame would elicit greater occupational self-efficacy change, arguing its dual provision of mastery and social modelling information would have a broader impact. Finally, Hypothesis 3 (H3) suggested that both frames would produce equal effects. There was very little evidence to suggest that the progress reference frame elicited greater occupational self-efficacy change than the social reference frame, which ruled out H1. While both H2 and H3 found similar levels of evidential support through the

Bayesian analysis suggesting that either the two reference frame elicited equal change, or the social reference frame elicited greater change than the progress reference frame.

The exploratory research question addressed in Chapter 3 was:

'When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in direction of change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?'

This question extended the primary line of inquiry to examine the directional change in occupational self-efficacy. Guided by the same theoretical underpinnings that shaped our primary research question and hypotheses, the exploratory hypotheses aimed to examine the directional effects of the progress and social reference frames on occupational self-efficacy.

Three exploratory hypotheses were proposed. Exploratory Hypothesis 1 (eH1) posited that the progress reference frame would elicit a more positive directional change in occupational self-efficacy than the social reference frame. Exploratory Hypothesis 2 (eH2) posited that the progress reference frame would elicit a less positive directional change in occupational self-efficacy compared to the social reference frame. Finally, Exploratory Hypothesis 3 (eH3) posited that both the progress and social reference frames would elicit an equal directional change in occupational self-efficacy. Both eH1 and eH3 had comparable levels of evidential support through the Bayesian analysis, suggesting that either the two reference frames elicited equal directional change or the progress reference frame elicited a more positive directional change. Meanwhile, there was limited evidence to suggest that the social reference frame elicited a more positive directional change in occupational self-efficacy than the progress reference frame, thereby ruling out eH2 as viable.

Chapter 4 shifts its focus from the occupational self-efficacy considerations of Chapter 3 to the practical implications of LAD designs on dashboard interaction among workplace learners. The differential analysis in this chapter, as it was in Chapter 3, is conducted between two distinct groups: one group was provided with an LAD with a Progress reference frame, while the other group received an LAD with a Social reference frame. By maintaining the digital learning environment and LADs from Chapter 3, this chapter extends the research scope to explore how these dashboard designs influence user interaction among workplace learners, thereby offering valuable insights that can assist designers in improving LAD designs.

This chapter is structured around an overarching research question—'How do reference frames influence LAD interaction?'—which is further dissected into targeted research questions and their accompanying hypotheses.

Three sets of research questions, accompanied by their respective hypotheses, guide the study. The first research question investigates whether the time spent reviewing LADs differs between groups exposed to the Progress and Social reference frames. Three hypotheses are proposed for this first research question. Hypothesis 1.1 (H1.1) suggests that individuals in the Progress LAD group will spend more time on LADs with a reference frame compared to their Social LAD counterparts. Hypothesis 1.2 (H1.2) counters this, positing less time will be spent by the Progress LAD group. Hypothesis 1.3 (H1.3) asserts that the time spent by both groups will be equal.

The second research question investigates if the total time spent reviewing the LAD offering detailed task feedback varies between the two groups. Three competing hypotheses are also offered here. Hypothesis 2.1 (H2.1) predicts the Progress LAD group will spend more time on detailed task feedback than the Social LAD group. Hypothesis 2.2 (H2.2) claims the opposite, anticipating less time spent by the Progress LAD group, while Hypothesis 2.3 (H2.3) holds that no between-group difference will exist.

Lastly, the third research question investigates the levels of engagement with LADs between the two groups. Hypothesis 3.1 (H3.1) anticipates that the Progress LAD group will engage more frequently with the LADs. Hypothesis 3.2 (H3.2) suggests that the Social LAD group will engage more frequently. Hypothesis 3.3 (H3.3) proposes equal levels of engagement between the two groups.

The findings indicate that workplace learners who interacted with an LAD featuring a progress reference frame spent less time reviewing their dashboards with a reference frame compared to those who interacted with an LAD with a social reference frame (H1.2). Conversely, learners exposed to the progress reference frame showed a greater inclination to spend more time on dashboards that provided detailed task feedback (H2.1).

Moreover, despite the Progress LAD group spending less time with dashboards with a reference frame, the analysis provided evidence to reject the hypothesis that learners in the Progress LAD group would engage less frequently with the dashboards than those in the Social LAD group (H3.2). The engagement levels for this group were either equivalent to (H3.3) or surpassed those of the group with a Social reference frame (H3.1).

Chapter 5 initiates a contextual shift, redirecting our attention from workplace learners to higher education students. While maintaining the research design elements covered in Chapter 3, this chapter introduces an additional variable: 'direction of comparison,' operationalised as downward, lateral, or upward relative to a predetermined ‘point of comparison’, which varies depending on the specific reference frame—either progress or social. The chapter tests ten hypotheses that incorporate both the type of reference frame (progress, social) and the operationalised direction of comparison (downward, lateral, upward) using Bayesian analysis.

The results of Chapter 5 indicate that the hypothesis, formulated as 'H5: td > sd > tl = sl = su > tu,' where 't' denotes 'temporal comparison,' 's' denotes 'social comparison,' 'd' denotes for 'downward,' 'l' denotes 'lateral,' and 'u' denotes 'upward.' This finding emphasises the dominant role of mastery experiences derived from temporal comparisons as a source of self-efficacy information in both downward and upward conditions. Notably, in the upward social condition, learners might perceive their relative underperformance compared to peers as not indicative of their own capability. Here, social modelling derived from social comparison emerges as an equally impactful source of self-efficacy information, equal to mastery experiences and social modelling information, in the lateral conditions.

In addition to H5, evidence was also found for a secondary hypothesis, denoted as 'H8: sd > td > tl = sl = su > tu.' This hypothesis suggests a similar sequence in decreasing self-efficacy change but initiates the sequence from downward social to downward temporal conditions. Unlike the most supported hypothesis, this secondary finding points to social modelling as the primary source of self-efficacy information in downward conditions. Conversely, in upward conditions, mastery experiences appear as the stronger source of self-efficacy information. Consistent with the most-supported hypothesis, the impact of relatively poorer performance in the upward social condition aligns with the effects of both mastery experiences and social modelling as sources of self-efficacy information in lateral conditions.

In summarising the findings across our empirical investigations, we observe nuanced relationships between LAD design on learner preferences, interaction, and self-efficacy. Chapter 2 reveals learners' preferences for progress reference frames and identifies a potential relationship between these preferences and self-assessed SRL skills. Subsequent chapters build upon this by revealing the complexities of the impact of the progress and social reference frames on self-assessed occupational self-efficacy (Chapter 3), as well as on the time and engagement levels learners invest in interacting with dashboards (Chapter 4). Importantly, Chapter 5 contributes to the discourse by underscoring the shifts in academic self-efficacy among higher education students, influenced by the progress and social reference frame. Taken together, these findings form a cohesive narrative that advances our understanding of how the design features of LADs have varied effects across the domains of learner preferences, interaction, and self-efficacy. As we turn our attention to the theoretical implications of these complex findings, we aim to engage in a synthesis that addresses the overarching research question: What is the impact of LAD design on learner preferences, interaction, and self-efficacy in educational and training settings?

6.3 Theoretical Implications

This dissertation contributes to the field of learning analytics by enhancing its theoretical foundations, a dimension often underdeveloped in technical domains in the educational sciences (Costa et al., 2019; Serdyukov, 2017). It adopts a multi-theoretical approach, drawing from SRL Theory (Zimmerman, 2002), Self-Efficacy Theory (Bandura, 1997), Social (Festinger, 1954; Gerber, 2020) and Temporal Comparison Theory (Albert, 1977; Zell & Alicke, 2010), Goal Origin Theory (Hollenbeck & Brief, 1987; Seo et al., 2018), and (Pintrich, 2000b; Pintrich et al., 2003). This broad theoretical base enables a detailed examination of how different LAD designs impact variables relevant to learning.

Building upon initial insights in the existing literature about progress, social, and achievement reference frames (Jivet et al., 2017; Wise, 2014), this dissertation strengthens the theoretical foundations of these constructs. While existing research has discussed the social reference frame in relation to social comparison theory (Gerber, 2020), it has offered limited insights into the progress reference frame's connection to temporal comparison theory (Albert, 1977; Wilson & Shanahan, 2020). By addressing this gap through a detailed exploration of the relationship between temporal comparison theory and the progress reference frame, a more nuanced understanding of how the progress and social reference frames may differentially affect learning is now available.

Continuing in this vein of strengthening the theoretical foundations of reference frame research and design, the dissertation also extends the scope of the achievement reference frame. It employs Goal Origin Theory (Hollenbeck & Brief, 1987; Seo et al., 2018) to examine both self-set and assigned goals, thereby broadening the existing focus, which has largely centred on Achievement Goal Orientation Theory (Pintrich et al., 2003). This complementary approach leads to the introduction of two novel categories: the internal achievement and external achievement reference frame.

Moreover, this work advances the field by connecting the design elements of LADs with the types of self-efficacy information learners use when forming self-efficacy beliefs (Bandura, 1997). This link enables learning analytics designers and researchers to create dashboards that are more effectively grounded in theory, particularly for influencing self-efficacy, a critical component in both academic and workplace performance (Rigotti et al., 2008; Zimmerman et al., 1992).

Another key contribution of this dissertation is the exploration of how the type of reference frame and the direction of comparison affect learners' self-efficacy beliefs (Suls et al., 2002). These finding offer a novel perspective for understanding and improving the design and effectiveness of LADs.

6.4 Practical Implications

The research presented in this dissertation offers valuable practical implications that are relevant to diverse stakeholders within the educational ecosystem. These implications serve as initial, flexible guidelines that necessitate ongoing review and adaptation, particularly given the ever-changing technological landscape and the dynamic nature of the field of learning analytics and educational technology more broadly.

6.4.1 For Learners

This dissertation highlights the potential benefits of LADs in supporting the SRL cycle. The practical implication for learners is clear: When they have access to LADs, they are encouraged to incorporate the dashboard's information into their SRL processes. For instance, examining one's progress over time through the dashboard can provide useful data for future planning and goal setting. Learners should understand that tracking these trends enables them to adjust their learning strategies and allocate resources more effectively.

Learners are advised to be cognisant of how each criterion type might shape their self-reflection phase processes. Normative criteria offer a social context but may also cultivate undue competitiveness. Past-performance criteria allow for tracking personal growth over time but may lack contextual anchors. Mastery criteria emphasise specific learning outcomes but might neglect the broader scope of learning. Therefore, by being aware of the impacts of these different criteria, learners can more effectively interpret and use the feedback they receive, enhancing their capacity for meaningful self-reflection.

Moreover, learners should be aware that dashboard designs can induce particular goal orientations. For example, some dashboards may steer learners toward performance-oriented goals, which might not always be appropriate. Being mindful of this, learners have the opportunity to intervene by cognitively reframing their goals to be, for example, more mastery-oriented when such an orientation better aligns with their learning objectives.

In summary, learners are encouraged to engage with the data presented on LADs, incorporating it into their SRL cycles for more targeted forethought, performance, and self-reflection phase processes. Such engagement not only maximises the utility of the technology but also fosters a more self-regulated approach to learning.

6.4.2 For Designers of Learning Analytics Dashboards

Designers engaged in the conceptualisation and deployment of LADs should carefully consider the multiple phases of the SRL cycle and their associated processes (Panadero & Alonso-Tapia, 2014). Each phase introduces unique requirements that ought to be accounted for in the timing and content of LADs. For example, learners with less advanced SRL skills might benefit more from having goals assigned to them rather than setting their own. Such a design choice can permit these learners to divert their cognitive efforts to other crucial aspects of this phase, like formulating strategies.

