Deliverable 2.4. Theoretical model and design guidelines for (re)designing blended learning

ALO!-project - Work package 2:

PhD students: Stijn Van Laer[1] & Ruth Boelens[2]

Supervisors: Jan Elen[1] & Bram De Wever[2]

[1] KU Leuven, Center for Instructional Psychology and Technology

[2] Ghent University, Department of Educational Studies

Preface

For deliverable 2.4, work package 2 further elaborated the theoretical models proposed and empirically investigated these models to develop and test design guidelines for blended learning. In this deliverable we focus on the use of cues for reflection, cues for calibration, interaction, and open-ended tasks. Four empirical studies were administered to examine the link between learning and instruction. Below you can find a short summary of the four studies, followed by each of them.

The Effect of Cues for Reflection on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

Stijn Van Laer, Lai Jiang, & Jan Elen (KU Leuven, Centre for Instructional Psychology and Technology)

Literature on blended learning emphasizes the importance of self-regulation for learning in blended learning environments and the role of learners' self-reflection for self-regulated learning. Despite this, evidence is inconclusive about what effect support for reflection actually has on learners' self-regulated learning. Little insight has been provided into why previous findings are inconclusive, or how to overcome this. Our study investigates whether cues for reflection that are designed in line with current literature affect learners' self-regulated learning. We investigate this effect by examining changes in learners' learning behaviour and outcomes. A pre-post and control-experimental design was applied in a blended learning environment in which learners in the experimental condition received cues for reflection, while learners in the control group did not. Learners' behaviour was analysed using event sequence analysis and learners' learning outcomes using mixed ANOVA. The results show that although no significant differences were observed in cue-for-reflection use between conditions, learners in the experimental condition used significantly more sequences related to self-testing and monitoring others. As regards learners' learning outcomes, we observed a significant increase in unfavourable motivational outcomes. This paper discusses these unexpected results in terms of their theoretical and practical implications and provides recommendations for future research. We conclude that if cues for reflection are designed in line with current literature, self-reflection is evoked, and affects self-regulated learning. However, for this self-regulated learning to be effective, learners need to be guided towards the strategies required to meet instructional expectations successfully. Further investigations of this hypothesis could allow us to progress towards less inconclusive results on the effect of cues for reflection.

Van Laer, S., Jiang, L., & Elen, J. (2018) The Effect of Cues for Reflection on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes. *The Internet and Higher Education*, under review.

The Effect of Cues for Calibration on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

Stijn Van Laer & Jan Elen (KU Leuven, Centre for Instructional Psychology and Technology)

Literature on blended learning emphasizes the importance of self-regulation for learning in blended learning environments and the role of learners' calibration. Although literature on calibration is clear on its importance for self-regulated learning, it provides inconclusive insight in the effect of support for calibration on learners' self-regulated learning. One under-investigated avenue might be learners' ability to enact on the cues provided. In order to establish a more accurate picture of the effect of support for calibration on self-regulated learning, our study investigates whether providing cues for calibration affect learners' self-regulated learning, and whether this effect is different for learners with different metacognitive abilities. We investigate this effect by examining changes in learners' learning behaviour and outcomes. A pre-post design with one control and two experimental conditions was applied in a blended learning environment. Learners in the experimental conditions received either functional validity feedback (F-condition) or functional and cognitive validity feedback (FC-condition). Learners in the control condition did not receive any cues. Learners' behaviour was analysed using event sequence analysis. Learners' post-test learning scores were subjected to multivariate analysis of

covariance, with condition and learners' metacognitive ability as independent variables. The results show a significant and unexpected impact of condition and learners' metacognitive abilities on learners' learning behaviour and outcomes. This manuscript discusses the unexpected results in terms of their theoretical and practical implications and provides recommendations for future research. We conclude that when cues for calibration are provided through functional and cognitive validity feedback, learners' calibration capabilities will increase. Yet for this to result in goal-directed self-regulated learning, learners' need to be supported on how to apply the cognitive and metacognitive strategies needed.

Van Laer, S., & Elen, J. (2018) The Effect of Cues for Calibration on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes. *Computers and Education*, under review.

How can teachers support student interaction during computer-supported collaborative learning? An exploratory case study in a higher education setting for hands-on learners

Ruth Boelens & Bram De Wever (University of Ghent, Department of Educational Studies)

This study focused on how teacher support might influence students' interactions during collaboration. As little research on computer-supported collaborative learning (CSCL) is conducted in contexts that include students with a background in vocational or technical secondary education, an exploratory case study was carried out with 1 teacher and 10 students with prior craft knowledge aiming to become vocational teachers. The purpose was to make visible, with empirical examples, how teacher support might contribute to more productive student interaction during CSCL. General understanding on the quality of student interaction during collaboration was obtained through qualitative content analysis, while a detailed interpretative analysis of the dialogues between the teacher and the groups made it possible to trace different interaction patterns. The results suggest that it was hard for both students and the teacher to empower productive interactions. Two interaction patterns showed to be effective: the pattern in which the teacher takes the role of the devil's advocate, and the pattern in which the teacher provides gradual assistance. The latter finding suggests that not only the type of questions, but also the order of the questions asked by the teacher, is important.

Boelens, R., & De Wever, B. (under review). How can teachers support student interaction during computer-supported collaborative learning? An exploratory case study in a higher education setting for hands-on learners. *British Journal of Educational Technology*.

Conjecture mapping to support hands-on adult learners in open-ended tasks

This case reports on a teacher education course that aimed to support vocationally educated adults, referred to as hands-on learners, to accomplish open-ended tasks. Conjecture mapping was used to identify the most salient design features, and to test if, how, and why these course features supported learners. Inspired by ethnographic approaches, sustained engagement and multiple data sources were used to explain the effects of the course design on participants' behavior and perceptions: student and teacher interviews, observations, and artifacts. The results reveal that almost all of the proposed design features stimulated the participants toward the intended enactment processes, which in turn yielded the intended learning outcomes. For instance, worked examples (i.e., design feature) not only engendered the production of artifacts that meet high standards (i.e., enactment process) because they clarify the task requirements, but also fostered a safe structure (i.e., enactment process) by providing an overall picture of the task. The conjecture map resulting from this study provides a theoretical frame to describe, explain, and predict how specific course design features support hands-on adult learners in open-ended tasks, and assists those who aim to implement open-ended tasks in similar contexts.

Boelens, R., McKenney, S., & De Wever, B. (under review). Conjecture mapping to support hands-on adult learners in open-ended tasks. *Journal of the Learning Sciences*.

The Effect of Cues for Reflection on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

Abstract

Literature on blended learning emphasizes the importance of self-regulation for learning in blended learning environments and the role of learners' self-reflection for self-regulated learning. Despite this, evidence is inconclusive about what effect support for reflection actually has on learners' self-regulated learning. Little insight has been provided into why previous findings are inconclusive, or how to overcome this. Our study investigates whether cues for reflection that are designed in line with current literature affect learners' self-regulated learning. We investigate this effect by examining changes in learners' learning behaviour and outcomes. A pre-post and control-experimental design was applied in a blended learning environment in which learners in the experimental condition received cues for reflection, while learners in the control group did not. Learners' behaviour was analysed using event sequence analysis and learners' learning outcomes using mixed ANOVA. The results show that although no significant differences were observed in cue-for-reflection use between conditions, learners in the experimental condition used significantly more sequences related to self-testing and monitoring others. As regards learners' learning outcomes, we observed a significant increase in unfavourable motivational outcomes. This paper discusses these unexpected results in terms of their theoretical and practical implications and provides recommendations for future research. We conclude that if cues for reflection are designed in line with current literature, self-reflection is evoked, and affects self-regulated learning. However, for this self-regulated learning to be effective, learners need to be guided towards the strategies required to meet instructional expectations successfully. Further investigations of this hypothesis could allow us to progress towards less inconclusive results on the effect of cues for reflection.

Keywords: cues for reflection; self-regulated learning; learning outcomes; learning behaviour; computer log files

Highlights:

- Cues for reflection evoked changes in learners' self-regulated learning
- Performance avoidance outcomes and behaviour increased
- Cues were used similarly by learners with different characteristics
- Tailoring cues to current guidelines alone seems insufficient to improve learning
- Cues for reflection need to be accompanied by cues for strategy use

1. Introduction

The current literature on technology-enhanced learning emphasizes the importance of self-regulation in blended learning (e.g., Boekaerts, 1999; Greene & Azevedo, 2007; Vohs & Baumeister, 2016) and the role of learners' self-reflection in self-regulated learning (e.g., Lin, Coburn, & Eisenberg, 2016; Pajares & Schunk, 2001) based on the belief that learners in blended learning environments need to be able to deal with varying degrees of autonomy and to judge and adapt their learning to the learning outcomes imposed on them. Although instructional interventions fostering self-regulated learning have been investigated widely in different educational settings (e.g., Arrastia-Chisholm, Torres, & Tackett, 2017; Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015), evidence remains inconclusive regarding what effect support for reflection actually has on learners' self-regulated learning (Roessger, 2014). Results indicate positive effects (e.g., Bannert, 2006), no effects (e.g., van den Boom, Paas, & van Merriënboer, 2007) and negative effects (e.g., Furberg, 2009). The literature described above provides little insight into why the findings are inconclusive, or how to overcome this. In order to establish a more accurate picture of the effect of support for reflection on self-regulated learning, this study aims to enrich current insights by investigating whether cues for reflection that are designed in line with the literature on designing such cues to foster self-regulated learning do actually affect self-regulated learning – and through self-reflection – in a blended learning environment. We operationalize self-regulated learning as changes in learners' learning behaviour and outcomes. Investigating learning behaviour and outcomes provides insights not only into learners' self-regulated learning, but also into the nature of cues' effects (Gašević, Dawson, & Siemens, 2015). In the next part of the introduction we elaborate on blended learning and the conceptualization of self-regulated learning and present a theoretical basis for designing reflection cues intended to evoke self-regulated learning. In the final part of the introduction we focus on the significance of relationships between self-regulated learning, learning behaviour and outcomes for investigating changes in self-regulated learning.

1.1. Blended learning

Blended learning is a popular concept. A common aspect in many definitions of blended learning is that it combines online and face-to-face learning. Hence, it is assumed that blended learning environments combine the advantages of both modes of delivery (Graham, Henrie, & Gibbons, 2014; McCutcheon, Lohan, Traynor, & Martin, 2015). In line with this, blended learning is defined as learning in an instructional context which is characterized by a deliberate combination of online and classroom-based interventions to instigate and support learning (Boelens, Van Laer, De Wever, & Elen, 2015). Blended learning as a notion is widely used in higher and adult education (Allen, Seaman, & Garrett, 2007); K-12 education (Picciano, Seaman, Shea, & Swan, 2012); and corporate training (Bonk, 2017; Spring & Graham, 2017). Over the years, blended learning has been the focus of many research studies (Drysdale, Graham, Spring, & Halverson, 2013). The majority of studies on blended learning have focused either on comparing blended and face-to-face learning (Halverson, Graham, Spring, Drysdale, & Henrie, 2014) or on the characteristics learners need to thrive in such environments (Deschacht & Goeman, 2015). With regard to the latter, research for example has identified that learners with high amounts of verbal ability and self-efficacy (Lynch & Dembo, 2004) and learners with high self-regulatory capabilities (e.g., Kizilcec, Perez-Sanagustin, & Maldonado, 2017; Kuo, Walker, Schroder, & Belland, 2014) often perform better in blended learning environments compared to learners who lack these capabilities. Despite the importance of these types of research, hardly any research propels the quest for empirical evidence to support the design of such environments in which less 'capable' learners can also find success (Van Laer & Elen, 2018).

1.2. Self-regulated learning

Blended learning, like any learning, is an activity performed by learners rather than something happening to them as result of instruction (e.g., Bandura, 1989; Oliver & Trigwell, 2005) and so can be seen as a self-regulated process in which learners' regulate their behaviour according to the instructional demands (Zimmerman & Schunk, 2001). This is evidenced by a substantial body of literature showing that scores on performance-related variables are strongly positively correlated and have causal relations with scores on self-regulated-learning-related variables (e.g., Daniela, 2015; Lin et al., 2016). Over the past three decades, various self-regulated learning theories have been proposed (see: Puustinen & Pulkkinen, 2001). Each of these theories describes a similar cyclic process of self-regulatory phases, often consisting of (a) a forethought, (b) an enacting, and (c) an evaluation phase. In relation to this, the metacognitive processes occurring in each of these phases cannot be directly observed as they manifest in cognitive behaviours (i.e., learners' learning behaviour) and behavioural consequences (i.e., learners' learning outcomes) (Veenman & Alexander, 2011). For instance, when a learner recalculates the outcome of a mathematical equation, it is assumed that a monitoring or evaluation process must have preceded this overt cognitive activity of recalculation. Finally with regard to variables influencing learners' self-regulated learning, recent theories regard self-regulated learning as a concept superordinate to metacognition, influencing learners' cognitive, motivational and metacognitive characteristics and so future self-regulated learning (Veenman, Van Hout-Wolters, & Afflerbach, 2006). Each of the models also sees the influence of the instructional context as key influence to the development of self-regulated learning (Zimmerman, 2013). Taking both the influence of variables internal and external to the learner into account, self-regulated learning can be seen as dynamic in nature over time. Because of this nature and the inferential characteristics of self-regulated learning continues measurements and inferences based on learners' learning behaviour and outcomes are needed to capture learners' self-regulated learning.

1.3. Self-reflection and self-regulated learning

Borkowski, Carr, Rellinger, and Pressley (1990) define learners' self-reflection as a strategy or skill that operates on other strategies. For example, when learners recognize that a particular cognitive strategy (e.g., making a concept map) not seems to lead to retention, they might or might not switch to another strategy (e.g., self-questioning). By reflecting on their own learning, learners become aware of their learning processes and possible alternative strategies. This example illustrates that self-regulated learning includes a self-reflective phase in which performance measured in terms of internal or external feedback is evaluated (Winters, Greene, & Costich, 2008). Ample indications show self-reflection's role in self-regulated learning by the creation of the perception of choice and awareness of the need for alternatives, which are both critical elements for self-regulated learning (e.g., Järvelä, Järvenoja, Malmberg, Isohätälä, & Sobocinski, 2016; Nicol & Macfarlane-Dick, 2006).

1.4. Supporting self-reflection for self-regulated learning

Based on the relationship between self-reflection and self-regulated learning and in accordance with Butler (1998) and Winters et al. (2008) we assume that cues for reflection used to provoke learners' reflections are well suited to evoke learners' self-regulated learning. Additionally, there is no theoretical reason why self-regulated learning would not be at stake in blended learning environments, on the contrary (e.g., Lehmann, Hähnlein, & Ifenthaler, 2014; Lord, Chen, Cheng, Tai, & Pan, 2017). Nonetheless, empirical studies seem not to be able to demonstrate conclusive results when it comes to the effect of cues for reflection on learners' reflection and so learners' self-regulated learning. In their extensive reviews, Roessger (2014) and Kori, Pedaste, Leijen, and Mäeots (2014) reported inconclusive results and provided no conclusions on what kind of cues under what configuration supports self-reflection most. Moreover, hardly any explanations are given for the mixed effects of cues for reflection. Only few studies claim causality or correlation, revealing a positive relationship between cues for reflection and learners' self-regulated learning, and elaborate on the possible nature of these findings (e.g., Pannese & Morosini, 2014; Renner et al., 2014). Without insights in the nature of the effects of cues for reflection on learners' reflection and self-regulated learning, teachers and instructional designers remain dependent on inconsistent conceptual claims that cueing self-reflection may improve self-regulated learning (e.g., Burke, Scheuer, & Meredith, 2007; Mezirow, 2000; Van Woerkom, 2004).

Literature focusses on two types of support when it comes to cues for self-reflection (e.g., Pannese & Morosini, 2014; Renner et al., 2014). On the one hand there are cues for reflection for cognitive support and on the other hand there are those for metacognitive support. Cognitive support refers to cues to evoke learners' reflection on the understanding of content (Reiser et al., 2001). Metacognitive support refers to cues for reflection evoking learners to reflect on their use of metacognitive strategies (Kori et al., 2014). Current research reporting positive effects of cues for reflection on learners' self-regulated learning indicates cues for reflection focussing on jointly the cognitive and metacognitive support seem more successful than cues solely allowing learners to engage either with cognitive or metacognitive support (e.g., Chen, Wei, Wu, & Uden, 2009; Davis, 2003; Sonnenberg & Bannert, 2015; van den Boom, Paas, van Merriënboer, & van Gog, 2004). Combining cognitive and metacognitive cues for reflection may help learners identify support tools, information, and monitor task engagement potentially enhancing learners' self-regulated learning (Butler & Winne, 1995).