Additionally, designers need to account for learner preferences in LAD designs, as engagement with a dashboard may be influenced by how learners perceive it (Jivet et al., 2018; Nicol, 2020). A one-size-fits-all approach may not suffice, and designers should aim for a level of customisation or adaptability to meet diverse learner needs.

Further, designers must consider the importance of the selection of reference frames, which can have nuanced yet potentially substantial impacts on key learning-related variables such as self-efficacy and dashboard interaction (Bodily et al., 2018). Designers need to be conscious of how different types (i.e., temporal or social) and directions of comparisons—whether they be downward, lateral, or upward—might affect these determinants during the LAD design process (Collins, 1996; Suls et al., 2002).

In summary, designers are urged to carefully navigate the intricacies of the SRL cycle, learner preferences, and reference frame selection. By doing so, they can create LADs that not only offer valuable insights but also foster effective SRL experiences.

6.4.3 For Educational Researchers

This research has noteworthy practical implications for educational researchers, by highlighting the efficacy of comparative judgment in extracting user design preferences (Pollitt, 2012). Employing this technique, the present dissertation demonstrates that statistical analyses can quantify both the confidence level and the reliability of the design preferences elicited.

Moreover, the dissertation presents Bayesian informative hypothesis evaluation as an alternative to traditional null hypothesis testing using p values to determine significance (Hoijtink et al., 2019; Van Lissa et al., 2021). This approach is particularly advantageous when theoretical and empirical evidence give rise to competing plausible hypotheses.

Notably, the ease of implementing Bayesian methods has significantly increased with the advent of specialised software packages. While the ‘bain’ package in R provides the computational capabilities, it is JASP that offers a user-friendly interface akin to SPSS, making the adoption of Bayesian methods more accessible for a broader audience. Furthermore, the free and open source nature of these software packages removes the financial barriers often associated with commercial software like SPSS.

In sum, this dissertation introduces methodological innovations that enrich current common practices in educational research, providing both robust and accessible ways to conduct studies in this field.

6.4.4 For a Wider Educational Context

While these implications are neither exhaustive nor prescriptive, they are intended to inform decision-making across varied educational settings. The choice of design elements, particularly reference frames, can substantially shape the utility and effectiveness of LADs in meeting educational objectives (Verbert et al., 2020). As such, these guidelines should be regularly reviewed and adapted to reflect developments in educational technology and pedagogical approaches.

6.5 Limitations

The research presented in this dissertation offers valuable insights into the potential of LADs within the educational landscape, however, certain limitations merit attention and should inform future research.

A primary method of data collection used in this research was self-reported questionnaires. While these tools are valuable and widely accepted in the research community, they carry certain constraints (Tempelaar et al., 2020a). The potential for bias in participants' responses due to factors like social desirability is worth noting. To address these potential biases, future studies might consider integrating physiological response data or directly measuring the impact of LAD interventions on key learning related variables.

While certain individual differences, such as SRL skills, were accounted for in some chapters, others were overlooked. For instance, factors such as past academic performance, workplace experiences, or cognitive abilities could influence the observed outcomes (Tempelaar et al., 2020b). Introducing these factors in future research could provide a more comprehensive understanding in the effects of LAD designs.

Another limitation to note is the relatively brief duration over which the studies were conducted, which limits the ability to observe long-term changes in self-efficacy and evolving patterns of use related to LADs (White & Arzi, 2005). Future studies adopting a longitudinal approach could potentially improve the validity of the studies by highlighting trends that appear over prolonged durations.

Furthermore, the dissertation did not investigate factors related to the acceptance and use of technology and their influence on the outcome variables studied (Udeozor et al., 2021). An exploration integrating these dimensions in subsequent studies might offer a deeper understanding of the relationship between technology adoption and the influence of different LAD designs.

Lastly, the research did not employ measures to precisely gauge the visual engagement of learners with the LADs (Ha et al., 2015). Gaining a deeper understanding of learners' visual interactions with the dashboards would have added another valuable layer to the findings. There is potential in harnessing technology, such as eye-tracking, in future research to gain a better understanding on where learners' focus when interfacing with LADs (Clay et al., 2019).

6.6 Points on the Horizon

The field of educational technology is an ever-evolving landscape, continually influenced by methodological innovations and emerging technologies. Learning analytics, as a specialised domain within this broader field, is no exception. As the discipline matures, new avenues for research and development come into focus, necessitating ongoing theoretical research and the development of practical tools to address emerging questions and practical challenges. A range of research opportunities lie ahead, inviting focused study of, for example, the following topics.

6.6.1 Optimising Performance

Investigations into the influence of different reference frames on performance outcomes are important. Such research has the potential to offer empirical evidence that can inform both instructional strategies and the design of LADs, thus optimising educational performance. For instance, a study by (Garbers et al., 2023) did not find significant changes in learning outcomes when employing social comparison-based behavioural nudges through learning analytics in a graduate public health program. This result contributes valuable empirical data, aiding in the elimination of ineffective strategies and refining the direction of future instructional methods and LAD designs.

6.6.2 Unique Contextual Needs

Immersive learning environments, such as virtual reality, represent an underexplored yet increasingly relevant domain in learning analytics. A study by Christopoulos et al. (2020) advances this subfield by proposing a four-dimensional theoretical framework designed for virtual reality supported instruction, providing structural elements for the development of learning analytics systems. This framework highlights the needs of learners in immersive environments and offers actionable insights for customising LADs. Accordingly, future research should focus on building upon these contributions to more effectively tailor LADs to the unique needs of learners in immersive settings.

6.6.3 Longitudinal Studies

Longitudinal studies should be prioritised to gain a deeper understanding of long-term changes in self-efficacy and other key learning related variables. A relevant example can be found in a study that conducted a four-year longitudinal analysis using predictive learning analytics at a distance learning university (Herodotou et al., 2020). This extensive study offers valuable insights into the factors affecting large-scale adoption of learning analytics and serves as a model for the kind of longitudinal research that is needed to deeply understand key variables in the educational landscape.

6.6.4 AI's Role in Learning Analytics

The emerging opportunity to incorporate Artificial Intelligence (AI) into learning analytics systems marks a significant shift in the field. The role of AI tools, such as large language models and other generative technologies, in education, opens new avenues for making dashboards more adaptive and personalised. The integration of advanced analytics methods, such as predictive and prescriptive analytics, amplifies this potential further. A recent study elaborates on the concept of hybrid intelligence and introduces a detect-diagnose-act framework as the core function of AI in education (Molenaar, 2022). It suggests that a coordinated, interdisciplinary dialogue among various stakeholders, including educators, researchers, entrepreneurs, and policymakers, is essential for maximising the benefits of AI in learning analytics.

6.7 Conclusion

This dissertation has shed light on various aspects of learning analytics, particularly focusing on the design of LADs for digital learning environments in workplace and higher education settings. The research explored learner preferences and analysed the impact of different LAD designs on both occupational and academic self-assessed self-efficacy, as well as how these designs influenced dashboard use, providing insights that can be instrumental in shaping future practices.

By doing so this dissertation opens up new avenues for future research, underlining the significance of thoughtful LAD design that respects the complex interplay of various factors, including the impact on dashboard use, to ensure a positive impact on learning. The journey of understanding and harnessing learning analytics is ongoing, and this work is a valuable steppingstone on that path. At the core of this efforts lies the fundamental goal of enhancing the experiences and outcomes of learners, which must remain our primary focus.

English Summary

The current era is marked by a digital transformation that has been continuously altering the way individuals lead their lives, engage in work, and pursue education. The European Union's dedication to navigating this shift is evident through their substantial investment in the CHARMING project, which has provided resources to investigate the design of learning analytics for lifelong learning. This dissertation is a direct outcome of that project and investigates topics that sit at the intersection of technology and learning. Throughout, the importance of effective learning design alongside the innovative use of technology is emphasised. While technology offers immense potential to enhance learning experiences, it is crucial to design digital learning environments carefully. This can be accomplished in part by leaning on the wealth of established knowledge accumulated from decades of educational scientific inquiry. This dissertation aims to build on this established knowledge and contribute to it further by investigating identified gaps in that knowledge.

Chapter 1 serves as the foundation of the dissertation. It begins by underscoring the societal significance and relevance of learning analytics research, showing how it fits into the broader evolving digital landscape. Central to this chapter is the presentation of the problem statement, which focuses on the knowledge gap related to learning analytics design and the overarching research question:

How does learning analytics dashboard (LAD) design influence learner preferences, interaction, and self-efficacy in training and education?

The chapter introduces 'Learning Analytics' as a vital area for further exploration, highlighting the gaps in knowledge and emphasising the importance of learning analytics reference frames. To ground the research in recognised academic literature, it presents 'An Operational Framework for Learning Analytics Reference Frames' to systematically

deconstruct reference frames into discernible components. Supporting this, a comprehensive theoretical framework is offered, which provides a roadmap guiding the subsequent chapters. In essence, this chapter sets the stage for an in-depth examination of learning analytics dashboard design and their influence on key learning-related variables.

Chapter 2 investigates the growing prominence of LADs as tools for delivering feedback to workplace learners. As their adoption increases, two pivotal research questions emerge.

1: In the context of an immersive learning environment, what are workplace learner preferences for learning analytics reference frames in LADs designed for before, during and after task performance?

2: In the context of an immersive learning environment, how are workplace learner preferences for learning analytics reference frames in LADs related to their perceived SRL (self-regulated learning) skills?

The effectiveness of feedback is often contingent on how it is perceived by the learner, making the exploration of these questions crucial. The specific preferences of workplace learners regarding dashboard designs remain under-explored, especially when considering the phases of the SRL cycle: forethought, performance, and self-reflection. Furthermore, the relationship between these design preferences and the learner's SRL skills has not been thoroughly investigated. The chapter underscores the importance of investigating these preferences for varied dashboard designs, specifically those designed with the following reference frames: progress, social, internal achievement, and external achievement.

The empirical study documented in this chapter involved seventy participants from a chemical process apprenticeship program. Through an adaptive comparative judgement technique, the study sought to ascertain their preferences for four distinct dashboard designs, each corresponding to a specific SRL phase. Concurrently, a questionnaire was employed to gauge their perceived SRL skills. Rooted in both social and temporal comparison theory as well as goal-setting theory, these dashboard designs were subjected to statistical analyses to investigate the relationship between dashboard preferences and self-perceived SRL skills.

The results, processed using multinomial logistic regressions, revealed interesting patterns. The progress reference frame emerged as a favoured choice both before and after task performance, while the social reference frame was least preferred in these stages. Early evidence suggests a significant observation that learners with heightened self-assessed SRL skills displayed a pronounced inclination towards the progress reference frame before task execution, in contrast to those with no distinct preference. In alignment with existing research, the findings of this chapter emphasise the importance of exercising caution when integrating social comparison elements into dashboard designs intended for feedback.

Chapter 3 delves into the role of LADs within virtual reality simulation-based training environments in the Chemical industry. Central to this is how these dashboards use different reference frames—namely, the progress and social reference frame—to provide feedback. Specifically, the progress reference frame offers learners a comparison with their own past performance, while the social reference frame allows for comparison with the performance of their peers. The chapter further examines the potential of these dashboards to impact occupational self-efficacy through processes associated with the self-reflection phase of the SRL cycle. To support this, a theoretical framework is introduced to articulate the concept of learning analytics reference frames. Central to this investigation are two pivotal research questions:

1: When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in total change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?