Although it is of vital importance to target cues for reflection so learners will be directed to the right types of information, cue-use research shows that this is only one part of the challenge. They evidence that not all learners equally use and benefit from cues provided (e.g., Lust, Vandewaetere, Ceulemans, Elen, & Clarebout, 2011; Rashid & Asghar, 2016; Winne & Hadwin, 1998). In general three characteristics of cues for reflection can be extracted from literature: (1) timing, (2) focus, and (3) integration. In relation to the timing of cues for reflection, literature indicates that cues for reflection can occur at three different moments centred around the task at hand (e.g., Farrall, 2007; Mann, Gordon, & MacLeod, 2009): before the task, during the task, and after the task. Cues for reflection provided before the task aim to trigger learners' proactive reflection (e.g., Lehmann et al., 2014). Cues presented during the task aim to trigger learners' reflection while learners are performing a task and encourage learners to reflect upon the need to alter, amend, or change what they are doing in order to adjust to changing circumstances (e.g., Lavoué, Molinari, Prié, & Khezami, 2015). Finally cues for reflection after the task aim to encourage learners to reflect on what they have done to discover what metacognitive strategies they used and perceived as successful (e.g., Embo, Driessen, Valcke, & Van Der Vleuten, 2014; Hatton & Smith, 1995)

As we aim to impact learners' self-regulated learning it is important to identify which part of the self-regulated learning process should be targeted to trigger learners' to use their most effective metacognitive strategies. Winne and Hadwin (1998) suggested that the first two phases of self-regulation, namely task identification and goal-setting and planning are most prone (e.g., Feyzi-Behnagh et al., 2014; Rubenstein, Callan, & Ridgley, 2017) to changes by interventions as they seem to be most directional for change in learning (Wehmeyer & Shogren, 2017). With regard to task identification learners need to identify the task at hand, the information supporting the execution of the task, and the strategies needed to solve the task. In relation to goal-setting and planning learners need to identify the steps needed to complete the task at hand, how to plan for successful completion of the task, and which cognitive and metacognitive strategies to use. Keeping this in mind, combining on the one hand cognitive and metacognitive information and on the other hand task identification and goal-setting and planning seem to make up the most optimal content of cues for reflection to foster self-regulated learning (e.g., Bannert, 2006; Bannert & Mengelkamp, 2013; Chen et al., 2009; Davis, 2000, 2003; Sonnenberg & Bannert, 2015; van den Boom et al., 2004).

As indicated before, different authors in the past (e.g., Aleven, Stahl, Schworm, Fischer, & Wallace, 2003) recognized the problem of suboptimal use of cues. Greene and Azevedo (2007) and Winne and Jamieson-Noel (2002) indicated that the lack or inadequate use of cues might be related to the considerations learners need to make when deciding on the use of a cue. If learners are not able to make functional decisions about the use of cues for reflection, they should benefit from a learning environment that provides embedded cues. Embedded cues are defined as unavoidable by the learner. However, embedding cues may not necessarily solve all problems (Clarebout & Elen, 2009). It cannot be guaranteed that learners make use of the support and use the cues as intended (Land & Greene, 2000).

1.5. Investigating learners' self-regulated learning

Due to self-regulated learning's dynamic nature and its inferential characteristics, continues measurements and inferences based on learners' learning behaviour and outcomes are needed to capture learners' self-regulated learning. Investigating both learners' learning behaviour and outcomes does not only provide insights in learners' self-regulated learning, but also in the nature of cues' effects (e.g., Bannert, Molenaar, Azevedo, Järvelä, & Gašević, 2017).

1.5.1. Learning behaviour and self-regulated learning

Learners' learning behaviour is defined as the behavioural traces gathered from a learner during instructional processes (Entwistle & Peterson, 2004). Due to the nature of self-regulated learning continues measurements in the form of on-line measures gained substantial interest. These types of measurements are based on actual learners' learning behaviour. Typical online measures include observational methods, thinking-aloud protocols, eye-movement, or log file registration. The unobtrusiveness of some of these methods (i.e., log file registration) enables researchers to track learning events with high ecologically validity. Tracing methods such as log file analysis also allow researchers to track learning events in a nonlinear environment and "re-play" learners' behaviour. Such an approach might allow a fuller account of how learning outcomes come to be.

The idea that self-regulation unfolds in different phases suggests a cyclical relationship among the components of self-regulated learning. In this respect Cleary, Callan, and Zimmerman (2012) coined the term 'sequential phases of regulation'. This cyclical relationship between the components of self-regulated learning combined with the use of log file data is constituted in the term 'event sequence' to describe patterns of learners' behaviour. Although 'event sequence' in relation to self-regulated learning implies specific theoretical assumptions and methodological approaches, both "event" and "sequence" are words often used in different fields of research to describe all sorts of ordered events comprised in patterns (e.g., Abbott, 1995; Suthers & Verbert, 2013). In the light of self-regulated learning the first distinction to be made between trace data types is whether the basic information they contain relates to a state or an event. Simply put, each change of state is an event, and each event implies a change of state (Müller, Studer, Gabadinho, & Ritschard, 2010). For example in relation to log files, a state could be being on a page, while clicking the calendar tool in the online learning environment would be an event that changes the state to being on a different page. Log files are only able to capture events as we are (without triangulation) not able to determine what learners' are actually doing between two consecutive events (e.g., getting coffee, reading, processing). Another distinction is whether the order of events or states is logged. If this is the case, the data is sequenced;

if not, it is perceived as an item set. Once the data is collected, the investigation of event sequence data can be divided into three main types: pattern mining, pattern pruning and interactive visualization design (Liu, Dev, Dontcheva, & Hoffman, 2016). In studies investigating differences in learner behaviour the focus lies on pattern mining defined as the identification of meaningful event sequences (patterns) (e.g., Azevedo et al., 2016; Bannert, Molenar, Azevedo, Järvelä, & Gašević, 2017; Bannert et al., 2015; Siadaty, Gašević, & Hatala, 2016; Sonnenberg & Bannert, 2015). The mining of patterns consists of two dimensions: ordering of events and containment. When the order of events is preserved, the pattern is a sequential pattern. Sub-sequences in this respect are parts of a sequence whose elements also appear in the same order elsewhere. In other words, sub-sequences are unique sets of ordered events shared by a threshold number of learners. Containment relates to support for a sub-sequence in the sample. Support for a sub-sequence is the number (or percentage) of sub-sequences matching other learners' sub-sequences. A frequent sub-sequence is a sub-sequence that is present in at least the threshold number of times among learners. Following the identification of sub-sequences, statistical trials can be performed to ascertain whether significant differences in the occurrence of sub-sequences can be linked to conditions internal or external to the learner.

1.5.2. Learning outcomes and self-regulated learning

Learners' learning outcomes can be conceived as the outcomes of instruction (e.g., Endedijk, Brekelmans, Verloop, Sleegers, & Vermunt, 2014). To operationalize learning outcomes the Winne and Hadwin model (1998) for self-regulated learning was used. This model proposes beliefs, dispositions, and styles; motivational factors and orientation; domain knowledge; knowledge of task; and knowledge of study tactics and strategies as variables influencing or influenced by learners' self-regulated learning. In this study we focus on three main learning outcomes (1) domain knowledge; (2) goal orientation; and (3) academic self-concept. These learning outcomes were chosen as they have each been investigated widely in terms of their relationship to self-regulated learning. In what follows, we relate each of them to research on self-regulated learning. The first, domain knowledge, relates to learners' knowledge about a task (Greene & Azevedo, 2007). As widely demonstrated in the literature on expertise, the more extensive learners' domain knowledge is, the less they need to search for, use, and regulate metacognitive strategies when grappling with complex tasks or when trying to learn information in the domain (e.g., Lesgold et al., 1988; Song, Kalet, & Plass, 2016). With regard to goal orientation, Pintrich (2000) and Eccles and Wigfield (2002) operationalized goal orientation in mastery and performance goals, along with their approach and avoidance forms. Most of the research on mastery goal orientations has focused on the approach form and has almost universally found more use of cognitive elaboration, organization strategies, and more frequent help-seeking behaviour (e.g., Duffy & Azevedo, 2015; Kitsantas, Steen, & Huie, 2017; Midgley, 2014). Little research seems to exist on the mastery-avoidance orientation. Nonetheless Wolters, Pintrich, and Karabenick (2005) and Elliot and McGregor (2001) did find an association between this orientation and test anxiety, consistent with the characterization of mastery-avoidant learners as perfectionists. With regard to the performance-approach authors argue they can lead to some productive strategy behaviour (e.g., Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002; Kitsantas et al., 2017; Mega, Ronconi, & De Beni, 2014), whereas others claim that the effects of this form are still unknown (e.g., Dompnier, Darnon, & Butera, 2013; Midgley, Kaplan, & Middleton, 2001; Senko, Durik, Patel, Lovejoy, & Valentiner, 2013). Also with regard to performance-avoidance results are ambiguous. Some research states this orientation is associated with negative outcomes, such as the use of fewer cognitive strategies (Pintrich, 2000). Nonetheless there is evidence for the increased use of cognitive strategies to test one's own abilities and comparison with others (e.g., Collazo, Elen, & Clarebout, 2015; Crippen, Biesinger, Muis, & Orgill, 2009). Finally, academic self-concept is defined as an individual's perception of self within academia (Elliot & Dweck, 2013). Academic self-concept contains two concepts. The academic competence component relates to whether learners feel that academic subjects are easy for them and whether they believe to be good at them, whilst academic effort relates to whether learners like or dislike to go to school, and like to study different subjects (Liu & Wang, 2005). A strong and positive influence of academic self-concept on the number and types of metacognitive strategies used could be observed in literature. Those with stronger academic self-concepts explore learning environments more vigorously, whereas those who have lower self-concepts retreated and concentrated on simple cognitive strategies (e.g., Kuo et al., 2014).

1.6. Problem statement and hypotheses

Literature emphasizes the importance of self-regulation for learning in blended learning environments and the role of learners' self-reflection for self-regulated learning. Nonetheless, evidence is inconclusive on the use of cues for reflection and their effect on self-regulated learning. Given the inconclusiveness guidelines for interventions are difficult to outline, hence new approaches are needed to better understand the underlying mechanisms that may explain the inconclusive results. To get more profound insights in the effect of cues for reflection, this study investigates if cues for reflection in blended learning environments designed in line with the timing, focus, and integration principles extracted from current literature on the design of cues for reflection to foster self-regulated learning affect learners' self-regulated learning through changes in their learning behaviour (operationalized through event sequences) and outcomes (operationalized through prior domain knowledge and domain knowledge, goal orientation, and academic self-concept). This leads us to three hypotheses:

- **Hypothesis 1:** "Cues for reflection will evoke more goal-directed behaviour in learners' learning behaviour (operationalized through event sequences) when they are designed in line with current insights on the design and implementation for cues for reflection."

- **Hypothesis 2:** "Cues for reflection in blended learning environments will positively affect learners' learning outcomes (operationalized in domain knowledge, goal orientation, and academic self-concept) when they are designed in line with current insights on the design and implementation for cues for reflection."

- **Hypothesis 3:** "Relating learners' learning outcomes to changes in learners' learning behaviour provides a more complete account of cues for reflection's effect on learners' self-regulated learning."

2. Method

2.1. Participants

The participants in this study were 41 adults from a centre 'second-chance' adult education in Belgium. There were 25 women (61.00%) and 16 men (39.00%). In total, 14.60% of the participants were younger than 20 years of age; 39.00% were between 20 and 30; 31.70% were between 31 and 40; and 14.60% were between 41 and 50 years of age. They were familiar with the domain of mathematics to some extent, but before the experiment they had not acquired the basic principles of statistics, which was the subject of the study task in the experiment. The subject matter was expected to be entirely new to them. This was controlled for in a prior domain knowledge test dealing with basic knowledge as presented in the study task. None of the participants was able to achieve the maximum score on the test's questions. It was concluded that the students could be divided over the experimental groups at random. All 41 adults voluntarily participated in the study, nonetheless some (different numbers for different analyses) were excluded along the way because of incomplete records.

2.2. Content and course description

2.2.1.Content

In second chance education, ‘module two - mathematics: basic statistics’ is one of the three modules of the obligatory ‘mathematics cluster’ in the ‘basis education’ track. The course is both theoretical (statistics concepts) and practical (calculating the different key figures). The course content consists of eight topics: (1) quantitative and qualitative characteristics, (2) representative surveys, (3) descriptive tables, (4) presentation of statistical data using ICT, (5) using grouped data, (6) centred measures, (7) variance measures, and (8) standard deviation. The course content was provided through Moodle.

2.2.2.Course description

The course ‘Module two - mathematics: basic statistics’ was organized in the second semester of the 2016-2017 school year. The course consisted of 62.50% (5 sessions) online instruction and 37.50% (3 sessions) of face-to-face instruction. It was started face-to-face with an introductory lesson containing general information with regard to the course and a short introduction to each of the topics. Next, the course ran for three weeks online, until the first face-to-face session (week 5). After this session there were another two weeks of online sessions. The course ended after eight weeks, at which point the instructor administered a classroom-based online exam.

2.3. The blended learning environment

2.3.1.Experimental and control environment

The experimental and control environment were identical (except for the cues for reflection). With regard to the online component of the blended learning environment a Moodle course was developed in a co-design fashion between instructor and researcher, addressing the course content. For each topic an identical outline was used. This outline consisted of following elements: the goals of the topic, an introduction (examples from practice), the course content (theory), followed by exercises, tasks, and assessments. In the experimental condition cues for reflection were added to each of the topics. For each of these elements learners interactions’ was tracked. The structure of the face to face contact moment of week 5 was similar to the online ones. It also included the mentioning of the goal of the lesson at the start, used an example as introduction, provided the learners with theory, and concluded the lesson with exercises and a formative test. The differences with the online component for this lesson was that week 5 was more focussing on all the topics addressed during the first three online weeks.

2.3.2.Cue for reflection design

The cues for reflection as integrated in the experimental learning environment were provided in two different formats, one for the online learning environment (Moodle) and one for the face to face contact moments. Although the formats (online of face to face) were different the principles used for the design were the same (see Appendix 1). Cues for reflection were provided at three different moments for each of the topics. The first cue was provided before the introduction of the theory. The second cue appeared before the learners started the exercise related to the topic. The final cue was given at the end of the topic after the completion of the assessment. The cues for reflection focused on the first two phases of self-regulation namely (1) task identification and (2) goal-setting and planning. With regard to task identification learners were asked to reflect on the content and task at hand, the information supporting the execution of the task, and the strategies needed to solve the task. In relation to goal-setting and planning learners were asked to reflect on the steps needed to complete the task at hand, how to plan for successful completion of the task, and which metacognitive strategies to use. Finally, each of the cues was embedded in the environment at the same level of the other content items. Learners were signalled that the content hidden under the link of the cue for reflection might help their learning. The operationalization of this design in the online learning environment was done by using feedback forms, including the 'megaphone' icon in Moodle. Each page looked similar. In the face to face contact moment cues for reflection were presented through whole classroom interactions between the learners and the instructor. During the co-design phase reflection cues were selected and integrated in both the online learning environment and the design of the face to face contact moments. This was done based on a list of reflection cues provided by the researcher, based on current literature (see Appendix 1).

2.4. Data structure of the log files

To investigate the suitability of using event sequence analysis insight are needed in the data structure of the log file data. Each action made by a learner within the online course was registered resulting in a time stamped event (TSE) database with as column headers the time stamp of the action, personal identifier of the user, and event name. To keep the gathered data as ecologically valid as possible a data-driven approach was chosen. In line with this approach the raw data was used to perform the event sequence analysis, no recoding or transformations took place. Both environments (experimental and control) were identical and included eight standard event names (see Table 1). Each of these event names refers to an attribute of the environment, data reported on the attributes hence refer to specific (series of) events. In the experimental condition a ninth attribute was available because of the integration of the cues for reflection, Feedback. As in the investigation of the relationship between cues for reflection and learners' behaviour the aim was to identify which sub-sequences occurred significantly more in which condition, sub-sequences including the Feedback attribute were exclude from the analysis as they only occur in the experimental condition. The feedback attribute itself was only used for the investigation of learners' cue-for-reflection use.

Table 1. *Actions traced in the online learning environment.*

Attribute	Description
Course	Landing page of the course. On this page learner found an overview of the entire course, links to each of the topics addressed, and links to the discussion forum.
File	Downloadable content pages elaborating on the topic addressed. These attributes were glossary sheets for making calculations.
Folder	Collection of extra, not mandatory materials related to the content, consisting of alternative software, examples, etc.
Forum	Discussion forum including the viewing of the forum, posting of information, and all other interactions related to the forum.
Link	External resources, related to the course and not mandatory. Containing extra examples or exercises.
Page	Organizational information about the course. Contains information about the goals, study help, etc.
Task	Exercises learners needed to perform on the different topics. Each of these tasks was obligatory and contributes to the grades learners could obtain for participation.

Test	Assessment in the different topics or at the end of the course. Obligatory and contributes to learners grades obtained for the course.
Feedback	Cues for reflection and posed as an open ended question to the learner (which were free to either answer or don't answer)

Finally, for the face to face contact moments trace data with regard to the learners was also collected via the online learning environment as the learners did use it as the main and only source for information during the face to face contact moment. Additionally, each of the contact moments was video and audio recorded to assure the rigor of the intervention. Each of the recordings was confronted with the accompanying (and agreed upon) lesson plan. If deviations would occur, the data of that lesson would be discarded. No deviations were observed, so no data needed to be discarded.