2: When controlling for workplace self-reflection as a phase of the SRL cycle, are there between group differences in direction of change to occupational self-efficacy between pre-test and post-test for workplace learners who receive LADs with a progress reference frame compared to LADs with a social reference frame?

In this experimental study, 42 chemical operator employees, aged between 18 and 55 years and each with at least one year of experience, were engaged. Using a two-group design and Bayesian informative hypothesis evaluation, the study aimed to address the research questions, each supported by three competing hypotheses related to changes in occupational self-efficacy. The main findings indicated that dashboards with progress reference frames might not produce greater changes to self-efficacy than those with social reference frames, but both could elicit equal change. Additionally, dashboards with social reference frames might induce a more significant change to occupational self-efficacy than those with progress frames. Exploratory results suggested that dashboards with progress frames might prompt a more positive directional change than those with social frames or might produce an equal directional change.

These insights deepen the understanding of occupational self-efficacy beliefs within the Chemical industry and the potential for dashboards with different reference frames to differentially influencing skill development. The results can help guide the design of targeted interventions and training programs to bolster self-efficacy. From a practical standpoint, these findings emphasise the importance of thoughtfully choosing reference frames in LADs due to their potential impact on workplace learner occupational self-efficacy.

Chapter 4 builds upon the findings of Chapter 3 by shifting its focus to the analysis of log-file data, aiming to understand how chemical plant employees engage with LAD designs within a virtual reality simulation-based training environment. As in Chapter 3, these designs use the progress and social reference frames. The progress reference frame provides past performance data, while the social reference frame offers average peer group performance data. These designs, framed by achievement goal orientation theory and temporal and social comparison theories, offer different points of comparison. Central to this chapter's investigation are three research questions:

1: Are there between group differences in total time spent reviewing LADs with a reference frame?

2: Are there between group differences in total time spent reviewing detailed task feedback?

3: Are there between group differences in engagement with LADs?

To address these questions, the study assessed 42 participants' time spent on reviewing the dashboards, time dedicated to reviewing detailed task feedback, and the frequency of engagement with the LADs. This engagement was quantified by tracking the number of times specific features, such as detailed task feedback and assessment formulas, were selected. Findings from the study suggest that participants receiving a progress reference frame tend to spend less time on their main LAD containing the reference frame compared to those exposed to a social reference frame. Conversely, those with a progress reference frame appear to dedicate more time to reviewing detailed task feedback LADs. Such patterns suggest that progress reference frames might foster mastery goal orientation behaviours, while social reference frames seem to lean towards promoting performance goal orientation behaviours.

Chapter 5 moves from an investigation of workplace learners in the chemical industry to the setting of higher education and university students. In this context, as digital learning environments become increasingly prevalent, the importance of LADs as tools for delivering feedback has become more evident. These LADs often use progress and social frames to offer a context for students' self-evaluations. Central to this chapter is the research question:

How does learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students?

To address this, the study hypothesised that both the type of comparison—temporal or social—and the direction of comparison—downward, lateral, or upward—would have distinct impacts on learners' academic self-efficacy. A 2x3 mixed factorial switching replications design was implemented, considering dashboard design (either progress or social) and comparison direction (downward, lateral, or upward) as the main factors. Through Bayesian Informative Hypothesis Evaluation, changes in academic self-efficacy across six different conditions were compared.

Participation included 147 university students aged between 18-54. The results indicated that changes in academic self-efficacy were influenced by both the type and direction of comparison. Notably, temporal downward comparisons seemed to foster more positive changes in academic self-efficacy than social downward comparisons. Conversely, temporal upward comparisons appeared to lead to more pronounced negative changes. When exposed to upward comparison situations, participants seemed to give less weight to social comparison information in shaping their academic self-efficacy beliefs, though more research is needed to further validate these observations.

These findings emphasise the critical role of LAD design, shedding light on how diverse comparison types and directions can influence academic self-efficacy. By highlighting the potential impact of natures of comparison (i.e., type and direction) within LADs on academic self-efficacy, this chapter offers valuable insights into designing more effective LADs, enriching the conversation about the influence of digital learning environment designs on academic self-efficacy.

Chapter 6 begins with a discussion of the main research findings, revealing that learners exhibit clear preferences for specific reference frames, which may relate to self-perceived SRL skills. Furthermore, evidence suggests varied effects of these frames on learning-related variables in workplace and higher education settings, including impacts on academic and occupational self-efficacy as well as learner use of dashboards. Delving into theoretical implications, the chapter underscores the multi-theoretical approach of the dissertation, strengthening the theoretical base of progress, social, and achievement reference frames and introducing two new types: internal and external achievement reference frames.

The dissertation connects the design elements of LADs with the kind of self-efficacy information that learners use when shaping their self-efficacy beliefs, such as mastery experiences and social modelling information. This supports learning analytics designers and researchers in designing dashboards that are better informed by established educational and psychological theory. A key contribution is exploring the impact of reference frame types and comparison directions on learners' self-efficacy beliefs, paving the way for a fresh perspective in enhancing the understanding and optimisation of LAD design and effectiveness.

Chapter 6 then discusses the practical implications for various educational stakeholders, emphasising the importance of encouraging learners to actively engage with LAD data within their SRL cycles to improve learning processes. For LAD designers, it's crucial to consider learner preferences and the phases of the SRL cycle, ensuring the creation of dashboards that not only provide insights but also enhance SRL experiences. For educational researchers, the chapter highlights methodological advancements that refine existing research methodologies. These insights guide decision-making across varied contexts, underlining the need for continuous evaluation and recalibration of LADs in pace with technological and pedagogical advancements.

The chapter also addresses the limitations in the research presented. The primary data collection relied on self-reported questionnaires, which may introduce biases like social desirability. While controlling for individual differences like self-assessed SRL skills, other factors such as past academic performance or cognitive abilities were not included, which could influence outcomes. The research's design limits insights into long-term changes in self-efficacy and LAD use, suggesting that a longitudinal approach could reveal different patterns over time. Additionally, factors related to technology acceptance were not explored, leaving room for future studies to investigate the relationship between technology adoption and LAD design impact. The absence of precise measurements of learners' visual engagement with LADs was noted, indicating the potential use of eye-tracking technologies in future research.

In the "Points on the Horizon" section, Chapter 6 highlights the ever-evolving nature of educational technology and learning analytics, driven by methodological innovations and new technological arrivals. This calls for persistent theoretical research and the creation of practical tools informed by robust theory to address emerging challenges. Opportunities lie in examining the influence of different reference frames on performance outcomes, refining instructional techniques and LAD design for optimal educational results. Immersive learning environments, especially virtual reality, are identified as promising yet underexplored environments for learning analytics. The chapter also emphasises the potential of integrating Artificial Intelligence, particularly large language models, in learning analytics, opening a horizon of possibilities for more responsive and personalised dashboards.

Samenvatting

In het huidige tijdperk zorgt een digitale transformatie ervoor dat de manier waarop individuen hun leven leiden, aan het werk gaan en onderwijs volgen continue verandert. De betrokkenheid van de Europese Unie bij deze transformatie blijkt duidelijk uit hun substantiële investering in het CHARMING-project, dat onder andere het ontwerp van learning analytics voor een leven lang leren onderzoekt. Dit proefschrift is een direct resultaat van dat project en onderzoekt onderwerpen die zich op het snijvlak van technologie en leren bevinden. Overal wordt het belang van effectief instructieontwerp naast het innovatieve gebruik van technologie benadrukt. Hoewel technologie een enorm potentieel biedt om leerervaringen te verbeteren, is het van cruciaal belang om digitale leeromgevingen zorgvuldig te ontwerpen. Dit kan door gevestigde kennis die is opgebouwd gedurende decennia van onderwijswetenschappelijk onderzoek te gebruiken. Dit proefschrift heeft tot doel voort te bouwen op deze gevestigde kennis en er verder aan bij te dragen door te focusen op de hiaten.

Hoofdstuk 1 vormt de basis van het proefschrift. Het begint met de maatschappelijke betekenis en relevantie van learning analytics-onderzoek, en laat zien hoe dit past in het bredere transformerende digitale landschap. Centraal in dit hoofdstuk staan de probleemstelling, die zich richt op de kenniskloof van het ontwerp van learning analytics, en de overkoepelende onderzoeksvraag:

Hoe beïnvloedt het ontwerp van een learning analytics dashboard (LAD) de voorkeuren, het gebruik en de self-efficacy van lerenden in training en onderwijs?

Het hoofdstuk introduceert 'learning analytics' als de onderzoeksfocus van dit proefschrift, identificeert de hiaten in kennis en benadrukt het belang van learning analytics reference frames. Verder presenteert het 'Een operationeel raamwerk voor learning analytics reference frames' om reference frames systematisch te beschrijven gebaseerd op theorie. De theorie biedt een routekaart voor de volgende hoofdstukken. Kortom, dit hoofdstuk vormt de basis voor onderzoek naar het ontwerp van learning analytics dashboards en hun invloed op belangrijke leeruitkomsten.

Hoofdstuk 2 onderzoekt de LADs als instrument voor het geven van feedback aan werkplek lerenden met behulp van twee centrale onderzoeksvragen:

1: Wat zijn de voorkeuren van werkplek lerenden voor learning analytics reference frames in LAD's die zijn ontworpen voor vóór, tijdens en na de taakuitvoering in de context van een immersieve leeromgeving?

2: Hoe zijn de voorkeuren van werkplek lerenden voor learning analytics reference frames in LAD's gerelateerd aan hun zelfbeoordeelde SRL-vaardigheden (zelfregulerend leren) in de context van een immersieve leeromgeving?

De effectiviteit van feedback hangt vaak af van hoe de lerende deze ervaart, waardoor het onderzoeken van voorkeuren cruciaal is. De specifieke voorkeuren van werkplek lerenden met betrekking tot dashboardontwerpen blijven in de literatuur onderbelicht, vooral als we de fasen van de SRL-cyclus in ogenschouw nemen: vooruitdenken, presteren en zelfreflectie. Bovendien is de relatie tussen voorkeuren en zelfbeoordeelde SRL-vaardigheden niet grondig onderzocht. Hoofdstuk 2 onderzoekt dit voor verschillende dashboardontwerpen: de progress, social, internal achievement en external achievement reference frames. Deze ontwerpen zijn geworteld in social and temporal comparison theory en goal-setting theory.

Zeventig trainees in de chemische industrie deden mee aan dit onderzoek. Met een adaptive comparative judgement-techniek werden de deelnemers' voorkeuren vastgesteld voor de vier verschillende dashboardontwerpen, die elk overeenkomen met een specifieke

SRL-fase. Tegelijkertijd werd een vragenlijst gebruikt om hun SRL-vaardigheden te meten. De relatie tussen dashboardvoorkeuren en SRL-vaardigheden werd onderzocht.

De resultaten, geanalyseerd met multinomiale logistische regressies, brachten interessante patronen aan het licht. Het progress reference frame kwam naar voren als de favoriete keuze, zowel voor als na de taakuitvoering, terwijl het social reference frame in deze fasen het minst de voorkeur kreeg. Early evidence suggests a significant observation that learners with heightened self-assessed SRL skills displayed a pronounced inclination towards the progress reference frame before task execution, in contrast to those with no distinct preference. In lijn met bestaand onderzoek, laat dit hoofstuk zien dat voorzichtigheid geboden is bij het integreren van sociale vergelijkingselementen in LAD’s die bedoeld zijn voor feedback.