2.5. Instruments

2.5.1.Prior domain knowledge and domain knowledge

During the pre-test phase a performance based prior domain knowledge test was administered to investigate learners' prior domain knowledge. This prior domain knowledge test represented the content of the entire course consisting of fifteen questions (multiple choice and open questions). The test consisted of questions related to each of the eight topics. The test was scored on fifteen points and recalculated to a score out of 100. The same test was used as post-test to measure learners' domain knowledge.

2.5.2. Goal orientation

Learners' goal orientation was measured by using the merged version of two questionnaires of Elliot and Church (1997) and Elliot and McGregor (2001) for measuring learners' goal orientation as constructed by Lust (2012). Whereas the initial questionnaire of Elliot and Church (1997) measured solely three dimensions of goal orientation (mastery approach, performance avoidance, and performance approach), the revised questionnaire Elliot and McGregor (2001) incorporated the fourth dimension of mastery avoidance as well. These two questionnaires were merged into one that consisted out of 21 items (Mastery goal approach (MGA) (6 items), Mastery avoidance approach (MAA) (4 items), Performance goal approach (PGA) (5 items), Performance avoidance approach (PAA) (6 items)). Answers for the items were given on a 5-point Likert type scale ranging from 1 (strongly disagree) to 5 (strongly agree).

2.5.3.Academic self-concept

The Academic self-concept (ASC) scale comprised two 10-item subscales: learning confidence and learning effort. The learning confidence subscale assessed learners' feelings and perceptions about their academic competence (Liu, Wang, & Parkins, 2005). Example items included 'I am good in most of my course subjects' and 'most of my classmates are smarter than I am' (negatively worded). The learning effort subscale assessed learners' commitment to and involvement and interest in schoolwork. Example items included 'I study hard for my tests' and 'I often feel like quitting my courses' (negatively worded). Answers for the items were given on a 5-point scale ranging from 1 (strongly disagree) to 5 (strongly agree).

2.5.4.The quality of the instruments

Traditional reliability analysis were used in order to investigate the quality of the measurement instruments. The reliabilities of the different measurement instruments were measured through Cronbach's alpha. The measurement instruments are reliable when the Cronbach' s alpha is above the threshold of .70 (Cortina, 1993). Table 2 depicts the Cronbach's alpha values of the different scales. The values are all above .70, hence the measurements seem reliable.

Table 2. Pre and post r*eliability analysis per construct.*

Latent variable	Construct	α pre	α post
Cognition	Prior domain knowledge (PDK) (15 items)	.76	.83
Goal orientation	Mastery goal approach (MGA) (6 items)	.80	.73
	Mastery avoidance approach (MAA) (4 items)	.87	.85
	Performance goal approach (PGA) (5 items)	.88	.83
	Performance avoidance approach (PAA) (6 items)	.80	.80
Academic self-concept	Learning effort (LE) (10 items)	.85	.71
	Learning confidence (LC) (10 items)	.82	.75

2.6. Procedure

Learners' were randomly assigned to two separate but identical learning environments, either the control or the experimental condition. All (n=41) learners attending the course were invited to complete the online pre-test questionnaire and prior domain knowledge test before starting the module. The learners got 60 minutes to complete the questionnaire and prior domain knowledge test during the first face to face and introductory contact moment of the course. The subsequent weeks (eight in total), the learners in the experimental condition received cues for reflection (as described earlier) in the online and face to face learning environment. After the completion of the intervention, in the last face to face contact session the learners of both groups completed the online post-test questionnaire and domain knowledge test, in their classroom. The learners were given 60 minutes to complete them. The learners did not receive any other form of instruction on the course topic during the time period between the pre-test and the post-test. For the matching of the pre-test and post-test questionnaire, prior domain knowledge test and domain knowledge test, and behavioural traces learners' anonymized student IDs were used.

2.7. Analysis

2.7.1.Comparison of pre-test scores

With regard to the pre-test comparison of the experimental and control group as to prior domain knowledge, goal-orientation and academic self-concept assumptions were tested using one-way analysis of variance (ANOVA) and one-way multivariate analysis of variance (MANOVA) measures.

2.7.2.Cue-for-reflection use

For the investigation of the cue-for-reflection use we first applied descriptive statistics to identify the overall and mean use of the cues. Secondly a one-way multivariate analysis of variance (MANOVA) was used to investigate differences in cue use among learners' with different pre-test scores. As third a repeated Measures ANOVA was administered to investigate differences in timing of cue use (prior, during, or after a task) among learners' with different pre-test scores. Fourth and final an ANOVA build from the Trend Line Model was used to investigate a decrease in cue use over time.

2.7.3.Investigation of learning behaviour

The event sequence analysis consisted of two major steps (e.g., Cicchinelli et al., 2018; Zhou, 2016). First an exploratory sequence analysis was done by the identification of frequent event sub-sequences. Secondly, an explanatory approach was taken by the identification of discriminant frequent event sub-sequences. The latter analysis was based on the condition learners were in. This to identify what sub-sequences occurred significantly more in which condition. The same approach was taken based on the impacted learning outcome (see: 2.7.4). Here it was done to distinguish if the same differences in sub-sequences observed among conditions also could be observed within the significant impacted learning outcome (Hypothesis 3). The data was imported in R-statistics and analysed using the TraMineR package (Gabadinho, Ritschard, Mueller, & Studer, 2011).Frequent event sub-sequences: exploratory sequence analysis

An identical approach to identify frequent event sub-sequences was used as Jovanović, Gašević, Dawson, Pardo, and Mirriahi (2017) and Van Laer and Elen (2016). Both studies emphasize the importance of two parameters when identifying frequent event sub-sequences. The first one is the time constraint (Studer, Mueller, Ritschard, & Gabadinho, 2010). As we followed a data-driven approach while investigating the ecological order of events, we chose to set this parameter on one. This indicates that only events that actually occurred following each other will be included instead of events further apart in time. The second one is the relative threshold number of times (pMinSupport) a sub-sequence occurs among the different learners (Müller et al., 2010). In this study this parameter was arbitrarily set on .25 (25%) to assure frequent sub-sequences occurred at least in 25% of the participants.

2.7.3.2. Discriminant frequent event sub-sequences per condition and outcome

The identification of discriminant frequent event sub-sequence happened in line with Kim and Shute (2015) and with Grover et al. (2017). The significant discriminating ability of the sub-sequences was first based on differences between conditions learners were in and secondly on the impacted learning outcome. For these analysis chi-square tests were used (Studer et al., 2010). To be able to calculate the discriminating abilities of a frequent sub-sequence two arguments are needed (a) a sub-sequence (subseq) object containing the sub-sequences considered for discriminating the groups and (b) the variable that defines the groups (groups) (Garza, 2016). The former was defined using the method described in 2.7.3.1 and the latter based first on the condition and secondly on the outcomes of the statistical trails on significant changes in learning outcomes. A Chi-square test is used to investigate the significance of the relationship between the observed and expected occurrence of a frequent sub-sequence for each value of the measured variables. Finally, the effect sizes are calculated based on the Cramer's V. The Cramer's V expresses the relationship between a certain discriminating frequent sub-sequence and the learners' characteristics and is reported in a value between zero and one. The closer to one the higher the relation. Cohen (1988) refers to small (≤.30), medium (≥.30 and ≤50), and large (≥.50) effect sizes.

2.7.4.Investigation of learning outcomes

In order to examine the effect of the instructional intervention on learners' learning outcomes, a 2 (groups: experimental and control) × 2 (testing time: pre-test and post-test) mixed design ANOVA was conducted. Before conducting this test the variables were tested for normality (Shapiro–Wilks' test), sphericity (Mauchly's Test of Sphericity), and homogeneity of variances (Levene's test). The main advantage of a mixed ANOVA design in this ecologically valid study is that unlike the traditional repeated measures approaches that discard all results on any subject with even a single missing measurement, mixed versions allow other data of such subjects to be used as long as the missing data meets the so-called missing-at-random definition (Seltman, 2012).

3. Results

3.1. Pre-test comparison of the experimental and control group

Based on initial comparison of pre-test measurements for each variable, no significant differences between the pre-test scores of the experimental and control group could be found (see Table 3) (respectively prior domain knowledge test (ANOVA) (F (1, 23) = 2.881, *p* = .10), goal orientation and academic self-concept questionnaire (MANOVA) (F (6, 29) = 1.272, *p* = .30))). Of the 41 learners, 23 learners (control (n=13) and experimental (n=10)) completed the prior domain knowledge test. The goal orientation and academic self-concept questionnaire was completed by 36 (control (n=18) and experimental (n=18)).

Table 3. *Pre-test comparison experimental and control group.*

Latent variables	Construct	df	F	Sig.	Ctrl. M	Ctrl. SD	Exp. M	Exp. SD
Cognition	*PDK*	1, 23	2.88	.10	12.73	6.42	17.26	7.32
Goal orientation	*MGA*	1, 36	2.28	.41	4.07	.14	4.23	.14
	MAA	1, 36	.47	.83	3.26	.24	3.33	.24
	PGA	1, 36	.01	.92	2.75	.31	2.80	.31
	PAA	1, 36	.60	.45	3.32	.24	3.06	.24
Academic self-concept	*LE*	1, 36	1.13	.30	1.85	.18	2.11	.18
	LC	1, 36	.01	.92	2.58	.20	2.61	.20

3.2. Cues for reflection use

To be able to assign changes to either learners' learning outcomes or learning behaviour we need to be sure that learners were actually exposed to the cues for reflection. Descriptive statistics showed that on average learners consulted sixteen of twenty-four cues for reflection. Individual learners consulted on average each cue three times. Repeated Measures ANOVA did indicate there was no significant difference among learners with regard to timing of individual cue use (prior, during, or after a task) (F (2, 8) = 2.168, *p* = .19). Nor were significant differences found through an ANOVA build from the Trend Line Model for a decreased use of the cues over time (F(2, 16) = 1.76, *p* = .14) or based on a one-way MANOVA based on differences in learners' characteristics (F(2, 16) = .51, *p* = .81).

3.3. Hypothesis 1: Cues for reflection affect learners' learning behaviour

3.3.1. Frequency of event occurrence

With regard to the frequency of use of the attributes available in the online learning environment statistically significant differences (F (9, 20) = 6.130, p < .001, Λ = .450, η_p^2 = .73) were found between the control and the experimental condition as determined by a one-way multivariate analysis of variance (MANOVA) (see Table 4). Of the 41 learners overall, only 28 logged in to the online learning environment and so were included for further analysis (control (n=14) and experimental (n=14)). Only for the Folder attribute no significant differences were retrieved (p = .73).

Table 4. *Frequency and significance of attribute occurrence for the experimental and control condition.*

Attribute	Ctrl. M	Ctrl. SD	Exp. M	Exp. SD	df	F	Sig.	η_p^2
Course	65.00	65.73	180.00	152.34	1, 28	8.485	.007	.23
File	35.95	42.86	124.80	125.11	1, 28	8.385	.007	.23
Folder	-	-	-	-	-	-	.730	-
Forum	30.65	50.76	93.70	103.32	1, 28	5.116	.032	.15
Link	1.50	1.79	4.50	3.03	1, 28	11.707	.002	.30
Page	3.80	4.81	23.20	25.37	1, 28	11.272	.002	.29
Task	10.70	12.47	28.60	21.31	1, 28	9.493	.007	.23
Test	52.10	51.48	108.30	69.96	1, 28	6.245	.019	.18
Overall	201.20	214.94	564.30	470.142	1, 28	8.584	.007	.24

3.3.2.Frequency of significant discriminant sub-sequence occurrence (between conditions)

The learners (n=28) included in the event sequence analysis generated 10163 events over the timespan of eight weeks. A total of 688 frequent sub-sequences were extracted. First we will describe the nature of the discriminant sub-sequences, next we will discuss the differences between the experimental and the control condition. Based on the assignment of the learners to either the control or the experimental condition (independent variable) 80 significant discriminant sub-sequences were identified (see Appendix 2). The analysis of the 80 significant discriminant sub-sequences showed that a sub-sequence contained between two and thirteen events. The occurrence of each of the attributes in the 80 significant discriminant sub-sequences was: Course (83%), File (53%), Test (49%), Forum (18%), Task (15%), Page (6%), and Link (4%). No Folder attributes occurred. Course events were mostly followed by Task (33%) and File (32%) events, whereas Task and File events were often preceded by Forum events (18%) or the consultation of other Task (15%), or File (18%) events. Finally, with regard to Test events, these were most often proceeded by the consultation of other Test (39%) and followed by Course (23%) events. Based on this analysis we categorized the significant discriminant sub-sequences in three main categories.

The first category relates to significant discriminant sub-sequences involving File and Task events which in many cases (see description online learning environment) consisted of assignments and tasks. Results show that the sub-sequence containing File events occurred significantly (between $\chi(1) = 3.906$, p = .048, V= .36 and $\chi(1) = 11.271$, p < .001, V= .61) more (standardized residuals between 1.13 and 2.65) in the experimental, compared to the control condition. Examination of the Cramer's V scores indicate according to Cohen (1988) effects from medium (≥.30 and ≤50) to large (≥.50). Similar findings were found in relation to Task events with Chi squared test (between $\chi(1) = 3.906$, $p = .048$, $V = .36$ and $\chi(1) = 8.750$, $p = .003$, $V = .54$), effect size (≥.30 and ≤50), and standardized residuals (between 1.23 and 2.31). The second category relates to Forum events, focussing on communication. Results show that the sub-sequence containing Forum events occurred significantly (between $\chi(1) = 3.906$, $p = .048$, $V = .36$ and $\chi(1) = 6.158$, $p < .013$, $V = .45$) more (standardized residuals between 1.50 and 2.04) in the experimental, compared to the control condition. Examination of the Cramer's V score indicate medium (≥.30 and ≤50) effects (Cohen, 1988). The third and final category relates to Test events, focussing on formal assessment. Results show that sub-sequences containing Test events occurred significantly (between $\chi(1) = 3.906$, $p = .048$, $V = .36$ and $\chi(1) = 8.856$, $p < .003$, $V = .54$) more (standardized residuals between 1.23 and 2.01) in the experimental, compared to the control condition. Examination of the Cramer's V score indicate medium (≥.30 and ≤50) effects (Cohen, 1988).

3.4. Hypothesis 2: Cues for reflection positively affect learners' learning outcomes.

3.4.1. Effect of time on learners' learning outcomes

The results of the mixed design ANOVA revealed that both the control group and the experimental group demonstrated a statistically significant increase ($MD = 41.29$; $SE = 6.04$, $p < .001$) in mean domain-knowledge (DK) scores across the two time points ($F(1, 8) = 46.716$, $p < .001$, $\eta_p^2 = .85$). The effect size value suggested a large practical significance (Cohen, 1988). Also the mean score for goal orientation and more specifically performance goal approach (PGA), was significantly impacted over time ($F(1, 8) = 6.564$, $p = .034$, $\eta_p^2 = .45$). A significant decrease ($MD = -.64$; $SE = .25$, $p = .034$) in mean score was found. Finally, the mean score for learning confidence (LC) within learners' academic self-concept was impacted significantly ($F(1, 8) = 7.498$, $p = .026$, $\eta_p^2 = .48$). A significant increase ($MD = 1.06$; $SE = .39$, $p = .026$) in mean score on confidence was found. No other significant effects of time were found.

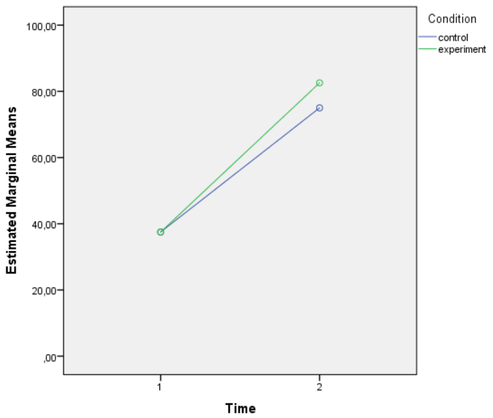

Figure 1: Differences in means of domain knowledge

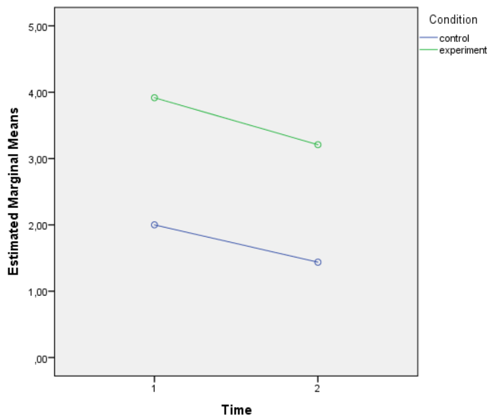

Figure 2: Differences in means of performance goal approach

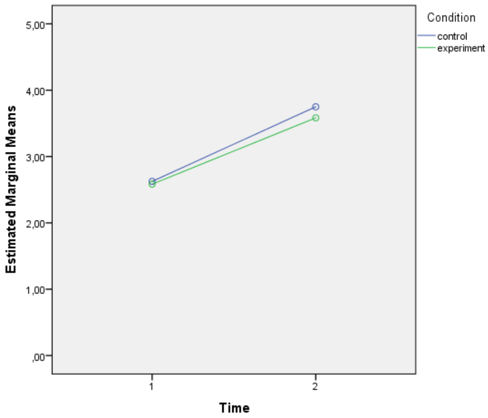

Figure 3: Differences in means of learning confidence

3.4.2.Effect of condition on learners' learning outcomes

The results of the mixed design ANOVA also revealed that for the between subject analysis the experimental group had only a statistically significant different (F (1, 8) = 26.396, $p < .001$, $\eta_p^2 = .77$) mean score (higher) for performance goal approach (PGA) (*MD* = 1.84; *SE* = .36, $p < .001$).

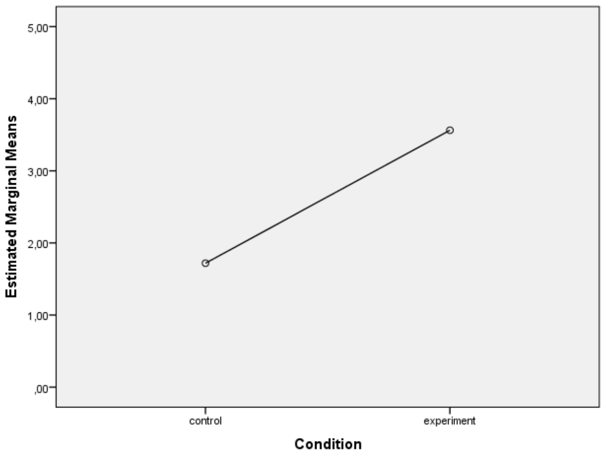

Figure 4: Differences in means of PGA

3.4.3.Effect of time and condition on learners' learning outcomes

With regard to the interaction effect of time and condition on learners' learning outcomes there was only one significant interaction effect between the intervention type (experimental - control) and the testing time (pre-test - post-test) found for learners' performance avoidance approach (PAA), F (1, 8) = 7.374, $p = .026$, $\eta_p^2 = .48$. The effect size value suggested a large practical significance (Cohen, 1988). Where the control group (*MD* = -.68) decreased, the experimental group (*MD* = .50) increased.

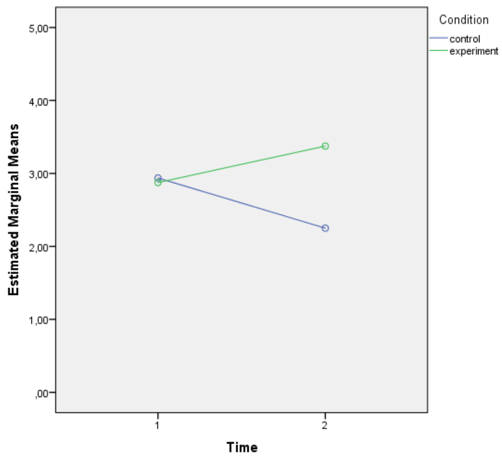

Figure 5: Differences in means of PAA

3.5. Hypothesis 3: There is a relationship between learning behaviour and learning outcomes

To investigate if the differences in significant discriminant sub-sequences between conditions are related to differences in performance avoidance approach (PAA) (low (1) – high (5)), the same analysis was done as when addressing Hypothesis 1. Here, too, we describe the nature of the discriminant sub-sequences followed by a discussion on the differences between learners with different performance avoidance approach scores. A total of 30 significant discriminant sub-sequences were identified (see Appendix 3). The analysis of the 30 significant discriminant sub-sequences showed that sub-sequences contained between two and seventeen events. The occurrence of each of the attributes in the 80 significant discriminant sub-sequences was: Course (73%), File (73%), Test (67%), Forum (7%), Task (27%), Page (13%), and Link (13%). No Folder attributes occurred. Also here Course events were most often followed by Task (28%) and File (28%) events, whereas Task and File events were often preceded by Forum events (22%) or the consultation of other Task (28%), or File (61%) events. Finally, with regard to Test events, these were most often proceeded by the consultation of other Test (33%) and followed by Course (72%) events. The same three categories as when analysing the impact of the condition were used. Below, Table 5 provides a detailed account of the key figures for each category.

Table 5. *Frequency of significant discriminant sub-sequence occurrence (PAA).*

Attribute in sub-sequence	Summary
File	Occurred significantly (between χ(4) = 9.574, *p* = .048, *V*= .73 and χ(4) = 13.846, *p* < .008, *V*= .88) more when the score of PAA increased from low (1) (standardized residuals = -0.52) to high (5) (standardized residuals = -1.37). Cramer's V score indicate large (≥.50) effects.
Task	Similar findings were found for: Chi squared test (between χ(4) = 9.574, *p* = .048, *V*= .72 and χ(4) = 13.041, *p* = .011, *V*= .76), effect size (between ≥.50) and standardized residuals (between -0.52 and 2.37).
Forum	Occurred significantly (χ(4) = 10.286, *p* = .036, *V*= .76) more when the score of PAA increased from low (1) (standardized residuals = -0.47) to high (4) (standardized residuals = -2.31). Cramer's V score indicate large (≥.50) effects.
Test	Occurred significantly (between χ(4) = 10.393, *p* = .034, *V*= .76 and χ(4) = 15.195, *p* < .004, *V*= .92) more when the score of PAA increased from low (1) (standardized residuals = -0.58) to high (4)

(standardized residuals = -1.15). Cramer's V score indicate large (≥.50) effects.

4. Discussion

In this section, we first summarize the results obtained. After analysing the data it became clear that learners in the experimental and control conditions did not differ significantly from each other at the start of the eight week intervention, that all learners in the experimental condition were exposed to the cues for reflection, and that no significant differences in use among these learners could be observed.

With regard to Hypothesis 1, the results showed that learners in the experimental condition made significantly more use of sub-sequences consisting of File and Task attributes related to assignments and tasks, Forum attributes related to communication, and Test attributes referring to assessment. In relation to Hypothesis 2, the results of post-tests for both conditions showed a significant increase in domain knowledge and learning confidence and a decrease in performance goal approach. Learners in the experimental condition who received cues for reflection scored significantly higher on performance goal approach compared to the learners in the control condition. As for the interaction effect between time and condition, learners in the experimental condition scored significantly higher on performance avoidance approach (PAA) compared to the controls. Finally, as regards Hypothesis 3, the results showed that when differences in learners' behaviour per condition were compared to differences in learners' behaviour per PAA score, the same differences in behaviour were found. This suggests that a change in learning behaviour might be linked to learning outcomes.

4.1. Findings

A change in learners' self-regulated learning was only observed through learners' behaviour related to self-testing and monitoring others via the use of Forum attributes and via performance avoidance approach (PAA) goal orientation. This appears to be in line with Crippen et al. (2009), whose results indicate that cues for reflection prompted learners to test their own performance against others and thus attempt to reduce the anxiety associated with a potential failure (perhaps triggered by the question to reflect upon their own behaviour). As a result, learners might seek (1) information on others' performance through the consultation of discussion forums and (2) to demonstrate that they are not doing worse than others by using tests (Collazo et al., 2015). In line with this, our findings seem to indicate that changes in learners' learning behaviour can be linked to changes in learners' learning outcomes. Subsequently, this study shows that cues for reflection designed according to the timing, focus, and integration guidelines extracted from current literature affected learners' self-regulated learning in an unexpected manner.

4.2. Exploration of unexpected findings

Despite the wide range of studies that have indicated the importance of reflection for self-regulated learning and increased learning outcomes (e.g., Johnson, Azevedo, & D'Mello, 2011; Kramarski & Gutman, 2006), only a limited number of empirical studies seem to be able to demonstrate positive results. This is especially true when it comes to cues for reflection intended to evoke self-regulated learning through self-reflection. The findings of the study presented here add to the inconclusive nature of the investigation of the effect of cues for reflection on learners' self-regulated learning. Although the reason for these unexpected findings is unclear, the literature suggests two sets of possible explanations: firstly, the influence of learners' characteristics (e.g., Bannert & Reimann, 2012) and the design of the cues for reflection (e.g., Bannert, 2009), and secondly, the additional cognitive and metacognitive capacities needed to act upon the cues presented (Veenman, 1993).

4.2.1.Learners' characteristics and the design of cues for reflection

In the introduction to this manuscript we addressed cues' timing (e.g., Farrall, 2007; Mann et al., 2009), focus (e.g., Sonnenberg & Bannert, 2015; Winne & Hadwin, 1998), and integration (e.g., Aleven et al., 2003; Greene & Azevedo, 2007) as potential mechanisms to overcome issues related to no or sub-optimal cue use. Yet, there is no guarantee that learners (1) make use of the support and/or (2) use the cues as intended (Land & Greene, 2000). With regard to the former, the results of the study presented in this manuscript show that a mismatch between learners' characteristics and the design of the cues for reflection seems highly unlikely, as the results of the description of learners' cue use indicated that learners (with different learner characteristics) in the experimental condition did not differ significantly in their frequency or temporal use of the cues for reflection. With regard to the use of cues as intended, the results show that the intervention evoked self-regulated learning observed through changes in learners' learning behaviour and learning outcomes. The combination of these findings strengthens our hypothesis that in the study presented cues for reflection were sufficiently well designed to evoke self-regulated learning, but that they also triggered unintended behaviour and outcomes.

4.2.2.Additional cognitive and metacognitive capacities needed

In line with our findings, van den Boom et al. (2004) found no effect of metacognitive strategy use when only cues for reflection were provided to learners with low self-regulated learning capabilities. However, in conditions where cues for reflection were combined with different forms of feedback, a significant increase in metacognitive strategy use was observed. Similar findings were presented by Krause and Stark (2010), who demonstrated that feedback interventions alone clearly enhanced learning outcomes, whereas conditions including cues for reflection had no significant effect on learning. The meta-analysis of Ardasheva, Wang, Adesope, and Valentine (2017) reveals that the overall effects of cues with regard to the use of specific cognitive and metacognitive strategies were large (.87) and that this is specifically the case when interventions adopt a reflection-oriented model targeting metacognitive strategies (Dabarera, Renandya, & Zhang, 2014; Takallou, 2011; Vandergrift & Tafaghodtari, 2010). In the light of self-regulated learning theory (see: Puustinen & Pulkkinen, 2001) the observations described above might be explained as follows. Even when learners can reflect on cognitive (Reiser et al., 2001) and metacognitive (Kori et al., 2014) strategies, they might not possess or be able to recall the metacognitive strategies needed to act in a way that will produce successful learning outcomes (e.g, Pintrich, 2002; Veenman, 1993). This will result in sub-optimal self-regulated learning. Based on this notion, we hypothesize that when cues for reflection are supported by cues directing learners towards appropriate metacognitive strategies, learners are more likely to evolve towards the behaviour needed to affect learning outcomes positively than when they only receive cues for reflection.

4.3. Summary

In line with the reasoning that not all learners use and perceive cues as intended by the instructor, the findings of the study presented show that the integration of cues for reflection in a blended learning environment did change learners' self-regulated learning. However, we only observed a change in self-regulated learning when it came to learners' behaviour in relation to self-testing and monitoring others and in their performance avoidance approach (PAA) goal orientation. To explain this finding, we hypothesized that when differences in learners' characteristics are taken into account when designing cues for reflection, learners' self-reflection will be evoked and self-regulated learning will take place. Yet for this self-reflection and self-regulated learning to be tuned to instructional expectations (i.e., design of the learning environment and learning goals), learners need to be guided to use metacognitive strategies that contribute to their performing the task at hand successfully.

5. Further directions and conclusions

In our view, the merit of the study presented lies in its fine-grained insights into the relationship between learners' self-regulated learning and cues for reflection. To obtain these insights, we first investigated learners' reflection-cue use based on learners' individual differences. Secondly, we operationalized self-regulated learning through learners' learning behaviour and outcomes. Although the use of online event measurements are nothing new, the link to learners' learning outcomes is not often made, neglecting the essence of self-regulation in an educational context. Investigating both learning behaviour and outcomes provides insights not only into learners' self-regulated learning, but also into the nature of cues' effects. In the study presented, we found that unexpected learning outcomes were related to behavioural indications, thus establishing a link between learners' self-regulated learning and cues for reflection. Finally, in the discussion of the results we unravelled the effect of the design and content of the cues for reflection provided and hypothesized that when differences in learners' characteristics are taken into account when designing cues for reflection, learners' self-reflection will be evoked and self-regulated learning will take place. For this self-reflection and self-regulated learning to be effective, however, learners subsequently need to be guided towards metacognitive strategies that will help them perform the task successfully.

5.1. Further directions

To enable us to build further on the theoretical and methodological insights of this study, some challenges need to be addressed. A first challenge to overcome is the sample-size fluctuation in our ecologically valid setting. A total of 41 learners were involved in the study – 20 in the control group and 21 in the experimental condition, which according to Field (2013) is an appropriate rule of thumb for testing the effect of a single condition. Nonetheless, there was a substantial amount of random missing data related to both the pre- and post-tests. Although the mixed analysis of variance did not include case-wise deletion, meaning the effect of missing values was minimized, the power of some of the statistics might be debatable. A second challenge relates to the use of a data-driven approach to analyse learners' learning behaviour and its' arbitrary parameter setting. As theoretical insights can be derived from the results of data-driven trails, contributing to such an approach may prove more promising than, for example, recoding events as (covert) metacognitive strategies or activities. In further research, this data-driven approach might be explored by experimenting with different parameter settings or using a combination of data-driven and theory-driven approaches. With regard to the latter, this could be achieved for example by recoding events or sequences based on a theoretical framework unrelated to self-regulation theory (for example a tool-use scheme), which would make the sub-sequences identified more meaningful (at least when it comes to explaining them).

5.2. Conclusions

The current lack of certainty regarding the effects of cues for reflection on learners' self-regulated learning means teachers and instructional designers remain dependent on inconsistent conceptual claims that cues for reflection may improve self-regulated learning. Studies such as the one presented in this manuscript could help both researchers and practitioners distinguish between the effect of cues for reflection, how different learners use them, and how learners react to them, resulting in particular behaviours and outcomes. Establishing more fine-grained links between learners' cue use, learning behaviour and learning outcomes could help us propel the investigation of intervention research.

6. Acknowledgements

Deleted for review.

7. Bibliography

Abbott, A. (1995). Sequence analysis: new methods for old ideas. *Annual review of sociology, 21*(1), 93-113. doi: DOI 10.1146/annurev.so.21.080195.000521

Aleven, V., Stahl, E., Schworm, S., Fischer, F., & Wallace, R. (2003). Help seeking and help design in interactive learning environments. *Review of Educational Research, 73*(3), 277-320. doi: Doi 10.3102/00346543073003277

Allen, I. E., Seaman, J., & Garrett, R. (2007). *Blending in: The extent and promise of blended education in the United States*: ERIC.

Ardasheva, Y., Wang, Z., Adesope, O. O., & Valentine, J. C. (2017). Exploring Effectiveness and Moderators of Language Learning Strategy Instruction on Second Language and Self-Regulated Learning Outcomes. *Review of Educational Research, 87*(3), 544-582. doi: 10.3102/0034654316689135

Arrastia-Chisholm, M. C., Torres, K. M., & Tackett, S. (2017). Using Reflection to Increase Self-Regulation Among Pre-Service Teachers. *Fostering Reflective Teaching Practice in Pre-Service Education*, 148.

Azevedo, R., Martin, S. A., Taub, M., Mudrick, N. V., Millar, G. C., & Grafsgaard, J. F. (2016). *Are Pedagogical Agents' External Regulation Effective in Fostering Learning with Intelligent Tutoring Systems?* Paper presented at the International Conference on Intelligent Tutoring Systems.

Bandura, A. (1989). Human agency in social cognitive theory. *American psychologist, 44*(9), 1175.

Bannert, M. (2006). Effects of reflection prompts when learning with hypermedia. *Journal of Educational Computing Research, 35*(4), 359-375.

Bannert, M. (2009). Promoting self-regulated learning through prompts. *Zeitschrift für Pädagogische Psychologie, 23*(2), 139-145. doi: 10.1024/1010-0652.23.2.139

Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts *International handbook of metacognition and learning technologies* (pp. 171-186): Springer.

Bannert, M., Molenaar, I., Azevedo, R., Järvelä, S., & Gašević, D. (2017). *Relevance of learning analytics to measure and support students' learning in adaptive educational technologies.* Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference.

Bannert, M., Molenar, I., Azevedo, R., Järvelä, S., & Gašević, D. (2017). *Relevance of learning analytics to measure and support students' learning in adaptive educational technologies.* Paper presented at the LAK.

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*(1), 193-211. doi: 10.1007/s11251-011-9167-4

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short-and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293-306.

Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International Journal of Educational Research, 31*(6), 445–457. doi: http://dx.doi.org/10.1016/S0883-0355(99)00014-2
Boelens, R., Van Laer, S., De Wever, B., & Elen, J. (2015). Blended learning in adult education: towards a definition of blended learning. Brussels.
Bonk, C. (2017). *Best Practices for Online and Blended Learning: Introducing the R2D2 and TEC-VARIETY Models*. Paper presented at the Fall Faculty Development Conference, Indianapolis.
Borkowski, J. G., Carr, M., Rellinger, E., & Pressley, M. (1990). Self-regulated cognition: Interdependence of metacognition, attributions, and self-esteem. *Dimensions of thinking and cognitive instruction, 1*, 53-92.
Burke, M. J., Scheuer, M. L., & Meredith, R. J. (2007). A dialogical approach to skill development: The case of safety skills. *Human Resource Management Review, 17*(2), 235-250.
Butler, D. L. (1998). The Strategic Content Learning approach to promoting self-regulated learning: A report of three studies. *Journal of Educational Psychology, 90*(4), 682-697. doi: 10.1037/0022-0663.90.4.682
Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research, 65*(3), 245-281. doi: http://dx.doi.org/10.2307/1170684
Chen, N.-S., Wei, C.-W., Wu, K.-T., & Uden, L. (2009). Effects of high level prompts and peer assessment on online learners' reflection levels. *Computers & Education, 52*(2), 283-291.
Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams.
Clarebout, G., & Elen, J. (2009). The complexity of tool use in computer-based learning environments. *Instructional Science, 37*(5), 475. doi: 10.1007/s11251-008-9068-3
Cleary, T. J., Callan, G. L., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalytic protocols. *Education Research International, 2012*. doi: 10.1155/2012/428639
Cohen, J. (1988). Statistical power analysis for the behavioral sciences . Hilsdale. *NJ: Lawrence Earlbaum Associates, 2*.
Collazo, N. A. J., Elen, J., & Clarebout, G. (2015). The multiple effects of combined tools in computer-based learning environments. *Computers in Human Behavior, 51*, 82-95. doi: 10.1016/j.chb.2015.04.050
Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology, 78*(1), 98.
Crippen, K. J., Biesinger, K. D., Muis, K. R., & Orgill, M. (2009). The role of goal orientation and self-efficacy in learning from web-based worked examples. *Journal of Interactive Learning Research, 20*(4), 385.
Dabarera, C., Renandya, W. A., & Zhang, L. J. (2014). The impact of metacognitive scaffolding and monitoring on reading comprehension. *System, 42*, 462-473. doi: 10.1016/j.system.2013.12.020
Daniela, P. (2015). The relationship between self-regulation, motivation and performance at secondary school students. *Procedia-Social and Behavioral Sciences, 191*, 2549-2553. doi: 10.1016/j.sbspro.2015.04.410
Davis, E. A. (2000). Scaffolding students' knowledge integration: Prompts for reflection in KIE. *International Journal of Science Education, 22*(8), 819-837. doi: Doi 10.1080/095006900412293
Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *The Journal of the Learning Sciences, 12*(1), 91-142. doi: Doi 10.1207/S15327809jls1201_4
Deschacht, N., & Goeman, K. (2015). The effect of blended learning on course persistence and performance of adult learners: A difference-in-differences analysis. *Computers & Education, 87*, 83-89. doi: 10.1016/j.compedu.2015.03.020

Dompnier, B., Darnon, C., & Butera, F. (2013). When performance-approach goals predict academic achievement and when they do not: A social value approach. *British Journal of Social Psychology, 52*(3), 587-596.
Drysdale, J. S., Graham, C. R., Spring, K. J., & Halverson, L. R. (2013). An analysis of research trends in dissertations and theses studying blended learning. *The Internet and Higher Education, 17*, 90-100. doi: 10.1016/j.iheduc.2012.11.003
Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior, 52*, 338-348.
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology, 53*(1), 109-132. doi: 10.1146/annurev.psych.53.100901.135153
Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *J Pers Soc Psychol, 72*(1), 218. doi: Doi 10.1037//0022-3514.72.1.218
Elliot, A. J., & Dweck, C. S. (2013). *Handbook of competence and motivation*: Guilford Publications.
Elliot, A. J., & McGregor, H. A. (2001). A 2× 2 achievement goal framework. *J Pers Soc Psychol, 80*(3), 501.
Embo, M., Driessen, E., Valcke, M., & Van Der Vleuten, C. P. (2014). Scaffolding reflective learning in clinical practice: a comparison of two types of reflective activities. *Medical Teacher, 36*(7), 602-607. doi: 10.3109/0142159X.2014.899686
Endedijk, M. D., Brekelmans, M., Verloop, N., Sleegers, P. J., & Vermunt, J. D. (2014). Individual differences in student teachers' self-regulated learning: An examination of regulation configurations in relation to conceptions of learning to teach. *Learning and Individual Differences, 30*, 155-162. doi: 10.1016/j.lindif.2013.12.005
Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International Journal of Educational Research, 41*(6), 407-428.
Farrall, S. (2007). Desistance studies vs. cognitive-behavioural therapies: Which offers most hope for the long term. *Dictionary of Probation and Offender Management. Cullompton: Willan Publishing, 178*.
Feyzi-Behnagh, R., Azevedo, R., Legowski, E., Reitmeyer, K., Tseytlin, E., & Crowley, R. S. (2014). Metacognitive scaffolds improve self-judgments of accuracy in a medical intelligent tutoring system. *Instructional Science, 42*(2), 159-181. doi: 10.1007/s11251-013-9275-4
Field, A. (2013). *Discovering statistics using IBM SPSS statistics*: Sage.
Furberg, A. (2009). Socio-cultural aspects of prompting student reflection in Web-based inquiry learning environments. *Journal of Computer Assisted Learning, 25*(4), 397-409.
Gabadinho, A., Ritschard, G., Mueller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1-37.
Garza, M. M. M. (2016). *Coevolution of Theory and Data Analytics of Digital Game-Based Learning.* Vanderbilt University.
Gašević, D., Dawson, S., & Siemens, G. (2015). Let's not forget: Learning analytics are about learning. *TechTrends, 59*(1), 64-71.
Graham, C. R., Henrie, C. R., & Gibbons, A. S. (2014). Developing models and theory for blended learning research. *Blended learning: Research perspectives, 2*, 13-33.
Greene, J. A., & Azevedo, R. (2007). A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Review of Educational Research, 77*(3), 334-372. doi: http://dx.doi.org/10.3102/003465430303953
Grover, S., Basu, S., Bienkowski, M., Eagle, M., Diana, N., & Stamper, J. (2017). A Framework for Using Hypothesis-Driven Approaches to Support Data-Driven Learning Analytics in Measuring Computational Thinking in Block-Based Programming Environments. *ACM Transactions on Computing Education (TOCE), 17*(3), 14. doi: 10.1145/3105910

Halverson, L. R., Graham, C. R., Spring, K. J., Drysdale, J. S., & Henrie, C. R. (2014). A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *The Internet and Higher Education, 20*, 20-34. doi: 10.1016/j.iheduc.2013.09.004

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology, 94*(3), 638-645. doi: 10.1037//0022-0663.94.3.638

Hatton, N., & Smith, D. (1995). Reflection in teacher education: Towards definition and implementation. *Teaching and Teacher Education, 11*(1), 33-49. doi: Doi 10.1016/0742-051x(94)00012-U

Järvelä, S., Järvenoja, H., Malmberg, J., Isohätälä, J., & Sobocinski, M. (2016). How do types of interaction and phases of self-regulated learning set a stage for collaborative engagement? *Learning and Instruction, 43*, 39-51.

Johnson, A. M., Azevedo, R., & D'Mello, S. K. (2011). The temporal and dynamic nature of self-regulatory processes during independent and externally assisted hypermedia learning. *Cognition and Instruction, 29*(4), 471-504. doi: 10.1080/07370008.2011.610244

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The Internet and Higher Education, 33*, 74-85. doi: 10.1016/j.iheduc.2017.02.001

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education, 87*, 340-356. doi: 10.1016/j.compedu.2015.07.009

Kitsantas, A., Steen, S., & Huie, F. (2017). The role of self-regulated strategies and goal orientation in predicting achievement of elementary school children. *International Electronic Journal of Elementary Education, 2*(1), 65-81.

Kizilcec, R. F., Perez-Sanagustin, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18-33. doi: 10.1016/j.compedu.2016.10.001

Kori, K., Pedaste, M., Leijen, Ä., & Mäeots, M. (2014). Supporting reflection in technology-enhanced learning. *Educational Research Review, 11*, 45-55.

Kramarski, B., & Gutman, M. (2006). How can self-regulated learning be supported in mathematical E-learning environments? *Journal of Computer Assisted Learning, 22*(1), 24-33.

Krause, U.-M., & Stark, R. (2010). Reflection in example-and problem-based learning: Effects of reflection prompts, feedback and cooperative learning. *Evaluation & research in education, 23*(4), 255-272.

Kuo, Y.-C., Walker, A. E., Schroder, K. E., & Belland, B. R. (2014). Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *The Internet and Higher Education, 20*, 35-50. doi: 10.1016/j.iheduc.2013.10.001

Land, S. M., & Greene, B. A. (2000). Project-based learning with the World Wide Web: A qualitative study of resource integration. *Educational Technology Research and Development, 48*(1), 45-66. doi: Doi 10.1007/Bf02313485

Lavoué, É., Molinari, G., Prié, Y., & Khezami, S. (2015). Reflection-in-action markers for reflection-on-action in Computer-Supported Collaborative Learning settings. *Computers & Education, 88*, 129-142. doi: 10.1016/j.compedu.2015.05.001

Lehmann, T., Hähnlein, I., & Ifenthaler, D. (2014). Cognitive, metacognitive and motivational perspectives on preflection in self-regulated online learning. *Computers in Human Behavior, 32*, 313-323.

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures.

Lin, B., Coburn, S. S., & Eisenberg, N. (2016). Self-regulation and reading achievement. *The cognitive development of reading and reading comprehension*, 67-86.

Liu, W., Wang, C., & Parkins, E. (2005). A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British Journal of Educational Psychology, 75*(4), 567-586. doi: 10.1348/000709905X42239

Liu, W. C., & Wang, C. K. J. (2005). Academic self-concept: A cross-sectional study of grade and gender differences in a Singapore secondary school. *Asia Pacific Education Review, 6*(1), 20-27. doi: 10.1007/bf03024964

Liu, Z., Dev, H., Dontcheva, M., & Hoffman, M. (2016). Mining, Pruning and Visualizing Frequent Patterns for Temporal Event Sequence Analysis.

Lord, A. Y., Chen, M.-P., Cheng, Y.-Y., Tai, K.-C., & Pan, W.-H. (2017). Enhancing nutrition-majored students' reflective judgment through online collective reflection. *Computers & Education, 114*, 298-308.

Lust, G. (2012). Opening the Black Box. Students' Tool-use within a Technology-Enhanced learning environment: An Ecological-Valid Approach.

Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., & Clarebout, G. (2011). Tool-use in a blended undergraduate course: In Search of user profiles. *Computers & Education, 57*(3), 2135-2144. doi: 10.1016/j.compedu.2011.05.010

Lynch, R., & Dembo, M. (2004). The relationship between self-regulation and online learning in a blended learning context. *The International Review of Research in Open and Distance Learning, 5*.

Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. *Advances in Health Sciences Education, 14*(4), 595-621. doi: 10.1007/s10459-007-9090-2

McCutcheon, K., Lohan, M., Traynor, M., & Martin, D. (2015). A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate nurse education. *Journal of Advanced Nursing, 71*(2), 255-270. doi: 10.1111/jan.12509

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of Educational Psychology, 106*(1), 121.

Mezirow, J. (2000). *Learning as Transformation: Critical Perspectives on a Theory in Progress. The Jossey-Bass Higher and Adult Education Series*: ERIC.

Midgley, C. (2014). *Goals, goal structures, and patterns of adaptive learning*: Routledge.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of Educational Psychology, 93*(1), 77.

Müller, N. S., Studer, M., Gabadinho, A., & Ritschard, G. (2010). *Analyse de séquences d'événements avec TraMineR.* Paper presented at the EGC.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Oliver, M., & Trigwell, K. (2005). Can 'blended learning'be redeemed. *E-learning, 2*(1), 17–26. doi: 10.2304/elea.2005.2.1.17

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception, 11*, 239-266.

Pannese, L., & Morosini, D. (2014). Serious games to support reflection in the healthcare sector. *Int. J. Serious Games, 1*(3), 5-14. doi: 10.17083/ijsg.v1i3.30

Picciano, A. G., Seaman, J., Shea, P., & Swan, K. (2012). Examining the extent and nature of online learning in American K-12 education: The research initiatives of the Alfred P. Sloan Foundation. *The Internet and Higher Education, 15*(2), 127-135. doi: 10.1016/j.iheduc.2011.07.004

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice, 41*(4), 219-+. doi: 10.1207/s15430421tip4104_3

Puustinen, M., & Pulkkinen, L. (2001). Models of Self-regulated Learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269-286. doi: 10.1080/00313830120074206

Rashid, T., & Asghar, H. M. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers in Human Behavior, 63*, 604-612. doi: 10.1016/j.chb.2016.05.084

Reiser, B. J., Tabak, I., Sandoval, W. A., Smith, B. K., Steinmuller, F., & Leone, A. J. (2001). BGuILE: Strategic and conceptual scaffolds for scientific inquiry in biology classrooms. *Cognition and instruction: Twenty-five years of progress*, 263-305.

Renner, B., Kimmerle, J., Cavael, D., Ziegler, V., Reinmann, L., & Cress, U. (2014). Web-based apps for reflection: a longitudinal study with hospital staff. *Journal of Medical Internet Research, 16*(3), e85. doi: 10.2196/jmir.3040

Roessger, K. M. (2014). The effect of reflective activities on instrumental learning in adult work-related education: A critical review of the empirical research. *Educational Research Review, 13*, 17-34. doi: 10.1016/j.edurev.2014.06.002

Rubenstein, L. D., Callan, G. L., & Ridgley, L. M. (2017). Anchoring the Creative Process Within a Self-Regulated Learning Framework: Inspiring Assessment Methods and Future Research. *Educational Psychology Review*, 1-25.

Seltman, H. J. (2012). Experimental design and analysis. *Pittsburgh: Carnegie Mellon University, 428*.

Senko, C., Durik, A. M., Patel, L., Lovejoy, C. M., & Valentiner, D. (2013). Performance-approach goal effects on achievement under low versus high challenge conditions. *Learning and Instruction, 23*, 60-68. doi: 10.1016/j.learninstruc.2012.05.006

Siadaty, M., Gašević, D., & Hatala, M. (2016). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics, 3*(1).

Song, H. S., Kalet, A. L., & Plass, J. L. (2016). Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *Journal of Computer Assisted Learning, 32*(1), 31-50.

Sonnenberg, C., & Bannert, M. (2015). Discovering the Effects of Metacognitive Prompts on the Sequential Structure of SRL-Processes Using Process Mining Techniques. *Journal of Learning Analytics, 2*(2015), 72-100.

Spring, K. J., & Graham, C. R. (2017). Thematic Patterns in International Blended Learning Literature, Research, Practices, and Terminology. *Online Learning, 21*(4).

Studer, M., Mueller, N. S., Ritschard, G., & Gabadinho, A. (2010). *Classer, discriminer et visualiser des séquences d'événements.* Paper presented at the EGC.

Suthers, D., & Verbert, K. (2013). *Learning analytics as a middle space.* Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge.

Takallou, F. (2011). The effect of metacognitive strategy instruction on EFL learners' reading comprehension performance and metacognitive awareness. *Asian EFL Journal, 13*(1).

van den Boom, G., Paas, F., & van Merriënboer, J. J. G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction, 17*(5), 532-548. doi: http://dx.doi.org/10.1016/j.learninstruc.2007.09.003

van den Boom, G., Paas, F., van Merriënboer, J. J. G., & van Gog, T. (2004). Reflection prompts and tutor feedback in a web-based learning environment: effects on students' self-regulated learning competence. *Computers in Human Behavior, 20*(4), 551-567. doi: 10.1016/j.chb.2003.10.001

Van Laer, S., & Elen, J. (2016). Adults' Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. *Technology, Knowledge and Learning*, 1-31.

Van Laer, S., & Elen, J. (2018). An Instrumentalized Framework for Supporting Learners' Self-regulation in Blended Learning Environments. In M. J. Spector, B. B. Lockee, & M. D. Childress (Eds.), *Learning, Design, and Technology: An International Compendium of Theory, Research, Practice, and Policy* (pp. in press). Cham: Springer International Publishing (in press).

Van Woerkom, M. (2004). The concept of critical reflection and its implications for human resource development. *Advances in developing human resources, 6*(2), 178-192.

Vandergrift, L., & Tafaghodtari, M. H. (2010). Teaching L2 learners how to listen does make a difference: An empirical study. *Language learning, 60*(2), 470-497. doi: 10.1111/j.1467-9922.2009.00559.x

Veenman, M. V., & Alexander, P. (2011). Learning to self-monitor and self-regulate. *Handbook of research on learning and instruction*, 197-218.

Veenman, M. V., Van Hout-Wolters, B. H., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1*(1), 3-14.

Veenman, M. V. J. (1993). *Intellectual ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*: Faculteit Psychologie, Universiteit van Amsterdam.

Vohs, K. D., & Baumeister, R. F. (2016). *Handbook of self-regulation: Research, theory, and applications*: Guilford Publications.