Hoofdstuk 3 gaat in op de rol van LAD's binnen op virtual reality-simulatie gebaseerde trainingsomgevingen in de chemische industrie. Centraal hierbij staat de manier waarop deze dashboards verschillende reference frames gebruiken, namelijk de progress en social reference frames, om feedback te geven. Concreet presenteert het progress reference frame een vergelijking tussen vorige prestaties en huidige prestaties, terwijl het social reference frame een vergelijking presenteert tussen eigen prestaties en (gemiddelde) peer prestaties. Het hoofdstuk onderzoekt verder hoe deze dashboards de self-efficacy op de werkplek beïnvloeden via zelfreflectie processen in de SRL-cyclus. Het introduceert een theoretisch raamwerk om het concept van learning analytics reference frames te beschrijven. Centraal in dit onderzoek staan twee onderzoeksvragen:

1 : Als er wordt gecontroleerd voor werkplek zelfreflectie als fase van de SRL-cyclus, zijn er dan verschillen in de totale verandering self-efficacy op de werkplek tussen pre-test en post-test voor werkplek lerenden die LAD's ontvangen met een progress reference frame vergeleken met werkplek lerenden die LAD’s ontvangen met een social reference frame?

2: Als er wordt gecontroleerd voor werkplek zelfreflectie als fase van de SRL-cyclus, zijn er dan verschillen in de veranderingsrichting van self-efficacy op de werkplek tussen pre-test en post-test voor werkplek lerenden die LAD's ontvangen met een progress reference frame vergeleken met werkplek lerenden die LAD's ontvangen met een social reference frame?

Bij dit experimentele onderzoek waren 42 operators van chemische installaties betrokken. Ze waren tussen de 18 en 55 jaar oud en hadden elk minimaal één jaar ervaring. Een onderzoeksontwerp met twee groepen en een Bayesiaanse informatieve hypothese-evaluatie werden gebruikt om de onderzoeksvragen te beantwoorden. De onderzoeksvragen werden elk vergezeld door drie concurrerende hypothesen die verband hielden met veranderingen in de self-efficacy op de werkplek. De belangrijkste bevindingen waren dat dashboards met progress reference frames waarschijnlijk geen grotere veranderingen in de self-efficacy teweegbrengen dan dashboards met social reference frames, maar dat beiden vergelijkbare veranderingen teweegbrengen. Bovendien zouden dashboards met social reference frames een grotere verandering in de self-efficacy op de werkplek teweeg kunnen brengen dan dashboards met progress reference frames. Verkennende analyses suggereerden dat dashboards met progress reference frames een positievere verandering teweeg zouden kunnen brengen dan dashboards met social reference frames, of dat ze een gelijke verandering zouden kunnen veroorzaken.

Deze inzichten verdiepen het begrip van self-efficacy op de werkplek binnen de chemische industrie en laten zien dat dashboards met verschillende reference frames de vaardighedenontwikkeling in potentie op verschillende manieren kunnen beïnvloeden. De resultaten kunnen helpen bij het ontwerpen van gerichte interventies en trainingsprogramma's om de self-efficacy op de werkplek te versterken. Vanuit praktisch oogpunt is het belangrijk om reference frames in LADs zorgvuldig te kiezen vanwege hun potentiële impact op de self-efficacy van werkplek lerenden.

Hoofdstuk 4 bouwt voort op de bevindingen uit Hoofdstuk 3 en verlegt de focus naar de analyse van logbestandsgegevens, met als doel te begrijpen hoe operators van chemische installaties de verschillende LAD-ontwerpen binnen een op virtual reality-simulatie gebaseerde trainingsomgeving gebruiken. Net als in Hoofdstuk 3, vergelijkt Hoofdstuk 4 LAD's met een progress of social reference frame ingekaderd door de goal theory en de temporal and social comparison theory. Het progress reference frame biedt een vergelijk tussen prestatiegegevens uit het verleden en de huidige prestaties, terwijl het social reference frame een vergelijk biedt tussen huidige prestaties en de gemiddelde prestaties van peers. Centraal in het onderzoek van dit hoofdstuk staan drie onderzoeksvragen:

1: Zijn er verschillen tussen groepen in de totale tijd besteed aan het beoordelen van LAD's met een reference frame?

2: Zijn er verschillen tussen groepen in de totale tijd die wordt besteed aan het beoordelen van gedetailleerde taakfeedback?

3: Zijn er verschillen tussen groepen in het gebruik van LAD's?

Om deze vragen te beantwoorden, keek het onderzoek naar de tijd die 42 deelnemers besteedden aan het beoordelen van de dashboards, de tijd die werd besteed aan het beoordelen van gedetailleerde taakfeedback en de frequentie van de interactie met de LAD's. Het gebruik werd verder gekwantificeerd door het aantal keren bij te houden dat specifieke kenmerken, zoals gedetailleerde taakfeedback en beoordelingsformules, werden geselecteerd. Uit het onderzoek blijkt dat deelnemers die een progress reference frame ontvangen, doorgaans minder tijd besteden aan hun LAD met daarin het reference frame, vergeleken met degenen die zijn blootgesteld aan een social reference frame. Omgekeerd lijken degenen met een progress reference frame meer tijd te besteden aan het bekijken van gedetailleerde taakfeedback. Dergelijke patronen suggereren dat progress reference frames het mastery goal orientation gedrag zouden kunnen bevorderen , terwijl social reference frames het performance goal orientation gedrag zouden kunnen bevorderen.

Hoofdstuk 5 gaat van een onderzoek onder werkplek lerenden in de chemische industrie naar een onder studenten uit het hoger onderwijs uitgevoerd onderzoek. In deze context, nu digitale leeromgevingen steeds gangbaarder worden, is het belang van LAD's als instrumenten voor het leveren van feedback duidelijker geworden. Hier maken LAD's vaak gebruik van vooruitgang en sociale kaders om een context te bieden voor de zelfevaluatie van studenten. Centraal in dit hoofdstuk staat de onderzoeksvraag:

Hoe beïnvloeden het type learning analytics reference frame en de vergelijkingsrichting binnen het reference frame de academische self-efficacy van studenten in het hoger onderwijs?

Om dit te onderzoeken werd in de studie verondersteld dat zowel het type vergelijking – temporeel of sociaal – als de richting van de vergelijking – neerwaarts, lateraal of opwaarts – duidelijke gevolgen zouden hebben voor de academische self-efficacy van lerenden. Er werd een 2x3 mixed factorial switching replications onderzoeksontwerp geïmplementeerd, waarbij het dashboardontwerp (progress of social) en de vergelijkingsrichting (neerwaarts, lateraal of opwaarts) als de belangrijkste factoren werden beschouwd. Via Bayesiaanse informatieve hypothese-evaluatie werden veranderingen in academische self-efficacy onder zes verschillende omstandigheden vergeleken.

147 universiteitsstudenten tussen 18 en 54 jaar namen deel aan het onderzoek. De resultaten gaven aan dat veranderingen in academische self-efficacy werden beïnvloed door zowel het type vergelijking als de richting van de vergelijking. Met name leken temporele,

neerwaartse vergelijkingen meer positieve veranderingen in de academische self-efficacy te bewerkstelligen dan sociale, neerwaartse vergelijkingen. Omgekeerd leken temporele, opwaartse vergelijkingen tot meer uitgesproken negatieve veranderingen te leiden. Wanneer ze werden blootgesteld aan opwaartse vergelijkingssituaties, leken de lerenden minder waarde toe te kennen aan sociale vergelijkingen bij het vormen van hun academische self-efficacy opvattingen. Er is meer onderzoek nodig om deze observaties verder te valideren.

Deze bevindingen benadrukken de cruciale rol van LAD-ontwerp en laten zien hoe diverse vergelijkingstypen en -richtingen de academische self-efficacy kunnen beïnvloeden. Dit hoofdstuk biedt waardevolle inzichten in het ontwerpen van effectievere LADs door de potentiële impact van de aard van vergelijking (dat wil zeggen, type en richting) binnen LADs op academische self-efficacy te benadrukken. Hierdoor wordt het gesprek over de invloed van digitale leeromgevingontwerpen op academische self-efficacy verrijkt.

Hoofdstuk 6 begint met een bespreking van de belangrijkste onderzoeksresultaten. Lerenden blijken duidelijke voorkeuren te hebben voor specifieke reference frames, die verband kunnen houden met zelfbeoordeelde SRL-vaardigheden. Bovendien hebben de verschillende frames verschillende effecten op leeruitkomsten (dat wil zeggen, de professionele en academische self-efficacy en het gebruik van dashboards door lerenden) op de werkvloer en in het hoger onderwijs. In de theoretische implicaties onderstreept het hoofdstuk de multitheoretische benadering van het proefschrift, waarbij de theoretische basis van progress, social en achievement reference frames wordt versterkt en twee nieuwe typen worden geïntroduceerd: internal en external achievement reference frames.

Het proefschrift verbindt de ontwerpelementen van LADs met het soort informatie over self-efficacy dat lerenden gebruiken bij het vormen van hun ideeën over self-efficacy, zoals mastery ervaringen en social modelling informatie. Dit ondersteunt learning analytics-ontwerpers en onderzoekers bij het ontwerpen van dashboards die geworteld zijn in onderwijs- en psychologische theorieën. Een belangrijke bijdrage is de focus op reference frame typen en vergelijkingsrichtingen in relatie tot self-efficacy, waardoor de weg wordt vrijgemaakt voor een nieuw perspectief bij het verbeteren van het begrip, de optimalisatie van het ontwerp en de effectiviteit van LADs.

Hoofdstuk 6 bespreekt vervolgens de praktische implicaties voor verschillende belanghebbenden in het onderwijs. Het is daarbij belangrijk dat lerenden worden aangemoedigd om actief bezig te zijn met LAD-gegevens binnen hun SRL-cycli om leerprocessen te verbeteren. Voor LAD-ontwerpers is het van cruciaal belang om rekening te houden met de voorkeuren van lerenden en de fasen van de SRL-cyclus, waardoor dashboards kunnen worden gecreëerd die niet alleen inzichten bieden, maar ook SRL-ervaringen verbeteren. Voor onderwijsonderzoekers belicht dit hoofdstuk methodologische ontwikkelingen die bestaande onderzoeksmethodologieën verfijnen. Deze inzichten begeleiden de besluitvorming in verschillende contexten, en benadrukken de noodzaak van voortdurende evaluatie en herijking van LAD’s parallel aan de technologische en onderwijswetenschappelijke vooruitgang.

Het hoofdstuk gaat ook in op de beperkingen van het gepresenteerde onderzoek. De primaire gegevensverzameling was gebaseerd op zelfgerapporteerde vragenlijsten, die sociale wenselijkheid kunnen introduceren. Hoewel er werd gecontroleerd voor individuele verschillen, zoals zelfbeoordeelde SRL-vaardigheden, werden andere factoren, zoals academische prestaties uit het verleden of cognitieve vaardigheden die de uitkomsten zouden kunnen beïnvloeden, niet meegenomen. Het ontwerp van het onderzoek beperkt het inzicht in langetermijnveranderingen in self-efficacy en LAD-gebruik, wat suggereert dat een longitudinale benadering in de loop van de tijd andere patronen zou kunnen opleveren. Bovendien werden factoren die verband houden met technologie-acceptatie niet onderzocht, waardoor er ruimte overblijft in toekomstige studies om de relatie tussen technologie-adoptie en de impact van LAD-ontwerp te onderzoeken. Er werd opgemerkt dat er geen nauwkeurige metingen zijn van de visuele betrokkenheid van lerenden bij LAD's, wat wijst op het potentiële gebruik van eye-trackingtechnologieën in toekomstig onderzoek.