Wehmeyer, M. L., & Shogren, K. A. (2017). Goal Setting and Attainment and Self-regulation *Handbook of Positive Psychology in Intellectual and Developmental Disabilities* (pp. 231-245): Springer.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in educational theory and practice, 93*, 27–30.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551-572.

Winters, F. I., Greene, J. A., & Costich, C. M. (2008). Self-regulation of learning within computer-based learning environments: A critical analysis. *Educational Psychology Review, 20*(4), 429-444. doi: 10.1007/s10648-008-9080-9

Wolters, C. A., Pintrich, P. R., & Karabenick, S. A. (2005). Assessing academic self-regulated learning. *What Do Children Need to Flourish?*, 251-270.

Zhou, M. (2016). Data pre-processing of student e-learning logs *Information Science and Applications (ICISA) 2016* (pp. 1007-1012): Springer.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational Psychologist, 48*(3), 135-147. doi: 10.1080/00461520.2013.794676

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*: Routledge.

Appendix 1: cues for reflection

Before the task

- Do you have any idea which element of statistics this topic will be about?
- This topic introduces … Do you know what this concept means?
- What other concepts relate to …?
- Today we will look at … Do you have any idea how this relates to the bigger picture of this course?
- Why would you need … to complete your individual project?
- What do you need to know about … to understand it and be able to apply it?
- Do you have any idea why … is important to your professional and personal context?
- How will you approach this topic?
- Hearing the word … what would you like to do with it?
- Hearing the word … what does it mean for you?
- How are you planning to relate … to the bigger picture?
- Do you think you have the skills needed to use the information presented here?
- What will you do when you identify a lack in information?
- How do you plan to overcome problems?
- Where will you get help if needed?
- What skills do you have to deal with this topic?
- What kind of issues do you see when trying to master the concept …?
- What actions will you take when you figure out the topic is not about what you thought it would be about?
- At what point will you feel you have mastered the topic?
- When do you believe you are taking the right actions to achieve mastery of the topic?
- Which steps do you want to take to master the concept …?
- How will you ensure you take the most suitable steps to master the concept …?

During the task

- Based on the first part of the task, is the task about what you thought it was about?
- If not, what will you do about this discrepancy?
- What different elements do you need to combine to complete this task?
- Do you possess each of these elements?
- How does this task relate to the tasks you were given before?
- Why do you need to do this task?
- How will this task help you to master the bigger picture?
- How is this task of importance to you?
- Does this task still fulfil the role you thought it would fulfil?
- Does the task still help you to achieve your goals?
- How does this task relate to your professional and personal context?
- What can you do to maximize this fit even more?
- Are the goals you set at the beginning of the task still the best ones or did you acquire new knowledge that means you need to reframe the goals?
- What modifications do you need to make to tune your initial approach to how the task evolves?
- What do you think this task will lead to?
- Is your plan still in line with your initial one?

- How do you deal with discrepancies between what you thought this task would be about and what you know now?
- Is your plan to approach this task still appropriate to achieve the goals you set?
- Did you encounter any obstacles while solving the task?
- How did you deal with obstacles?
- I the task unfolding as expected?
- How does the unfolding of the task relate to your approach?

After the task

- Was the result of the task what you expected it to be?
- Did your idea match the final demands of the task?
- Were you able to identify the different elements of the task as expected?
- Do you see at this point how the task relates to the overall aim of the course?
- Do you understand why the instructor provided you with the different elements of the task?
- Is it clear for you what the significance of the task was?
- After completing the task, do you see the relevance for real life?
- Knowing what you know now about the task, would you approach it the same way?
- How will you approach the next task, based on the task you just completed?
- Are the goals you set for this task fulfilled?
- Was your approach appropriate to achieve the goals?
- Did your plans unfold as expected?
- Were there obstacles in achieving the goals you set?
- How did you deal with obstacles?
- Was your approach to dealing with obstacles effective?
- Are there things you will do differently if you get a similar task?
- Was the path you took to achieving the goal successful for you?
- Were the steps taken to achieve the goal sufficient?
- Was there anything you learned that will change your approach to the next task?
- Which factors contributed to the success/failure of the task?
- What advice would you give other students with regard to the task?
- What will be your approach from now on?

Appendix 2: Frequency of significant discriminant sub-sequence occurrence (condition)

#	Subsequence	Support	p.value	statistic	index	Resid.1	Resid.2	Cramer's V
1	(Course)-(File)-(Course)-(Course)-(Course)	0.267	0.001	11.271	578.000	-1.876	2.654	0.613
2	(Course)-(Course)-(File)-(Course)	0.433	0.001	10.605	230.000	-1.585	2.242	0.595
3	(Course)-(File)-(File)-(Test)-(Test)-(Test)	0.467	0.003	8.856	205.000	-1.418	2.006	0.543
4	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)	0.467	0.003	8.856	206.000	-1.418	2.006	0.543
5	(Course)-(Task)-(Task)-(Course)	0.300	0.003	8.750	501.000	-1.633	2.309	0.540
6	(File)-(Course)-(File)-(File)	0.600	0.006	7.656	112.000	-1.155	1.633	0.505
7	(Course)-(File)-(File)-(Test)-(Test)	0.500	0.007	7.350	177.000	-1.265	1.789	0.495
8	(File)-(Course)-(Course)	0.633	0.011	6.477	93.000	-1.030	1.457	0.465
9	(File)-(Course)-(File)	0.633	0.011	6.477	94.000	-1.030	1.457	0.465

10	(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.633	0.011	6.477	101.000	-1.030	1.457	0.465
11	(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.633	0.011	6.477	102.000	-1.030	1.457	0.465
12	(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.633	0.011	6.477	104.000	-1.030	1.457	0.465
13	(Course)-(Test)-(Course)-(Page)	0.267	0.013	6.158	594.000	-1.443	2.041	0.453
14	(File)-(Link)-(Course)	0.267	0.013	6.158	615.000	-1.443	2.041	0.453
15	(File)-(Test)-(Course)-(Test)	0.267	0.013	6.158	619.000	-1.443	2.041	0.453
16	(Page)-(Course)-(Forum)-(Forum)	0.267	0.013	6.158	644.000	-1.443	2.041	0.453
17	(Test)-(File)-(Course)	0.267	0.013	6.158	659.000	-1.443	2.041	0.453
18	(Course)-(File)-(Course)-(Course)	0.433	0.013	6.126	232.000	-1.246	1.761	0.452
19	(File)-(Course)-(Test)	0.433	0.013	6.126	246.000	-1.246	1.761	0.452
20	(Forum)-(Forum)-(Course)-(Forum)	0.433	0.013	6.126	257.000	-1.246	1.761	0.452

21	(Forum)-(Forum)-(Forum)-(Course)-(Forum)	0.433	0.013	6.126	259.000	-1.246	1.761	0.452
22	(Course)-(File)-(File)-(Course)	0.533	0.014	6.044	152.000	-1.123	1.588	0.449
23	(Test)-(Test)-(File)	0.533	0.014	6.044	167.000	-1.123	1.588	0.449
24	(Test)-(Test)-(Test)-(File)	0.533	0.014	6.044	170.000	-1.123	1.588	0.449
25	(Course)-(File)-(Course)	0.667	0.020	5.419	78.000	-0.913	1.291	0.425
26	(File)-(File)-(File)-(File)	0.667	0.020	5.419	80.000	-0.913	1.291	0.425
27	(File)-(Task)	0.667	0.020	5.419	83.000	-0.913	1.291	0.425
28	(Course)-(Course)-(Page)	0.367	0.023	5.185	334.000	-1.231	1.741	0.416
29	(Course)-(Course)-(Page)-(Course)	0.367	0.023	5.185	335.000	-1.231	1.741	0.416
30	(Course)-(File)-(Course)-(File)-(File)	0.367	0.023	5.185	337.000	-1.231	1.741	0.416
31	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)	0.367	0.023	5.185	342.000	-1.231	1.741	0.416

32	(File)-(Course)-(Course)-(Course)	0.367	0.023	5.185	360.000	-1.231	1.741	0.416
33	(Forum)-(Forum)-(Forum)-(Forum)-(Course)-(File)	0.367	0.023	5.185	375.000	-1.231	1.741	0.416
34	(Forum)-(Forum)-(Forum)-(Forum)-(Forum)-(Course)-(Forum)	0.367	0.023	5.185	376.000	-1.231	1.741	0.416
35	(Course)-(Course)-(Course)-(File)	0.567	0.027	4.904	126.000	-0.990	1.400	0.404
36	(Forum)-(Course)-(Forum)	0.567	0.027	4.904	137.000	-0.990	1.400	0.404
37	(Test)-(File)	0.567	0.027	4.904	147.000	-0.990	1.400	0.404
38	(Course)-(Course)-(Test)-(Test)-(Test)-(File)	0.467	0.028	4.838	203.000	-1.091	1.543	0.402
39	(File)-(Course)-(File)-(File)-(Course)	0.467	0.028	4.838	211.000	-1.091	1.543	0.402
40	(File)-(File)-(Task)	0.467	0.028	4.838	216.000	-1.091	1.543	0.402
41	(Forum)-(Forum)-(Forum)-(Forum)-(Forum)-(Forum)	0.467	0.028	4.838	220.000	-1.091	1.543	0.402
42	(Task)-(Course)-(Test)	0.467	0.028	4.838	222.000	-1.091	1.543	0.402

43	(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.700	0.035	4.464	52.000	-0.802	1.134	0.386
44	(Course)-(Course)-(File)	0.700	0.035	4.464	53.000	-0.802	1.134	0.386
45	(Course)-(Test)-(Test)-(Test)-(Test)	0.700	0.035	4.464	60.000	-0.802	1.134	0.386
46	(Course)-(Test)-(Test)-(Test)-(Test)-(Test)	0.700	0.035	4.464	61.000	-0.802	1.134	0.386
47	(File)-(File)-(Test)-(Test)	0.700	0.035	4.464	63.000	-0.802	1.134	0.386
48	(Test)-(Test)-(Course)-(Course)-(Course)-(Test)	0.700	0.035	4.464	69.000	-0.802	1.134	0.386
49	(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)	0.700	0.035	4.464	70.000	-0.802	1.134	0.386
50	(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)	0.700	0.035	4.464	71.000	-0.802	1.134	0.386
51	(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)	0.700	0.035	4.464	72.000	-0.802	1.134	0.386
52	(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)	0.700	0.035	4.464	73.000	-0.802	1.134	0.386
53	(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)	0.700	0.035	4.464	74.000	-0.802	1.134	0.386

54	(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)	0.700	0.035	4.464	75.000	-0.802	1.134	0.386
55	(Course)-(File)-(Course)-(Test)	0.300	0.035	4.464	484.000	-1.225	1.732	0.386
56	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)-(Course)	0.300	0.035	4.464	487.000	-1.225	1.732	0.386
57	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)	0.300	0.035	4.464	488.000	-1.225	1.732	0.386
58	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)	0.300	0.035	4.464	489.000	-1.225	1.732	0.386
59	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)	0.300	0.035	4.464	490.000	-1.225	1.732	0.386
60	(Course)-(File)-(File)-(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.300	0.035	4.464	491.000	-1.225	1.732	0.386
61	(Course)-(File)-(Link)	0.300	0.035	4.464	494.000	-1.225	1.732	0.386
62	(Course)-(Test)-(Test)-(Test)-(Test)-(Test)-(Course)	0.300	0.035	4.464	506.000	-1.225	1.732	0.386
63	(File)-(Course)-(Forum)-(Forum)-(Forum)	0.300	0.035	4.464	511.000	-1.225	1.732	0.386
64	(File)-(File)-(File)-(Task)	0.300	0.035	4.464	519.000	-1.225	1.732	0.386

65	(Forum)-(Forum)-(Course)-(Page)-(Course)	0.300	0.035	4.464	534.000	-1.225	1.732	0.386
66	(Task)-(Course)-(Forum)	0.300	0.035	4.464	543.000	-1.225	1.732	0.386
67	(Task)-(Course)-(Forum)-(Forum)	0.300	0.035	4.464	544.000	-1.225	1.732	0.386
68	(Task)-(Course)-(Test)-(Test)	0.300	0.035	4.464	546.000	-1.225	1.732	0.386
69	(Course)-(File)-(File)-(Test)	0.600	0.048	3.906	108.000	-0.866	1.225	0.361
70	(Task)-(Task)-(Task)	0.600	0.048	3.906	116.000	-0.866	1.225	0.361
71	(Task)-(Task)-(Task)-(Task)	0.600	0.048	3.906	117.000	-0.866	1.225	0.361
72	(Task)-(Task)-(Task)-(Task)-(Task)	0.600	0.048	3.906	118.000	-0.866	1.225	0.361
73	(Task)-(Task)-(Task)-(Task)-(Task)-(Task)	0.600	0.048	3.906	119.000	-0.866	1.225	0.361
74	(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)	0.600	0.048	3.906	124.000	-0.866	1.225	0.361
75	(Course)-(Course)-(Test)-(Test)-(Test)-(Test)	0.400	0.048	3.906	279.000	-1.061	1.500	0.361

76	(File)-(Link)	0.400	0.048	3.906	298.000	-1.061	1.500	0.361
77	(Forum)-(Forum)-(Forum)-(Course)-(File)	0.400	0.048	3.906	305.000	-1.061	1.500	0.361
78	(Forum)-(Forum)-(Forum)-(Forum)-(Course)-(Forum)	0.400	0.048	3.906	307.000	-1.061	1.500	0.361
79	(Forum)-(Forum)-(Forum)-(Forum)-(Forum)-(Forum)-(Course)	0.400	0.048	3.906	308.000	-1.061	1.500	0.361
80	(Test)-(Test)-(Course)-(Test)-(Test)-(Course)	0.400	0.048	3.906	321.000	-1.061	1.500	0.361

Appendix 3: Frequency of significant discriminant sub-sequence occurrence (PAA)

#	Subsequence	Support	p.value	statistic	index	Resid.1	Resid.2	Resid.3	Resid.4	Resid.5	Cramer's V
1	(Test)-(Page)	0.28	0.008	13.846	1274	-0.53	2.37	-1.39	-0.52	1.37	0.877
2	(Course)-(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.023	11.354	1034	1.37	-0.91	-1.39	1.81	-0.53	0.794
3	(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.008	13.846	1260	-0.53	2.37	-1.39	-0.52	1.37	0.877
4	(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.008	13.846	1243	-0.53	2.37	-1.39	-0.52	1.37	0.877
5	(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.008	13.846	1208	-0.53	2.37	-1.39	-0.52	1.37	0.877
6	(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.008	13.846	940	-0.53	2.37	-1.39	-0.52	1.37	0.877
7	(Test)-(Test)-(Test)-(Test)-(Course)-(Course)-(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.28	0.008	13.846	957	-0.53	2.37	-1.39	-0.52	1.37	0.877
8	(File)-(Course)-(File)-(Task)	0.33	0.034	10.393	782	-0.58	2.00	-0.87	-0.71	1.15	0.760
9	(File)-(File)-(Course)-(File)-(Task)	0.33	0.034	10.393	916	-0.58	2.00	-0.87	-0.71	1.15	0.760

10	(Course)-(Test)-(Test)-(Test)-(Course)-(Test)-(Test)-(Test)-(Test)	0.28	0.028	10.879	1079	-0.53	2.37	0.04	-1.29	-0.53	0.777
11	(Test)-(Link)	0.33	0.034	10.393	906	-0.58	2.00	-0.87	-0.71	1.15	0.760
12	(Test)-(Link)-(Course)	0.28	0.023	11.354	1257	1.37	-0.91	-1.39	1.81	-0.53	0.794
13	(Test)-(Test)-(Link)	0.28	0.023	11.354	1258	1.37	-0.91	-1.39	1.81	-0.53	0.794
14	(Test)-(Test)-(Link)-(Course)	0.33	0.024	11.250	908	-0.58	0.00	-1.53	2.12	-0.58	0.791
15	(Test)-(Test)-(Page)	0.22	0.036	10.286	1756	-0.47	-0.82	-1.25	2.31	-0.47	0.756
16	(Course)-(Course)-(Test)-(Course)	0.22	0.036	10.286	1755	-0.47	-0.82	-1.25	2.31	-0.47	0.756
17	(File)-(File)-(File)-(Test)-(Test)-(Course)	0.39	0.004	15.195	710	-0.62	-0.15	-1.65	2.40	-0.62	0.919
18	(Page)-(Course)-(File)-(File)	0.28	0.023	11.354	1238	1.37	-0.91	-1.39	1.81	-0.53	0.794
19	(Page)-(Course)-(File)-(File)-(File)	0.28	0.023	11.354	1239	1.37	-0.91	-1.39	1.81	-0.53	0.794
20	(Course)-(Test)-(Test)-(Test)-(File)-(File)	0.72	0.028	10.879	167	0.33	0.57	-1.36	0.80	0.33	0.777