In het gedeelte 'Points on the horizon' belicht Hoofdstuk 6 de steeds veranderende aard van onderwijstechnologie en learning analytics, aangedreven door methodologische innovaties en nieuwe technologische ontwikkelingen. Dit vraagt om aanhoudend onderzoek en het creëren van praktische instrumenten, gebaseerd op robuuste theorieën, om opkomende uitdagingen aan te pakken. Mogelijkheden liggen in het onderzoeken van de invloed van verschillende reference frames op prestatieresultaten, het verfijnen van instructietechnieken en LAD-ontwerp voor optimale onderwijsresultaten. Immersieve leeromgevingen, vooral virtual reality, worden geïdentificeerd als veelbelovende maar nog onderontdekte omgevingen voor learning analytics. Het hoofdstuk benadrukt ook het potentieel van het integreren van kunstmatige intelligentie, met name grote taalmodellen, in learning analytics, waardoor een horizon van mogelijkheden wordt geopend voor responsievere en gepersonaliseerde dashboards.

References

Adams, K. (2004). Modelling success: Enhancing international postgraduate research students' self-efficacy for research seminar presentations. *Higher Education Research and Development*, *23*(2), 115–130. https://doi.org/10.1080/0729436042000206618

Agresti, A. (2002). *Categorical Data Analysis*. Wiley. https://doi.org/10.1002/0471249688

Albert, S. (1977). Temporal comparison theory. *Psychological Review*, *84*(6), 485–503. https://doi.org/10.1037/0033-295X.84.6.485

Ariely, D., & Wertenbroch, K. (2002). Procrastination, Deadlines, and Performance: Self-Control by Precommitment. *Psychological Science*, *13*(3), 219–224. https://doi.org/10.1111/1467-9280.00441

Audia, P. G., Locke, E. A., & Smith, K. G. (2000). The Paradox of Success: An Archival and a Laboratory Study of Strategic Persistence Following Radical Environmental Change. *Academy of Management Journal*, *43*(5), 837–853. https://doi.org/10.5465/1556413

Bandura, A. (1977). Self-efficacy: Toward a unifying theory of behavioral change. *Advances in Behaviour Research and Therapy*, *1*(4), 139–161. https://doi.org/10.1016/0146-6402(78)90002-4

Bandura, A. (1997). Self-efficacy: The exercise of control. In *Self-efficacy: The exercise of control.* W H Freeman/Times Books/ Henry Holt & Co.

Bandura, A. (2010). The Corsini Encyclopedia of Psychology. In I. B. Weiner & W. E. Craighead (Eds.), *The Corsini Encyclopedia of Psychology*. Wiley. https://doi.org/10.1002/9780470479216

Bandura, A. (2012). On the functional properties of perceived self-efficacy revisited. *Journal of Management*, *38*(1), 9–44. https://doi.org/10.1177/0149206311410606

Bandura, A., & Locke, E. A. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology*, *88*(1), 87–99. https://doi.org/10.1037/0021-9010.88.1.87

Barreiros, C., Leitner, P., Ebner, M., Veas, E., & Lindstaedt, S. (2023). *Students in Focus – Moving Towards Human-Centred Learning Analytics* (pp. 77–94). https://doi.org/10.1007/978-3-031-27646-0_5

Bartholomew, S. R., Yoshikawa, E., & Connolly, P. (2018). Exploring the Potential for Identifying Student Competencies in Design Education through Adaptive Comparative Judgment. *36th International Pupils' Attitudes Towards Technology Conference*, 187–194.

Bartimote-Aufflick, K., Bridgeman, A., Walker, R., Sharma, M., & Smith, L. (2016). The study, evaluation, and improvement of university student self-efficacy. *Studies in Higher Education*, *41*(11), 1918–1942. https://doi.org/10.1080/03075079.2014.999319

Beck, D., Morgado, L., & O'Shea, P. (2023). Educational Practices and Strategies with Immersive Learning Environments: Mapping of Reviews for using the Metaverse. *IEEE Transactions on Learning Technologies*, 1–23. https://doi.org/10.1109/TLT.2023.3243946

Beheshitha, S. S., Hatala, M., Gaševic, D., & Joksimovic, S. (2016). The role of achievement goal orientations when studying effect of learning analytics visualizations. *ACM International Conference Proceeding Series*, *25-29-Apri*, 54–63. https://doi.org/10.1145/2883851.2883904

Bennett, L., & Folley, S. (2020). Four design principles for learner dashboards that support student agency and empowerment. *Journal of Applied Research in Higher Education*, *12*(1), 15–26. https://doi.org/10.1108/JARHE-11-2018-0251

Bodily, R., Kay, J., Aleven, V., Jivet, I., Davis, D., Xhakaj, F., & Verbert, K. (2018). *Open learner models and learning analytics dashboards: A systematic review*. 41–50. https://doi.org/10.1145/3170358.3170409

Brusilovsky, P., Somyurek, S., Guerra, J., Hosseini, R., Zadorozhny, V., & Durlach, P. J. (2016). Open Social Student Modeling for Personalized Learning. *IEEE Transactions on Emerging Topics in Computing*, *4*(3), 450–461. https://doi.org/10.1109/TETC.2015.2501243

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

Carter, W. R., Nesbit, P. L., Badham, R. J., Parker, S. K., & Sung, L. K. (2018). The effects of employee engagement and self-efficacy on job performance: a longitudinal field study. *International Journal of Human Resource Management*, *29*(17), 2483–2502. https://doi.org/10.1080/09585192.2016.1244096

Castro, R. (2019). Blended learning in higher education: Trends and capabilities. *Education and Information Technologies*, *24*(4), 2523–2546. https://doi.org/10.1007/s10639-019-09886-3

Chen, F., Curran, P. J., Bollen, K. A., Kirby, J., & Paxton, P. (2008). An empirical evaluation of the use of fixed cutoff points in RMSEA test statistic in structural equation models. *Sociological Methods and Research*, *36*(4), 462–494. https://doi.org/10.1177/0049124108314720

Cherian, J., & Jacob, J. (2013). Impact of Self Efficacy on Motivation and Performance of Employees. *International Journal of Business and Management*, *8*(14), 80–88. https://doi.org/10.5539/ijbm.v8n14p80

Christopoulos, A., Pellas, N., & Laakso, M.-J. (2020). A Learning Analytics Theoretical Framework for STEM Education Virtual Reality Applications. *Education Sciences*, *10*(11), 317. https://doi.org/10.3390/educsci10110317

Clark, R. E. (1983). Reconsidering Research on Learning from Media. *Review of Educational Research*, *53*(4), 445–459. https://doi.org/10.3102/00346543053004445

Clay, V., König, P., & König, S. U. (2019). Eye tracking in virtual reality. *Journal of Eye Movement Research*, *12*(1). https://doi.org/10.16910/jemr.12.1.3

Cleary, T. J. (2009). Monitoring trends and accuracy of self-efficacy beliefs during interventions: Advantages and potential applications to school-based settings. *Psychology in the Schools*, *46*(2), 154–171. https://doi.org/10.1002/pits.20360

Collins, R. L. (1996). For better or worse: The impact of upward social comparison on self-evaluations. *Psychological Bulletin*, *119*(1), 51–69. https://doi.org/10.1037/0033-2909.119.1.51

Corrin, L., & De Barba, P. (2014). *Exploring students' interpretation of feedback delivered through learning analytics dashboards*. 629–633. https://www.scopus.com/inward/record.uri?eid=2-s2.0-84955260027&partnerID=40&md5=73dcc7021ff746947e6d32163c1c3df5

Corrin, L., & de Barba, P. (2015). How do students interpret feedback delivered via dashboards? *Proceedings of the Fifth International Conference on Learning Analytics And Knowledge*, *16-20-Marc*, 430–431. https://doi.org/10.1145/2723576.2723662

Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*(1), 98–104. https://doi.org/10.1037/0021-9010.78.1.98

Costa, C., Hammond, M., & Younie, S. (2019). Theorising technology in education: an introduction. *Technology, Pedagogy and Education*, *28*(4), 395–399. https://doi.org/10.1080/1475939X.2019.1660089

Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the Motivated Strategies for Learning Questionnaire. *Learning and Individual Differences*, *21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002

Davis, D., Jivet, I., Kizilcec, R. F., Chen, G., Hauff, C., & Houben, G.-J. (2017). Follow the successful crowd. *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, 454–463. https://doi.org/10.1145/3027385.3027411

Duncan, T. G., & McKeachie, W. J. (2005). The Making of the Motivated Strategies for Learning Questionnaire. *Educational Psychologist*, *40*(2), 117–128. https://doi.org/10.1207/s15326985ep4002_6

El-Beheiry, M., McCreery, G., & Schlachta, C. M. (2017). A serious game skills competition increases voluntary usage and proficiency of a virtual reality laparoscopic simulator during first-year surgical residents' simulation curriculum. *Surgical Endoscopy*, *31*(4), 1643–1650. https://doi.org/10.1007/s00464-016-5152-y

Elias, S. M., & MacDonald, S. (2007). Using past performance, proxy efficacy, and academic self-efficacy to predict college performance. *Journal of Applied Social Psychology*, *37*(11), 2518–2531. https://doi.org/10.1111/j.1559-1816.2007.00268.x

Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 achievement goal framework. *Journal of Personality and Social Psychology*, *80*(3), 501–519. https://doi.org/10.1037/0022-3514.80.3.501

Epton, T., Currie, S., & Armitage, C. J. (2017). Unique Effects of Setting Goals on Behavior Change: Systematic Review and Meta-Analysis. *Journal of Consulting and Clinical Psychology*, *85*(12), 1182–1198. https://doi.org/10.1037/ccp0000260

Erdoğan, A., Kesici, Ş., & Sahin, I. (2011). *Prediction of High School Students' Mathematics Anxiety by Their Achievement Motivation and Social Comparison*. *10*.

Fairall, D. G., & Rodgers, W. M. (1997). The Effects of Goal-Setting Method on Goal Attributes in Athletes: A Field Experiment. *Journal of Sport and Exercise Psychology*, *19*(1), 1–16. https://doi.org/10.1123/jsep.19.1.1

Farahmand, A., Dewan, M. A. A., & Lin, F. (2020). *Student-Facing Educational Dashboard Design for Online Learners*. 345–349. https://doi.org/10.1109/DASC-PICom-CBDCom-CyberSciTech49142.2020.00067

Fejoh, J., Boyede, M. A., & Momoh, A. M. (2018). Self- Efficacy and Self-Knowledge as Predictors of Workers' Job Performance in Food and Beverage Industries in the South-West, Nigeria. *Kampala International University KIU Journal of Humanities*, *3*(1), 159–166.

Ferla, J., Valcke, M., & Cai, Y. (2009). Academic self-efficacy and academic self-concept: Reconsidering structural relationships. *Learning and Individual Differences*, *19*(4), 499–505. https://doi.org/10.1016/j.lindif.2009.05.004

Festinger, L. (1954). A Theory of Social Comparison Processes. *Human Relations*, *7*(2), 117–140. https://doi.org/10.1177/001872675400700202

Fleur, D. S., van den Bos, W., & Bredeweg, B. (2023). Social comparison in learning analytics dashboard supporting motivation and academic achievement. *Computers and Education Open*, *4*, 100130. https://doi.org/10.1016/j.caeo.2023.100130

Fontana, R. P., Milligan, C., Littlejohn, A., & Margaryan, A. (2015). Measuring self-regulated learning in the workplace. *International Journal of Training and Development*, *19*(1), 32–52. https://doi.org/10.1111/ijtd.12046

Fukubayashi, N., & Fuji, K. (2021). Social Comparison on Social Media Increases Career Frustration: A Focus on the Mitigating Effect of Companionship. *Frontiers in Psychology*, *12*. https://doi.org/10.3389/fpsyg.2021.720960

Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Fracaro, S. G., & Kester, L. (2022). *Comparison with Self vs Comparison with Others: The Influence of Learning Analytics Dashboard Design on Learner Dashboard Use* (pp. 11–21). https://doi.org/10.1007/978-3-031-22124-8_2

Garbers, S., Crinklaw, A. D., Brown, A. S., & Russell, R. (2023). Increasing student engagement with course content in graduate public health education: A pilot randomized trial of behavioral nudges. *Education and Information Technologies*. https://doi.org/10.1007/s10639-023-11709-5

Garcia Fracaro, S., Bernaerts, K., Glassey, J., & Wilk, M. (2022). Effectiveness of Immersive Learning Technologies for Training of Operators in the Chemical Industry. In *Effectiviteit van immersieve leertechnologieën voor de opleiding van operatoren in de chemische industrie*.

Garcia Fracaro, S., Chan, P., Gallagher, T., Tehreem, Y., Toyoda, R., Bernaerts, K., Glassey, J., Pfeiffer, T., Slof, B., Wachsmuth, S., & Wilk, M. (2021). Towards design guidelines for virtual reality training for the chemical industry. *Education for Chemical Engineers*, *36*, 12–23. https://doi.org/10.1016/j.ece.2021.01.014

Garcia Fracaro, S., Tehreem, Y., Gallagher, T., Toyoda, R., Glassey, J., Bernaerts, K., & Wilk, M. (2022). Procedure and emergency virtual reality training in the chemical industry: Study of effectiveness. *Chemie Ingenieur Technik*, *94*(9), 1375–1376. https://doi.org/10.1002/cite.202255202

Gasevic, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends*, *59*(1), 64–71. https://doi.org/10.1007/s11528-014-0822-x

Gasevic, D., Jovanovic, J., & Siadaty, M. (2013). *The Social Semantic Web and Workplace Learning* (pp. 132–143).

Gedrimiene, E., Silvola, A., Pursiainen, J., Rusanen, J., & Muukkonen, H. (2020). Learning Analytics in Education: Literature Review and Case Examples From Vocational Education. *Scandinavian Journal of Educational Research*, *64*(7), 1105–1119. https://doi.org/10.1080/00313831.2019.1649718

Gerber, J. P. (2019). Social Comparison Theory. *Encyclopedia of Personality and Individual Differences*, *2015*, 5004–5010. https://doi.org/10.1007/978-3-319-24612-3_1182

Gerber, J. P. (2020). Social comparison theory. *Encyclopedia of Personality and Individual Differences*, 5004–5011.

Gerber, J. P., Wheeler, L., & Suls, J. (2018). A social comparison theory meta-analysis 60+ years on. *Psychological Bulletin*, *144*(2), 177–197. https://doi.org/10.1037/bul0000127

Greller, W., & Drachsler, H. (2012). *Translating Learning into Numbers: A Generic Framework for Learning Analytics Multimodal Learning Analytics for Collaborative Learning Understanding and Support View project*. *June 2014*. http://groups.google.com/group/learninganalytics

Guerra, J., Hosseini, R., Somyurek, S., & Brusilovsky, P. (2016). An Intelligent Interface for Learning Content. *Proceedings of the 21st International Conference on Intelligent User Interfaces*, *07-10-March-2016*, 152–163. https://doi.org/10.1145/2856767.2856784

Guyer, J. J., & Vaughan-Johnston, T. I. (2018). Social Comparisons (Upward and Downward). In V. Zeigler-Hill & T. K. Shackelford (Eds.), *Encyclopedia of Personality and Individual Differences* (pp. 1–5). Springer International Publishing. https://doi.org/10.1007/978-3-319-28099-8_1912-1

Ha, K., Jo, I.-H., Lim, S., & Park, Y. (2015). Tracking Students' Eye-Movements on Visual Dashboard Presenting Their Online Learning Behavior Patterns. In *Lecture Notes in Educational Technology* (Issue 9783662441879, pp. 371–376). Springer International Publishing. https://doi.org/10.1007/978-3-662-44188-6_51

Herodotou, C., Rienties, B., Hlosta, M., Boroowa, A., Mangafa, C., & Zdrahal, Z. (2020). The scalable implementation of predictive learning analytics at a distance learning university: Insights from a longitudinal case study. *The Internet and Higher Education*, *45*, 100725. https://doi.org/10.1016/j.iheduc.2020.100725

Hoijtink, H., Mulder, J., van Lissa, C., & Gu, X. (2019). A tutorial on testing hypotheses using the Bayes factor. *Psychological Methods*, *24*(5), 539–556. https://doi.org/10.1037/met0000201

Hollenbeck, J. R., & Brief, A. P. (1987). The effects of individual differences and goal origin on goal setting and performance. *Organizational Behavior and Human Decision Processes*, *40*(3), 392–414. https://doi.org/10.1016/0749-5978(87)90023-9

Honicke, T., Broadbent, J., & Fuller-Tyszkiewicz, M. (2023). The self-efficacy and academic performance reciprocal relationship: the influence of task difficulty and baseline achievement on learner trajectory. *Higher Education Research & Development*, 1–18. https://doi.org/10.1080/07294360.2023.2197194

Hough, J., Levan, D., Steele, M., Kelly, K., & Dalton, M. (2019). Simulation-based education improves student self-efficacy in physiotherapy assessment and management of paediatric patients. *BMC Medical Education*, *19*(1). https://doi.org/10.1186/s12909-019-1894-2

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Ifenthaler, D. (2018). How We Learn at the Digital Workplace. In *Digital Workplace Learning* (pp. 3–8). Springer International Publishing. https://doi.org/10.1007/978-3-319-46215-8_1

Jansen, R. S., van Leeuwen, A., Janssen, J., Kester, L., & Kalz, M. (2017). Validation of the self-regulated online learning questionnaire. *Journal of Computing in Higher Education*, *29*(1), 6–27. https://doi.org/10.1007/s12528-016-9125-x

Janson, M. P., Siebert, J., & Dickhäuser, O. (2022). Compared to what? Effects of social and temporal comparison standards of feedback in an e-learning context. *International Journal of Educational Technology in Higher Education*, *19*(1). https://doi.org/10.1186/s41239-022-00358-2

Jivet, I., Scheffel, M., Drachsler, H., & Specht, M. (2017). Awareness Is Not Enough: Pitfalls of Learning Analytics Dashboards in the Educational Practice. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education* (Vol. 10474, Issue i, pp. 82–96). Springer International Publishing. 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. *Internet and Higher Education*, *47*. 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

Jonathan, C., Koh, E., & Pei-Ling Tan, J. (2022). *Examining the effectiveness of self-referenced and peer-referenced learning analytics dashboards in enhancing students' self-efficacy: Taking individual differences into account*. http://hdl.handle.net/10497/24341

Karaoglan Yilmaz, F. G. (2022). Utilizing Learning Analytics to Support Students' Academic Self-efficacy and Problem-Solving Skills. *The Asia-Pacific Education Researcher*, *31*(2), 175–191. https://doi.org/10.1007/s40299-020-00548-4

Konert, J., Bohr, C., Bellhauser, H., & Rensing, C. (2016). PeerLA - Assistant for individual learning goals and self-regulation competency improvement in online learning scenarios. *Proceedings - IEEE 16th International Conference on Advanced Learning Technologies, ICALT 2016*, 52–56. https://doi.org/10.1109/ICALT.2016.100

Langley, A., Lawson, G., Hermawati, S., D'Cruz, M., Apold, J., Arlt, F., & Mura, K. (2016). Establishing the Usability of a Virtual Training System for Assembly Operations within the Automotive Industry. *Human Factors and Ergonomics In Manufacturing*, *26*(6), 667–679. https://doi.org/10.1002/hfm.20406

Latham, G. P., & Locke, E. A. (1991). Self-regulation through goal setting. *Organizational Behavior and Human Decision Processes*, *50*(2), 212–247. https://doi.org/10.1016/0749-5978(91)90021-K

Leon-Perez, J. M., Medina, F. J., & Munduate, L. (2011). Effects of self-efficacy on objective and subjective outcomes in transactions and disputes. *International Journal of Conflict Management*, *22*(2), 170–189. https://doi.org/10.1108/10444061111126693

Li, K., Hall, M., Bermell-Garcia, P., Alcock, J., Tiwari, A., & González-Franco, M. (2017). Measuring the Learning Effectiveness of Serious Gaming for Training of Complex Manufacturing Tasks. *Simulation & Gaming*, *48*(6), 770–790. https://doi.org/10.1177/1046878117739929

Lim, L., Joksimović, S., Dawson, S., & Gasevic, D. (2019). *Exploring students' sensemaking of learning analytics dashboards: Does frame of reference make a difference?* 250–259. https://doi.org/10.1145/3303772.3303804

Lishinski, A., & Yadav, A. (2021). Self-evaluation Interventions: Impact on Self-efficacy and Performance in Introductory Programming. *ACM Transactions on Computing Education*, *21*(3), 1–28. https://doi.org/10.1145/3447378

Locke, E. A., & Latham, G. P. (2002). Building a practically useful theory of goal setting and task motivation: A 35-year odyssey. *American Psychologist*, *57*(9), 705–717. https://doi.org/10.1037/0003-066X.57.9.705

Locke, E. A., Latham, G. P., Locke, E. A., & Latham, G. P. (2015). New Directions in Goal-Setting Theory New Directions in Goal-Setting Theory. *Psychological Science*, *15*(October), 265–268.

Long, P. (2011). *Proceedings of the 1st international conference on learning analytics and knowledge*. ACM.

Lunenburg, F. C. (2011). Self-Efficacy in the Workplace: Implications for Motivation and Performance, International. *Journal of Management, Business and Administration*, *14*(1), 6.

Luszczynska, A., Gutiérrez-Doña, B., & Schwarzer, R. (2005). General self-efficacy in various domains of human functioning: Evidence from five countries. *International Journal of Psychology*, *40*(2), 80–89. https://doi.org/10.1080/00207590444000041

Maccallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power Analysis and Determination of Sample Size for Covariance Structure Modeling. In *Psychological Methods* (Vol. 2).