21	(Test)-(Test)-(File)-(File)	0.28	0.048	9.574	994	-0.53	2.37	-0.68	-0.52	-0.53	0.729
22	(Test)-(Test)-(Test)-(File)-(File)	0.28	0.048	9.574	993	-0.53	2.37	-0.68	-0.52	-0.53	0.729
23	(Task)-(File)-(Course)	0.28	0.048	9.574	992	-0.53	2.37	-0.68	-0.52	-0.53	0.729
24	(Test)-(File)-(Forum)	0.28	0.048	9.574	991	-0.53	2.37	-0.68	-0.52	-0.53	0.729
25	(Test)-(File)-(Forum)-(Forum)	0.28	0.048	9.574	990	-0.53	2.37	-0.68	-0.52	-0.53	0.729
26	(Course)-(File)-(Task)-(Task)	0.22	0.011	13.041	1470	-0.47	2.86	-0.45	-1.15	-0.47	0.851
27	(Course)-(File)-(Task)-(Task)-(Task)	0.22	0.011	13.041	1458	-0.47	2.86	-0.45	-1.15	-0.47	0.851
28	(Course)-(File)-(Task)-(Task)-(Task)-(Task)	0.22	0.036	10.286	1694	-0.47	-0.82	-1.25	2.31	-0.47	0.756
29	(Course)-(File)-(Task)-(Task)-(Task)-(Task)-(Task)	0.61	0.029	10.787	288	0.50	-1.35	-0.13	1.22	-0.78	0.774
30	(Course)-(File)-(Task)-(Task)-(Task)-(Task)-(Task)-(Task)	0.61	0.029	10.787	287	0.50	-1.35	-0.13	1.22	-0.78	0.774

The Effect of Cues for Calibration on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

Abstract

Literature on blended learning emphasizes the importance of self-regulation for learning in blended learning environments and the role of learners' calibration. Although literature on calibration is clear on its importance for self-regulated learning, it provides inconclusive insight in the effect of support for calibration on learners' self-regulated learning. One under-investigated avenue might be learners' ability to enact on the cues provided. In order to establish a more accurate picture of the effect of support for calibration on self-regulated learning, our study investigates whether providing cues for calibration affect learners' self-regulated learning, and whether this effect is different for learners with different metacognitive abilities. We investigate this effect by examining changes in learners' learning behaviour and outcomes. A pre-post design with one control and two experimental conditions was applied in a blended learning environment. Learners in the experimental conditions received either functional validity feedback (F-condition) or functional and cognitive validity feedback (FC-condition). Learners in the control condition did not receive any cues. Learners' behaviour was analysed using event sequence analysis. Learners' post-test learning scores were subjected to multivariate analysis of covariance, with condition and learners' metacognitive ability as independent variables. The results show a significant and unexpected impact of condition and learners' metacognitive abilities on learners' learning behaviour and outcomes. This manuscript discusses the unexpected results in terms of their theoretical and practical implications and provides recommendations for future research. We conclude that when cues for calibration are provided through functional and cognitive validity feedback, learners' calibration capabilities will increase. Yet for this to result in goal-directed self-regulated learning, learners' need to be supported on how to apply the cognitive and metacognitive strategies needed.

Keywords: functional validity feedback; cognitive validity feedback; self-regulated learning; learning outcomes; learning behaviour; computer log files

Highlights:

- Calibration cues through functional and cognitive validity feedback are effective
- Combined they affect learners' judgement of learning and learning confidence
- Functional and cognitive validity feedback increases learners' judgement of learning
- Learners need to be supported to apply successful strategies to increase performance

1. Introduction

Learners in blended learning environments need to be able to deal with varying degrees of autonomy and to judge and adapt their learning to the learning outcomes imposed. Based on this assumption, current literature on technology-enhanced learning emphasizes the importance of self-regulation in blended learning (e.g., Boekaerts, 1999; Greene & Azevedo, 2007; Vohs & Baumeister, 2016) and more specifically the role of learners' calibration in the monitoring phase of self-regulated learning (e.g., Lin, Coburn, & Eisenberg, 2016; Pajares & Schunk, 2001). Although instructional interventions fostering self-regulated learning have been investigated widely in different educational settings (e.g., Arrastia-Chisholm, Torres, & Tackett, 2017; Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015), the actual effect of support for calibration on learners' self-regulated learning remains unclear (Panadero, Klug, & Järvelä, 2016). In general, literature investigating learners' calibration hypothesizes that learners are well calibrated if they perceive links between their learning behaviour, cues provided, information presented, and the task at hand, and when their perceptions reflect reality (Butler & Winne, 1995; Nelson, Narens, & Bower, 1990). In that case learners are equipped to effectively monitor their learning (DiFrancesca, Nietfeld, & Cao, 2016; Zimmerman, Schunk, & DiBenedetto, 2015). However, even when learners can calibrate external and internal feedback, they might not possess or be able to recall the cognitive and/or metacognitive strategies needed to act in a way that will produce increased learning outcomes (e.g., Pintrich, 2002; Veenman, 1993). This would result in sub-optimal self-regulated learning.

Although literature on calibration is clear on its importance for self-regulated learning, it provides insufficient insight into how to support learners' calibration and so self-regulated learning (Stone, 2000; Yang, Potts, & Shanks, 2017). In order to establish a more accurate picture of the effect of cues for calibration on learners' self-regulated learning, this study investigates whether cues for calibration do actually affect self-regulated learning in a blended learning environment, and whether this effect is different for learners with different metacognitive abilities. We operationalize self-regulated learning as changes in learners' learning behaviour and outcomes. Investigating learning behaviour and outcomes provides insights on learners' self-regulated learning, as well as on the nature of cues' effects (Gašević, Dawson, & Siemens, 2015). In the next part of the introduction, we elaborate on blended learning and the conceptualization of self-regulated learning and present a theoretical basis for designing cues for calibration intended to evoke learners' self-regulated learning through monitoring. In the final part of the introduction, we discuss the relationship between self-regulated learning, learning behaviour, and learning outcomes.

1.1. Blended learning

Blended learning is a well-established approach that is applied in various educational contexts (e.g., Bonk, 2017; Spring & Graham, 2017). A frequently recurring aspect among definitions of blended learning is its combination of online and classroom-based learning. Current conceptualizations of blended learning assume that blended learning environments combine the advantages of both modes of delivery (e.g., Graham, Henrie, & Gibbons, 2014; McCutcheon, Lohan, Traynor, & Martin, 2015). In line with this assumption, the current study defines blended learning as learning in an instructional context which is characterized by a deliberate combination of online and classroom-based interventions to instigate and support learning (Boelens, Van Laer, De Wever, & Elen, 2015). Many researchers have investigated different aspects of blended learning (Drysdale, Graham, Spring, &

Halverson, 2013), with the lion's share of studies focusing either on comparing different modes of delivery (Halverson, Graham, Spring, Drysdale, & Henrie, 2014) or on learners' compliance with blended learning environments (Deschacht & Goeman, 2015). With regard to the latter, research has identified for example that learners who have a high degree of control over their learning and intrinsic goal orientation (e.g., Kassab, Al-Shafei, Salem, & Otoom, 2015) and learners with high cognitive and metacognitive capabilities (e.g., Kizilcec, Pérez-Sanagustín, & Maldonado, 2017) often perform better in blended learning environments compared to learners who do not have these characteristics.

1.2. Self-regulated learning

Blended learning environments do not themselves result in learning, as all learning requires activity on the part of learners (Oliver & Trigwell, 2005). More precisely, blended learning is a self-regulated process in which learners regulate their behaviour according to the instructional demands (Zimmerman & Schunk, 2001). Literature shows that performance strongly positively correlates with self-regulated learning variables (e.g., Daniela, 2015; Lin et al., 2016). Self-regulated learning has been widely investigated and various theories have been proposed (see: Puustinen & Pulkkinen, 2001). Each of these theories describes self-regulated learning as cyclical, influenceable, and covert in nature. With regard to its cyclical nature, self-regulated learning roughly consists of three phases, namely (a) a forethought phase, (b) an enacting phase, and (c) an evaluation phase. Each theory also draws attention to the key role of internal (learner characteristics) and external factors (design of the learning environment) in the development of self-regulated learning (Zimmerman, 2013). This cyclical and influenceable nature of self-regulated learning means that continuous measurements are needed to capture learners' self-regulated learning. Finally, regarding the covert nature of self-regulated learning, the metacognitive processes that occur in the different phases of self-regulated learning cannot be observed directly. If a learner reviews her exercises after completing an online test, for example, it might be assumed that an evaluation or monitoring process must have preceded this overt cognitive reviewing activity. This example shows how self-regulated learning is manifested through overt cognitive behaviours (i.e., learners' learning behaviour) and behavioural consequences (i.e., learners' learning outcomes) (Veenman & Alexander, 2011). Given the dynamic nature of self-regulated learning, investigations would benefit from using continuous measurements and inferences that draw on learners learning' behaviour and learning outcomes.

1.3. Monitoring, calibration, and self-regulated learning

Learners working on a task need to monitor their actions to be sure their enactment leads to achieving the targeted learning outcomes. Monitoring is the cognitive operation influencing whether an action is taken or not (Muis, Winne, & Ranellucci, 2016; Winne & Jamieson-Noel, 2002). When learners interact with different tasks, information about changes in learning outcomes is monitored relative to learners' perceived changes. When discrepancies exceed an idiosyncratic threshold, self-regulating learners adjust their behaviour to eliminate the discrepancies (Winne & Hadwin, 1998). So for this adjustment to be effective, good calibration between perceptions of and actual changes in learning outcomes is needed. The better learners' calibration is, the more accurate monitoring will be (Stolp & Zabrucky, 2017). This accuracy is often referred to as judgement of learning (JOL) (Schraw, 2009). The information available to learners to calibrate and hence to monitor changes in learning outcomes has two main sources. Either changes in learning outcomes in reality (external feedback), or cognitive representations (internal feedback) by the learners of changes in learning outcomes (e.g., Ariel & Karpicke, 2018; Broadbent & Poon, 2015). For learners to be able to accurately calibrate, they have to process the external feedback along with the internal feedback (Winne & Hadwin, 1998). More precisely, learners have to compare (a) the internal feedback with the desired level of change in learning outcomes, (b) the external feedback with the desired level of change in learning outcomes, and (c) the internal feedback with the external feedback. Based on the results of this process, learners monitor their learning and select cognitive and metacognitive strategies (e.g., error correction strategies, revision activities, etc.) which may help to proceed them in the direction of the desired level of learning outcomes (Narciss, 2017).

1.4. Supporting learners' calibration for self-regulated learning

One reason for interest in learners' calibration is that learners do use the result of the comparison of internal and external feedback to make decisions about how to monitor and self-regulate learning. Thus, low levels of calibration can undermine effective regulation (e.g., Dunlosky & Rawson, 2012; Thiede, Anderson, & Therriault, 2003). Despite its importance, a substantial body of literature suggests learners are generally not especially good at accurately judging themselves (Bol & Hacker, 2001; Klassen, 2002). And even when they are, extensive reviews by Dunlosky and Thiede (2013) and Alexander (2013) show that this does not mean that learners have any particular insight into this aspect of cognition.

To be able to provide support to learners' calibration attempts and evoke monitoring, literature on cues for calibration proposes two approaches, outcome feedback and cognitive feedback. The simplest and most common type of feedback is outcome feedback (e.g., Delgadillo et al., 2017; Earley, Northcraft, Lee, & Lituchy, 1990; Paulson Gjerde, Padgett, & Skinner, 2017). This type of feedback is binary information describing whether or not results are correct. It contains no additional information (e.g., about task, tools provided, or support offered) other than the state of the current learning outcomes (Butler & Winne, 1995). Hence, outcome feedback provides minimal external support for learners about how to self-regulate their learning (Kluger & DeNisi, 1996). Alternatively, feedback can be elaborated to supply several different types of information. Cognitive feedback can provide learners with information that links the task, tools or support provided, and changes in learning outcomes (Butler & Winne, 1995). In line with this, research showed feedback providing validity-related information (i.e., cognitive feedback) was judged more effective than outcome feedback (e.g., Balzer & Doherty, 1989; Nadolski & Hummel, 2017; Ridder, McGaghie, Stokking, & Cate, 2015).

Cognitive feedback comes in two forms, namely functional validity feedback and cognitive validity feedback (e.g., Besser, 2016; Butler & Winne, 1995; Ernst & Steinhauser, 2017; Sedrakyan & Snoeck, 2016). Functional validity feedback, describes the relation between learners' estimate of change in learning outcomes and the actual change in learning outcomes (e.g., Bui & Loebbecke, 1996; Frysak, 2017; Popelka, 2015). For example, in an adaptive learning environment, learners might be asked to estimate their scores on a to-come test (in the form a JOL-cue). Then, after learners' estimates were compared to the actual score, functional validity feedback suggest to the learners, "You overestimated yourself, your score is 60% not 80%" (e.g., Mory, 2004). Cognitive validity feedback aims to evoke monitoring through the activation of learners' perceptions about the relationship between the different course components, information offered, cues provided, and potential change in learning outcomes (e.g., Chyung, 1996; Ellis, 2012). For example, in an adaptive learning environment, a learner who studies texts might be shown a cue: "You aren't using the advance organizer to guide your studying." (Butler & Winne, 1995). Cognitive validity feedback conveys the information that directs learners' further actions based on their estimate and actual performance.

Research investigating the use of functional and cognitive validity feedback does not lend uniform support to the effectiveness of these types of cues for calibration in practice. On the one hand it is reported that cognitive validity feedback helps learners distinguish those pieces of information most important to increase their learning outcomes (Popelka, 2015). Or when outcome or functional feedback is provided, learners tend to devote time and energy to compare themselves with others, rather than to develop ways to revise and make improvements to their learning (e.g., Nicol & Macfarlane-Dick, 2006; Van Popta, Kral, Camp, Martens, & Simons, 2017). Even a score accompanied by suggestions for improvement seemed to distract learners from addressing how they might improve their work (Wiliam & Thompson, 2007). On the other hand, cognitive validity information alone seems insufficient to support learners' monitoring, as without information on learners' change in learning outcomes, behaviour will not be goal-directed (Butler & Winne, 1995; Ellis, Carette, Anseel, & Lievens, 2014). In line with this inconclusive view, Gielen, Peeters, Dochy, Onghena, and Struyven (2010) showed that functional validity feedback was as effective as cognitive validity feedback.

1.5. Metacognitive skilfulness as influencing variable of calibration-cue use

Unarguably it is of vital importance to target cues for calibration so learners will eventually be directed to the right types of information. Nonetheless, cue-use research shows that this is only one part of the challenge. It evidences that not all learners equally use and benefit from cues provided (e.g., Lust, Vandewaetere, Ceulemans, Elen, & Clarebout, 2011; Rashid & Asghar, 2016; Winne & Hadwin, 1998). Learners often do not utilize well-developed cues (Cleary, Callan, & Zimmerman, 2012). These learners typically make ineffective, suboptimal learning choices (Segedy, 2014). One of the possible explanations of sub-optimal cue use provided by the self-regulated learning literature is that learners might lack the skills needed or the ability to activate the targeted cognitive or metacognitive strategies. Even when learners can calibrate external and internal feedback, they might not possess or be able to recall the cognitive and/or metacognitive strategies needed to act in a way that will produce increased learning outcomes (e.g., Pintrich, 2002; Veenman, 1993). This would result in sub-optimal self-regulated learning. A way to investigate this issue and partly explain differences in the effect of the cues provided might be the investigation of learners' metacognitive abilities in relation to the instructional interventions provided (e.g., Ardasheva, Wang, Adesope, & Valentine, 2017). Metacognitive skilfulness or the ability to regulate and control cognitive and metacognitive strategies is a combination of general cognitive and metacognitive strategies. By mapping learners' metacognitive skilfulness differences in the use of cues provided can be investigated (Veenman, Elshout, & Meijer, 1997). Doing so will provide us with a more fine-grained picture of the effect of cues for calibration.

1.6. Investigating learners' self-regulated learning

Due to self-regulated learning's dynamic nature and its covert nature, continuous measurements and inferences through learners' learning behaviour and learning outcomes are needed to capture learners' self-regulated learning. Investigating both learners' learning behaviour and outcomes provides insights not only into learners' self-regulated learning, but also into the nature of cues' effects (e.g., Bannert, Molenaar, Azevedo, Järvelä, & Gašević, 2017).

1.6.1. Learning behaviour and self-regulated learning

Behavioural traces gathered from learners during instructional processes can be categorized as learners' learning behaviour (Entwistle & Peterson, 2004). As support grows for the conception of self-regulated learning as a continuous process, so does interest in online measures such as thinking-aloud protocols, eye-movement tracking and log-file registration, which account for the dynamic nature of learners' learning behaviour. Some of these methods (i.e., log file registration) involve no direct interaction with the learner, and this unobtrusiveness enables researchers to trace learning events in ecologically valid settings. Thus, we can "re-play" learners' learning behaviour and obtain a better understanding of what leads to certain learning outcomes.