Makransky, G., & Petersen, G. B. (2021). The Cognitive Affective Model of Immersive Learning (CAMIL): a Theoretical Research-Based Model of Learning in Immersive Virtual Reality. *Educational Psychology Review*, *33*(3), 937–958. https://doi.org/10.1007/s10648-020-09586-2

Marshall, N., Shaw, K., Hunter, J., & Jones, I. (2020). Assessment by Comparative Judgement: An Application to Secondary Statistics and English in New Zealand. *New Zealand Journal of Educational Studies*, *55*(1), 49–71. https://doi.org/10.1007/s40841-020-00163-3

Marzouk, Z., Rakovic, M., Liaqat, A., Vytasek, J., Samadi, D., Stewart-Alonso, J., Ram, I., Woloshen, S., Winne, P. H., & Nesbit, J. C. (2016). What if learning analytics were based on learning science? *Australasian Journal of Educational Technology*, *32*(6), 1–18. https://doi.org/10.14742/ajet.3058

Matcha, W., Ahmad Uzir, N., Gasevic, D., & Pardo, A. (2019). A Systematic Review of Empirical Studies on Learning Analytics Dashboards: A Self-Regulated Learning Perspective. *IEEE Transactions on Learning Technologies*, *1382*(c), 1–1. https://doi.org/10.1109/tlt.2019.2916802

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

Mcintyre, K. P., & Eisenstadt, D. (2011). Social comparison as a self-regulatory measuring stick. *Self and Identity*, *10*(2), 137–151. https://doi.org/10.1080/15298861003676529

Molenaar, I. (2022). Towards hybrid human-AI learning technologies. *European Journal of Education*, *57*(4), 632–645. https://doi.org/10.1111/ejed.12527

Möller, J., & Marsh, H. W. (2013). Dimensional comparison theory. *Psychological Review*, *120*(3), 544–560. https://doi.org/10.1037/a0032459

Mussweiler, T. (2003). Comparison processes in social judgment: Mechanisms and consequences. *Psychological Review*, *110*(3), 472–489. https://doi.org/10.1037/0033-295X.110.3.472

Nicol, D. (2020). The power of internal feedback: exploiting natural comparison processes. *Assessment and Evaluation in Higher Education*, *0*(0), 1–23. https://doi.org/10.1080/02602938.2020.1823314

Nielsen, T., Makransky, G., Vang, M. L., & Dammeyer, J. (2017). How specific is specific self-efficacy? A construct validity study using Rasch measurement models. *Studies in Educational Evaluation*, *53*, 87–97. https://doi.org/10.1016/j.stueduc.2017.04.003

Nilsen, D., & Campbell, D. P. (1993). Self–observer rating discrepancies: Once an overrater, always an overrater? *Human Resource Management*, *32*(2–3), 265–281. https://doi.org/10.1002/hrm.3930320206

Osman, M. (2012). The effects of self set or externally set goals on learning in an uncertain environment. *Learning and Individual Differences*, *22*(5), 575–584. https://doi.org/10.1016/j.lindif.2011.09.012

Pan, W., Sun, L. Y., & Chow, I. H. S. (2011). The impact of supervisory mentoring on personal learning and career outcomes: The dual moderating effect of self-efficacy. *Journal of Vocational Behavior*, *78*(2), 264–273. https://doi.org/10.1016/j.jvb.2010.05.001

Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology*, *8*(APR), 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. https://doi.org/10.6018/analesps.30.2.167221

Panadero, E., Jonsson, A., & Botella, J. (2017). Effects of self-assessment on self-regulated learning and self-efficacy: Four meta-analyses. In *Educational Research Review* (Vol. 22, pp. 74–98). Elsevier Ltd. https://doi.org/10.1016/j.edurev.2017.08.004

Park, E., Ifenthaler, D., & Clariana, R. B. (2022). Adaptive or adapted to: Sequence and reflexive thematic analysis to understand learners' self-regulated learning in an adaptive learning analytics dashboard. *British Journal of Educational Technology*. https://doi.org/10.1111/bjet.13287

Pintrich, P. R. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ).*

Pintrich, P. R. (2000a). An Achievement Goal Theory Perspective on Issues in Motivation Terminology, Theory, and Research. *Contemporary Educational Psychology*, *25*(1), 92–104. https://doi.org/10.1006/ceps.1999.1017

Pintrich, P. R. (2000b). The Role of Goal Orientation in Self-Regulated Learning. In *Handbook of Self-Regulation* (pp. 451–502). Elsevier. https://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R., Conley, A. M., & Kempler, T. M. (2003). Current issues in achievement goal theory and research. *International Journal of Educational Research*, *39*(4–5), 319–337. https://doi.org/10.1016/j.ijer.2004.06.002

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

Poquet, O., Lim, L. A., & De Laat, M. (2022). Learning Analytics in the Corporate Sector: What Business Leaders Say? *IEEE Transactions on Learning Technologies*, *15*(5), 605–619. https://doi.org/10.1109/TLT.2022.3164783

Poyade, M., Eaglesham, C., Trench, J., & Reid, M. (2021). A Transferable Psychological Evaluation of Virtual Reality Applied to Safety Training in Chemical Manufacturing. *ACS Chemical Health & Safety*, *28*(1), 55–65. https://doi.org/10.1021/acs.chas.0c00105

Rets, I., Herodotou, C., Bayer, V., Hlosta, M., & Rienties, B. (2021). Exploring critical factors of the perceived usefulness of a learning analytics dashboard for distance university students. *International Journal of Educational Technology in Higher Education*, *18*(1). https://doi.org/10.1186/s41239-021-00284-9

Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: A systematic review and meta-analysis. *Psychological Bulletin*, *138*(2), 353–387. https://doi.org/10.1037/a0026838

Rigotti, T., Schyns, B., & Mohr, G. (2008). A short version of the occupational self-efficacy scale: Structural and construct validity across five countries. *Journal of Career Assessment*, *16*(2), 238–255. https://doi.org/10.1177/1069072707305763

Roberts, L. D., Howell, J. A., & Seaman, K. (2017). Give Me a Customizable Dashboard: Personalized Learning Analytics Dashboards in Higher Education. *Technology, Knowledge and Learning*, *22*(3), 317–333. https://doi.org/10.1007/s10758-017-9316-1

Ruble, D. N., & Flett, G. L. (1988). Conflicting goals in self-evaluative information seeking: developmental and ability level analyses. *Child Development*, *59*(1), 97–106. http://www.ncbi.nlm.nih.gov/pubmed/3342717

Ruiz-Calleja, A., Prieto, L. P., Ley, T., Rodríguez-Triana, M. J., & Dennerlein, S. (2017). Learning Analytics for Professional and Workplace Learning: A Literature Review. In É. Lavoué, H. Drachsler, K. Verbert, J. Broisin, & M. Pérez-Sanagustín (Eds.), *Data Driven Approaches in Digital Education* (Vol. 10474, pp. 164–178). Springer International Publishing. https://doi.org/10.1007/978-3-319-66610-5_13

Ryan, R. M., & Deci, E. L. (2000). Self-determination theory and the facilitation of intrinsic motivation, social development, and well-being. *American Psychologist*, *55*(1), 68–78. https://doi.org/10.1037/0003-066X.55.1.68

Sadri, G., & Robertson, I. T. (1993). *Self-efficacy and Work-related Behaviour : A Review and Meta-analysis*. *42*(2).

Schmidt, J. T., & Tang, M. (2020). Digitalization in Education: Challenges, Trends and Transformative Potential. In *Führen und Managen in der digitalen Transformation* (pp. 287–312). Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-28670-5_16

Schunk, D. H., & Pajares, F. (2002). The Development of Academic Self-Efficacy. In *Development of Achievement Motivation* (pp. 15–31). Elsevier. https://doi.org/10.1016/B978-012750053-9/50003-6

Schwendimann, B. A., Rodriguez-Triana, M. J., Vozniuk, A., Prieto, L. P., Boroujeni, M. S., Holzer, A., Gillet, D., & Dillenbourg, P. (2017). Perceiving Learning at a Glance: A Systematic Literature Review of Learning Dashboard Research. *IEEE Transactions on Learning Technologies*, *10*(1), 30–41. https://doi.org/10.1109/TLT.2016.2599522

Seo, E., Patall, E. A., Henderson, M. D., & Steingut, R. R. (2018). The Effects of Goal Origin and Implementation Intentions on Goal Commitment, Effort, and Performance. *The Journal of Experimental Education*, *86*(3), 386–401. https://doi.org/10.1080/00220973.2016.1277334

Serdyukov, P. (2017). Innovation in education: what works, what doesn't, and what to do about it? *Journal of Research in Innovative Teaching & Learning*, *10*(1), 4–33. https://doi.org/10.1108/JRIT-10-2016-0007

Shum, S. B., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-centred learning analytics. *Journal of Learning Analytics*, *6*(2), 1–9. https://doi.org/10.18608/jla.2019.62.1

Siemens, G., & Baker, R. S. J. D. (2012). Learning analytics and educational data mining: Towards communication and collaboration. *ACM International Conference Proceeding Series*, *December*, 252–254. https://doi.org/10.1145/2330601.2330661

Siemens, G., & Gasevic, D. (2012). Guest editorial - learning and knowledge analytics. *Educational Technology and Society*, *15*(3), 1–2.

Song, J. H., Chai, D. S., Kim, J., & Bae, S. H. (2018). Job Performance in the Learning Organization: The Mediating Impacts of Self-Efficacy and Work Engagement. *Performance Improvement Quarterly*, *30*(4), 249–271. https://doi.org/10.1002/piq.21251

Srinivasan, B., Iqbal, M. U., Shahab, M. A., & Srinivasan, R. (2022). Review of Virtual Reality (VR) Applications To Enhance Chemical Safety: From Students to Plant Operators. *ACS Chemical Health & Safety*, *29*(3), 246–262. https://doi.org/10.1021/acs.chas.2c00006

Stajkovic, A. D., & Luthans, F. (1998). Self-Efficacy and Work-Related Performance: A Meta-Analysis. *Psychological Bulletin*, *124*(2), 240–261. https://doi.org/10.1037/0033-2909.124.2.240

Suls, J., Martin, R., & Wheeler, L. (2002). Social Comparison: Why, With Whom, and With What Effect? *Current Directions in Psychological Science*, *11*(5), 159–163. https://doi.org/10.1111/1467-8721.00191

Susnjak, T., Ramaswami, G. S., & Mathrani, A. (2022). Learning analytics dashboard: a tool for providing actionable insights to learners. *International Journal of Educational Technology in Higher Education*, *19*(1). https://doi.org/10.1186/s41239-021-00313-7

Taber, K. S. (2018). The Use of Cronbach's Alpha When Developing and Reporting Research Instruments in Science Education. *Research in Science Education*, *48*(6), 1273–1296. https://doi.org/10.1007/s11165-016-9602-2

Tabuenca, B., Kalz, M., Drachsler, H., & Specht, M. (2015). Time will tell: The role of mobile learning analytics in self-regulated learning. *Computers and Education*, *89*, 53–74. https://doi.org/10.1016/j.compedu.2015.08.004

Talsma, K., Schüz, B., & Norris, K. (2019). Miscalibration of self-efficacy and academic performance: Self-efficacy ≠ self-fulfilling prophecy. *Learning and Individual Differences*, *69*, 182–195. https://doi.org/10.1016/j.lindif.2018.11.002

Taylor, R. (2012). *Review of the motivated strategies for learning questionnaire (MSLQ) using reliability generalization techniques to assess scale reliability* [Auburn University]. https://eric.ed.gov/?id=ED546666

Tehreem, Y., Fracaro, S. G., Gallagher, T., Toyoda, R., Bernaerts, K., Glassey, J., Abegao, F. R., Wachsmuth, S., Wilk, M., & Pfeiffer, T. (2022). May I Remain Seated: A Pilot Study on the Impact of Reducing Room-Scale Trainings to Seated Conditions for Long Procedural Virtual Reality Trainings. *2022 8th International Conference on Virtual Reality (ICVR)*, 62–71. https://doi.org/10.1109/ICVR55215.2022.9848222

Tempelaar, Rienties, B., & Nguyen, Q. (2020a). Individual differences in the preference for worked examples: Lessons from an application of dispositional learning analytics. *Applied Cognitive Psychology*, *34*(4), 890–905. https://doi.org/10.1002/acp.3652

Tempelaar, Rienties, B., & Nguyen, Q. (2020b). Subjective data, objective data and the role of bias in predictive modelling: Lessons from a dispositional learning analytics application. *PLOS ONE*, *15*(6), e0233977. https://doi.org/10.1371/journal.pone.0233977

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

Udeozor, C., Toyoda, R., Russo Abegão, F., & Glassey, J. (2021). Perceptions of the use of virtual reality games for chemical engineering education and professional training. *Higher Education Pedagogies*, *6*(1), 175–194. https://doi.org/10.1080/23752696.2021.1951615

Valle, N., Antonenko, P., Dawson, K., & Huggins-Manley, A. C. (2021). Staying on target: A systematic literature review on learner-facing learning analytics dashboards. *British Journal of Educational Technology*, *52*(4), 1724–1748. https://doi.org/10.1111/bjet.13089

Valle, N., Antonenko, P., & Valle, D. (2023). Dashboard Applications to Support Motivation: A Design Case. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology* (pp. 1–37). Springer International Publishing. https://doi.org/10.1007/978-3-319-17727-4_184-1

Valle, N., Antonenko, P., Valle, D., Sommer, M., Huggins-Manley, A. C., Dawson, K., Kim, D., & Baiser, B. (2021). Predict or describe? How learning analytics dashboard design influences motivation and statistics anxiety in an online statistics course. *Educational Technology Research and Development*, *69*(3), 1405–1431. https://doi.org/10.1007/s11423-021-09998-z

van Leeuwen, A., Strauß, S., & Rummel, N. (2023). Participatory design of teacher dashboards: navigating the tension between teacher input and theories on teacher professional vision. *Frontiers in Artificial Intelligence*, *6*. https://doi.org/10.3389/frai.2023.1039739

Van Lissa, C. J., Gu, X., Mulder, J., Rosseel, Y., Van Zundert, C., & Hoijtink, H. (2021). Teacher's Corner: Evaluating Informative Hypotheses Using the Bayes Factor in Structural Equation Models. *Structural Equation Modeling: A Multidisciplinary Journal*, *28*(2), 292–301. https://doi.org/10.1080/10705511.2020.1745644

Van Yperen, N. W., & Leander, N. P. (2014). The Overpowering Effect of Social Comparison Information. *Personality and Social Psychology Bulletin*, *40*(5), 676–688. https://doi.org/10.1177/0146167214523475

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly: Management Information Systems*, *27*(3), 425–478. https://doi.org/10.2307/30036540

Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J. L. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist*, *57*(10), 1500–1509. https://doi.org/10.1177/0002764213479363

Verbert, K., Ochoa, X., De Croon, R., Dourado, R. A., & De Laet, T. (2020). *Learning analytics dashboards: The past, the present and the future*. 35–40. https://doi.org/10.1145/3375462.3375504

Verhavert, S., Bouwer, R., Donche, V., & de Maeyer, S. (2019). A meta-analysis on the reliability of comparative judgement. *Assessment in Education: Principles, Policy and Practice*, *26*(5), 541–562. https://doi.org/10.1080/0969594X.2019.1602027

Verhavert, S., de Maeyer, S., Donche, V., & Coertjens, L. (2018). Scale Separation Reliability: What Does It Mean in the Context of Comparative Judgment? *Applied Psychological Measurement*, *42*(6), 428–445. https://doi.org/10.1177/0146621617748321

Wheeler, L., & Suls, J. (2020). A History of Social Comparison Theory. In *Social Comparison, Judgment, and Behavior* (pp. 5–31). Oxford University Press. https://doi.org/10.1093/oso/9780190629113.003.0001

White, R. T., & Arzi, H. J. (2005). Longitudinal Studies: Designs, Validity, Practicality, and Value. *Research in Science Education*, *35*(1), 137–149. https://doi.org/10.1007/s11165-004-3437-y

Whitehouse, C., & Pollitt, A. (2012). Centre for Education Research and Policy Using Adaptive Comparative Judgement to Obtain a Highly Reliable Rank Order in Summative Assessment. *Centre for Education Research and Policy*, *3644723*. www.cerp.org.uk

Wilson, A. E., & Ross, M. (2000). The frequency of temporal-self and social comparisons in people's personal appraisals. *Journal of Personality and Social Psychology*, *78*(5), 928–942. https://doi.org/10.1037/0022-3514.78.5.928

Wilson, A. E., & Shanahan, E. (2020). Temporal Comparisons in a Social World. In J. Suls, R. L. Collins, & L. Wheeler (Eds.), *Social Comparison, Judgment, and Behavior* (pp. 309–344). Oxford University Press. https://doi.org/10.1093/oso/9780190629113.003.0012

Winne, P. H. (2017). Learning Analytics for Self-Regulated Learning. *Handbook of Learning Analytics*, 241–249. https://doi.org/10.18608/hla17.021

Winstone, N. E., Nash, R. A., Parker, M., & Rowntree, J. (2017). Supporting Learners' Agentic Engagement With Feedback: A Systematic Review and a Taxonomy of Recipience Processes. *Educational Psychologist*, *52*(1), 17–37. https://doi.org/10.1080/00461520.2016.1207538

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*, 203–211. https://doi.org/10.1145/2567574.2567588

Wise, A. F., & Vytasek, J. (2017). Learning Analytics Implementation Design. In *Handbook of Learning Analytics* (pp. 151–160). Society for Learning Analytics Research (SoLAR). https://doi.org/10.18608/hla17.013

Wong, B. T. ming, Li, K. C., & Cheung, S. K. S. (2022). An analysis of learning analytics in personalised learning. *Journal of Computing in Higher Education*. https://doi.org/10.1007/s12528-022-09324-3

Wood, J. V. (1996). What is Social Comparison and How Should We Study it? *Personality and Social Psychology Bulletin*, *22*(5), 520–537. https://doi.org/10.1177/0146167296225009

Wortha, F., Azevedo, R., Taub, M., & Narciss, S. (2019). Multiple Negative Emotions During Learning With Digital Learning Environments – Evidence on Their Detrimental Effect on Learning From Two Methodological Approaches. *Frontiers in Psychology*, *10*. https://doi.org/10.3389/fpsyg.2019.02678

Zell, E., & Alicke, M. D. (2010). Comparisons over time: Temporal trajectories, social comparison, and self-evaluation. *European Journal of Social Psychology*, n/a-n/a. https://doi.org/10.1002/ejsp.737

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

Zimmerman, B. J. (2000b). Self-Efficacy: An Essential Motive to Learn. *Contemporary Educational Psychology*, *25*(1), 82–91. https://doi.org/10.1006/ceps.1999.1016

Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice*, *41*(2), 64–70. https://doi.org/10.1207/s15430421tip4102_2

Zimmerman, B. J. (2012). Development and Adaptation of Expertise: The Role of Self-Regulatory Processes and Beliefs. In K. A. Ericsson, N. Charness, P. J. Feltovich, & R. R. Hoffman (Eds.), *The Cambridge Handbook of Expertise and Expert Performance* (pp. 705–722). Cambridge University Press. https://doi.org/10.1017/cbo9780511816796.039

Zimmerman, B. J. (2013). From Cognitive Modeling to Self-Regulation: A Social Cognitive Career Path. *Educational Psychologist*, *48*(3), 135–147. https://doi.org/10.1080/00461520.2013.794676

Zimmerman, B. J. (2015). Self-Regulated Learning: Theories, Measures, and Outcomes. In *International Encyclopedia of the Social & Behavioral Sciences: Second Edition* (Second Edi, Vol. 21). Elsevier. https://doi.org/10.1016/B978-0-08-097086-8.26060-1

Zimmerman, B. J., Bandura, A., & Martinez-Pons, M. (1992). Self-Motivation for Academic Attainment: The Role of Self-Efficacy Beliefs and Personal Goal Setting. *American Educational Research Journal*, *29*(3), 663. https://doi.org/10.2307/1163261

Zimmerman, B. J., & Campillo, M. (2003). Motivating Self-Regulated Problem Solvers. In *The Psychology of Problem Solving* (pp. 233–262). Cambridge University Press. https://doi.org/10.1017/CBO9780511615771.009

Zimmerman, B. J., & Moylan, A. R. (2009). Self-regulation: Where metacognition and motivation intersect. In *Handbook of metacognition in education.* (pp. 299–315). Routledge/Taylor & Francis Group.

List of publications

Gallagher, T., Slof, B., van der Schaaf, M., Arztmann, M., Garcia Fracaro, S., & Kester, L. (2023). Learning analytics dashboard design: Workplace learner preferences for reference frames in immersive training in practice. S*ubmitted.* Department of Education, Utrecht University

Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2022). *Comparison with Self vs Comparison with Others: The Influence of Learning Analytics Dashboard Design on Learner Dashboard Use* (pp. 11–21). https://doi.org/10.1007/978-3-031-22124-8_2

Gallagher, T., Slof, B., Van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2023). Impact of Reference Frames on Learning Analytics Dashboard Use. S*ubmitted*. Department of Education, Utrecht University.

Gallagher, T., Slof, B., van der Schaaf, M., Hoijtink, H., & Kester, L. (2023). Learning analytics reference frame type and direction of comparison affect academic self-efficacy among higher education students. S*ubmitted.* Department of Education, Utrecht University.

Gallagher, T., Slof, B., van der Schaaf, M., Toyoda, R., Tehreem, Y., Garcia Fracaro, S., & Kester, L. (2023). Reference frames for learning analytics dashboards: The progress and social reference frame and occupational self-efficacy. Journal of Computer Assisted Learning, 1–19. https://doi.org/10.1111/jcal.12912

Garcia Fracaro, S., Chan, P., Gallagher, T., Tehreem, Y., Toyoda, R., Bernaerts, K., Glassey, J., Pfeiffer, T., Slof, B., Wachsmuth, S., & Wilk, M. (2021). Towards design guidelines for virtual reality training for the chemical industry. *Education for Chemical Engineers*, *36*, 12–23. https://doi.org/10.1016/j.ece.2021.01.014

Hu, Y., Gallagher, T., Wouters, P., van der Schaaf, M., & Kester, L. (2022). Game-based learning has good chemistry with chemistry education: A three-level meta-analysis. *Journal of Research in Science Teaching*. https://doi.org/10.1002/tea.21765

Tehreem, Y., Fracaro, S. G., Gallagher, T., Toyoda, R., Bernaerts, K., Glassey, J., Abegao, F. R., Wachsmuth, S., Wilk, M., & Pfeiffer, T. (2022). May I Remain Seated: A Pilot Study on the Impact of Reducing Room-Scale Trainings to Seated Conditions for Long Procedural Virtual Reality Trainings. *2022 8th International Conference on Virtual Reality (ICVR)*, 62–71. https://doi.org/10.1109/ICVR55215.2022.9848222