Cleary et al. (2012) describe the cyclical relationship among the components of self-regulated learning as 'sequential phases of regulation'. In relation to log file data, this cyclical nature is reflected in the term 'event sequence', which describes patterns of learners' learning behaviour. Both 'event' and 'sequence' are common terms for describing all sorts of patterns in different fields of research (e.g., Abbott, 1995; Suthers & Verbert, 2013). When it comes to self-regulation, a first distinction made between trace data types is whether the basic information they contain relates to a state or an event. An example of a state might be being on a content page, while clicking on a self-test link might be an event that changes the state to being on a self-test page. This means that each change of state is an event, and each event indicates a change in state (Müller, Studer, Gabadinho, & Ritschard, 2010). Log files only capture events, as we assume in this study that we cannot determine what learners are doing between two events. Another distinction is whether the order of events or states is logged. If the original order is logged by, for example, the inclusion of timestamps, the data is considered to be sequenced; otherwise, it is perceived as an item set.

Event sequence data analysis generally takes one of three forms: pattern mining, pattern pruning, and interactive visualization design (Liu, Dev, Dontcheva, & Hoffman, 2016). The current study focusses on investigating differences in learners' learner behaviour, so our focus lies on pattern mining defined as the identification of meaningful event sequences (patterns) (e.g., Azevedo et al., 2016; Bannert et al., 2015; Siadaty, Gašević, & Hatala, 2016; Sonnenberg & Bannert, 2015). Pattern mining has two main concerns – the order of the events and the containment of sub-sequences. A sequential pattern implies the order of events has been preserved. A sub-sequence is a part of a larger sequence, which also appears elsewhere. Containment relates to the amount of support there is for a particular sub-sequence in the sample; in other words, the number (or percentage) of sub-sequences matching other learners' sub-sequences. A sub-sequence is considered frequent if it occurs in at least the threshold number of learners' sequences. Following sub-sequence identification, links between significant differences in subsequence occurrence and conditions either internal or external to the learner can be investigated in statistical trials.

1.6.2. Learning outcomes and self-regulated learning

Instruction may result in a variety of learning outcomes (e.g., Endedijk, Brekelmans, Verloop, Sleegers, & Vermunt, 2014). In this study, we focus on four main learning outcomes: (1) domain knowledge; (2) goal orientation; (3) academic self-concept; and (4) judgement of learning. We selected these learning outcomes as their relationship to self-regulated learning has already been investigated extensively.

Domain knowledge, firstly, relates to learners' knowledge of the content involved in a particular task (Greene & Azevedo, 2007). It has been widely demonstrated in the literature on expertise that when learners have more extensive domain knowledge, they are less reliant on the need to identify, use, and regulate metacognitive strategies during complex tasks or the acquisition of new information in the domain (e.g., Lesgold et al., 1988; Song, Kalet, & Plass, 2016).

Goal orientation was operationalized by Pintrich (2000) and Eccles and Wigfield (2002) in mastery and performance goals, along with their approach and avoidance forms. Most research on mastery goal orientation has focused on the approach form, often finding increased use of cognitive elaboration and organization strategies and more frequent help-seeking behaviour (e.g., Duffy & Azevedo, 2015; Kitsantas, Steen, & Huie, 2017; Midgley, 2014). The mastery-avoidance orientation remains underexplored, though Wolters, Pintrich, and Karabenick (2005) and Elliot and McGregor (2001) did observe a correlation between this orientation and test anxiety, tying in with the characterization of mastery-avoidant learners as perfectionists. The performance approach has been linked by some authors to certain productive strategy behaviours (e.g., Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002; Kitsantas et al., 2017; Mega, Ronconi, & De Beni, 2014), whereas others have claimed that its effects remain unclear (e.g., Dompnier, Darnon, & Butera, 2013; Midgley, Kaplan, & Middleton, 2001; Senko, Durik, Patel, Lovejoy, & Valentiner, 2013). Results on performance avoidance are also ambiguous, with some research associating it with negative outcomes, such as the use of fewer cognitive strategies (Pintrich, 2000), and others producing evidence for learners' increased use of cognitive strategies to test their own abilities and compare themselves to other learners (e.g., Collazo, Elen, & Clarebout, 2015; Crippen, Biesinger, Muis, & Orgill, 2009).

Academic self-concept, the third outcome, can be defined as an individual's perception of self within academia (Elliot & Dweck, 2013) and incorporates the two distinct concepts of competence and effort. It measures the degree to which learners feel that academic subjects are easy and that they are good at them (competence), and the degree to which learners like or dislike going to school and studying different subjects (effort) (Liu & Wang, 2005). Previous studies have revealed that academic self-concept has a strong and positive influence on the variety of metacognitive strategies that learners use. A stronger academic self-concept leads to a deeper engagement with the learning environment, while a less well-developed self-concept is associated with retreat and concentration on simpler cognitive strategies (e.g., Kuo, Walker, Schroder, & Belland, 2014).

Finally, learners' judgement was operationalized in line with Schraw (2009), focussing on learners' precision of estimation of future performance compared to actual performance (Maki, Shields, Wheeler, & Zacchilli, 2005). When learners are able to accurately estimate their performance, they are more likely to take appropriate action and so regulate their learning (Butler & Winne, 1995). Learners who have a poor judgement of learning tend to make ineffective, suboptimal learning choices (Segedy, 2014).

1.7. Problem statement and hypotheses

While the literature emphasizes the importance of self-regulation for learning in blended learning environments on the one hand and the role of learners' monitoring through calibration for self-regulated learning on the other, evidence is inconclusive on the use of cues for calibration and their effect on self-regulated learning. Given this inconclusiveness, guidelines for interventions are difficult to outline. Hence, new approaches are needed to better understand the underlying mechanisms that may help to understand the inconclusive results. To get more profound insights in the effect of cues for calibration, this study investigates whether cues for calibration in blended learning environments foster self-regulated learning through changes in learners' learning behaviour and outcomes, and if this effect is different for learners with different metacognitive abilities. This leads us to four hypotheses:

- Hypothesis 1: “Cues for calibration affect learners’ learning behaviour.”
- Hypothesis 2: “Cues for calibration affect learners’ learning behaviour differently for learners with different levels of metacognitive skilfulness.”
- Hypothesis 3: “Cues for calibration positively affect learners’ learning outcomes.”
- Hypothesis 4: “Cues for calibration positively affect learners’ learning outcomes most when learners have high levels of metacognitive skilfulness.”

2. Method

2.1. Participants

The participants in this study were 151 learners taking a course on instructional psychology and technology as part of a Bachelor’s degree in Educational Sciences from a large Belgian university. There were 134 women (88.74%) and 17 men (11.26%), which is a representative sample of the entire student population within the Faculty of Psychology and Educational Sciences. The learners were between 19 and 58 years of age (*M*=21.87, *SD*= 6.84). They were familiar with the domain of instructional psychology and technology to some extent, but before the experiment they had not acquired insight in the texts of Anderson (2005) and Mayer (2004), which were the subject of the study task in the experiment. The subject matter was expected to be entirely new to them. This was controlled for in a prior domain knowledge test. None of the participants was able to achieve the maximum score on the test’s questions, the average score was 4.5/10. It was concluded that the students could be divided over the experimental groups at random. All voluntarily participated in the study, some (different numbers for different analyses) were excluded along the way because of incomplete records.

2.2. Content and module description

2.2.1.Content

In the course ‘instructional psychology and technology’, the module dealing with ‘educational practice’ was targeted. In this module, two texts are discussed. Through the first text written by Anderson (2005) the instructors introduce the ‘revised taxonomy of Bloom’ and aim to provide learners with insights about: (1) the importance of learning objectives, (2) the difficulties with regard to the formulation of such objectives, (3) the differences between the initial and the revised taxonomy of Bloom, (4) the link of assessment and instruction with the taxonomy, and (5) the revised taxonomy’s potential application. The second text written by Mayer (2004) is used by the instructors to evoke learners’ reflection on (1) the difference between ‘pure discovery learning’ and ‘guided discovery learning’, (2) research on both, and (3) the implications for education.

2.2.2.Module description

The module 'educational practice' was organized in the second semester of the 2017-2018 academic year. The module was provided in a blended learning format and consisted firstly of an online learning module in Moodle. Between two face-to-face contact sessions, learners had 28 days to progress through the environment and study the texts. In addition, learners were invited to participate in a two hour-long contact sessions dealing with the content of the online learning module after having studied in Moodle.

2.3. The blended learning format

2.3.1.Experimental and control environments

The experimental and control environments were identical (except for the cues for calibration). With regard to the online component of the blended learning environment, a Moodle course was developed in a co-design fashion between instructors and researcher. For each text, the outline was identical and consisted of the following elements: (1) goals of the text, (2) introduction (examples from practice), (3) text, (4) exercises, and (5) self-tests. With regard to the exercises and self-tests, each section of the texts was supported by practice exercises (not obligatory), preparing the learners for a self-test about the section addressed. Following each practice exercise an obligatory self-test was provided. After this test was submitted, learners could progress to the next section. Learners were allowed to choose a text to start with. During the studying of the texts learner control was limited, but after having finished the entire online learning module, learners could navigate freely through it. Also a (6) discussion forum was provided for course related discussions. In the experimental conditions cues for calibration were added to each section of the texts, to be more precisely before (in the form of a JOL- cue) and after (in the form of validity feedback) each self-test. With regard to the latter only functional validity feedback was given to learners in the functional validity feedback condition (F-condition), whereas learners in the functional and cognitive validity condition (FC-condition) received both functional and cognitive validity feedback. Figure 1 shows the design of each condition.

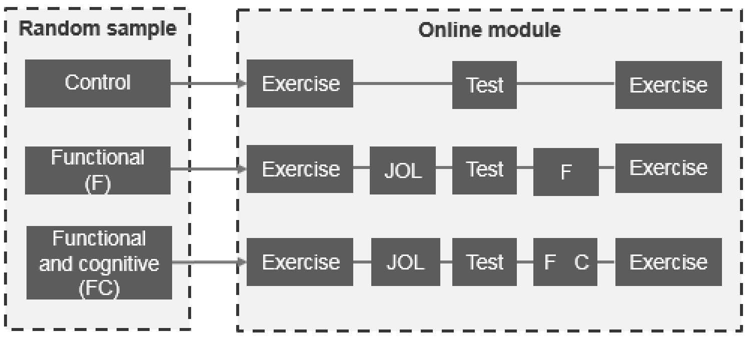

Figure 1. Visual representation of the intervention for each condition.

For each of the elements in the online learning module learners' interactions were tracked. The structure of the face-to-face contact moment (after the 28 days of independent study) was identical for all learners. Learners were divided in groups of 5 to 6 and were instructed to make (1) a concept map of both texts, (2) elaborate a multimedia knowledge clip explaining the elements of the concept map and their link with practice, and finally they were asked to come up with (3) exam questions targeting the different levels of the 'revised taxonomy of Bloom' and focussing on the content of the two texts. The results of these exercises were discussed in the entire group and after the sessions, all materials were made available to all learners for further study.

2.3.2. Cues for calibration

2.3.2.1. Judgment of learning cues

Both experimental conditions contained judgement of learning cues to generate input for the automatization of the validity feedback. In line with Schraw (2009) these cues provided the learners with a question about their expected performance on the upcoming self-test, potentially based on the practise exercises. The cues were embedded and so implemented in line with guidelines on cue-use as formulated by Clarebout and Elen (2009). Each of the cues was embedded in the environment at the same level of the other content items and presented prior to each self-test. Learners in the experimental conditions were obliged to estimate their score. The operationalization of this design in the online learning module was done by using tests, including the 'test' icon in Moodle. Each page looked similar. In the face-to-face contact moment, no cues were provided.

2.3.2.2. Functional and cognitive validity feedback cues

As mentioned before, both experimental conditions differed in the type of validity feedback cues learners received. Only functional validity feedback was given to learners in the F-condition, whereas learners in the FC-condition received both functional and cognitive validity feedback. Both the functional and cognitive validity feedback cues were developed in line with Balzer and Doherty (1989) and Butler and Winne (1995). The functional validly feedback cues focused on the relationship between learners' judgement of learning (absolute accuracy index) obtained via the JOL-cues and their actual performance (obtained via self-test scores). The accuracy of learners' judgement of learning was calculated using the absolute accuracy index formula proposed by Schraw (2009). This index was used as an indicator for assigning the appropriate functional validity feedback cue. This feedback contains any of the following labels: 'underestimate', 'estimate accurate', or 'overestimate' to the accuracy of prediction. Additionally learners were informed about their score on each self-test. A comparison of a learner's score was compared with the maximal score possible on the different self-tests. This was done by assigning either the label 'below 50% correct' or '50% or above 50% correct'. Combined this resulted in a personalized message for the learners stating: "You seem to [judgement of learning label] your ability and your score is [score label].".

The cognitive validity feedback cues pertained to learners' perceptions about the relationship between the instructional components and performance. Learners received information on the link between (1) the instructional components, (2) the cognitive and metacognitive strategies needed to use these components, and (3) on the potential impact of both on their performance. Both the functional and cognitive validity feedback cues were embedded in the environment at the same level of the other content items and provided after the completion of each self-test. Additionally, through a popup learners were informed that the content under the link of the cue might significantly help their learning. The operationalization of this design in the online learning module was done by using feedback forms, including the 'megaphone' icon in Moodle. Each page looked similar. In the face-to-face contact moment, no cues were provided.

2.4. Data structure of the log files

To investigate the suitability of using event sequence analysis, insight is needed in the data structure of the log file data. Each action made by learners within the online course was registered resulting in a time stamped event (TSE) database with as column headers the time stamp of the action, personal identifier of the user, and event name. A data-driven approach was chosen. Prior to the event analysis, no recoding or transformations took place. The three conditions (F-condition, FC-condition, and control condition) were identical and included six standard event names (see Table 1). Each of these event names refers to an attribute of the environment. Data reported on the attributes hence refer to specific (series of) events. In the F-condition two additional attributes were available, namely 'judgement of learning (JOL)' and 'functional validity feedback (F-feedback)'. In the FC-condition both were available too and a ninth one was available, 'cognitive validity feedback (C-feedback)'. As in the investigation of the relationship between cues for calibration and learners' behaviour the aim was to identify which sub-sequences occurred significantly more in which condition, sub-sequences including judgement of learning, functional validity, and cognitive validity feedback cues were excluded from the analysis as they only occur in the experimental conditions.

Table 1.

Actions traced in the online learning environment.

Attribute	***Description***
Course	Landing page of the course. On this page, learners found an overview of the entire course, links to each of the texts addressed, and links to the discussion forum.
Objective	Pages elaborating on the learning objectives aimed for by the different texts.
Text	Downloadable version of both texts addressed during the online learning module.
Forum	Discussion forum including the viewing of the forum, posting of information, and all other interactions related to the forum.
Exercise	Practise exercise for the support of learning prior to the self-test.
Self-test	Formative test on each section of the different texts under investigation. Obligatory to be able to progress to the next part of the online learning module.
JOL-cue	Formative test to estimate performance on the self-test for each section addressed in the different texts. Obligatory to be to progress to the next part of the online learning module.
F-feedback	Feedback page containing a functional validity feedback cue providing information and an open question (free to answer or not) aiming to evoke learners' calibration.
C-feedback	Feedback page containing a cognitive validity feedback cue providing information and an open question (free to answer or not) aiming to evoke learners' calibration.

Finally, in the face-to-face contact moments no trace data were gathered as all learners received the same instruction and did not interact with the Moodle environment.

2.5. Instruments

2.5.1.Prior domain knowledge and domain knowledge

During the pre-test phase a performance based prior domain knowledge test was administered to investigate learners' prior domain knowledge. This prior domain knowledge test containing ten multiple-choice questions (including an "I don't know option") represented the content of the module. The test consisted of questions related to both texts (five per text). The test was scored on ten points. The same test was used as post-test to measure learners' domain knowledge.

2.5.2.Goal orientation

Learners' goal orientation was measured by using the merged version of two questionnaires of Elliot and Church (1997) and Elliot and McGregor (2001) for measuring learners' goal orientation as constructed by Lust (2012). Whereas the initial questionnaire of Elliot and Church (1997) measured solely three dimensions of goal orientation (mastery approach, performance avoidance, and performance approach), the revised questionnaire Elliot and McGregor (2001) incorporated the fourth dimension of mastery avoidance as well. These two questionnaires were merged into one that contained 21 items (Mastery goal approach (MGA) (6 items), Mastery avoidance approach (MAA) (4 items), Performance goal approach (PGA) (5 items), Performance avoidance approach (PAA) (6 items)). Answers for the items were given on a 5-point Likert type scale ranging from 1 (strongly disagree) to 5 (strongly agree).

2.5.3.Academic self-concept

The Academic self-concept (ASC) scale comprised two 10-item subscales: learning confidence and learning effort. The learning confidence subscale assessed learners' feelings and perceptions about their academic competence (Liu, Wang, & Parkins, 2005). Example items are 'I am good in most of my course subjects' and 'most of my classmates are smarter than I am' (negatively worded). The learning effort subscale assessed learners' commitment to and involvement and interest in schoolwork (Liu et al., 2005).

2.5.4.Judgement of learning

Learners' judgement of learning was used in two ways. On the one hand as a dependent variable, on the other hand as the input for the adaptive feedback in the experimental conditions. For both purposes learners' judgement of learning was calculated in accordance with Schraw (2009). Learners' absolute accuracy was measured, indicating the precision of a single estimation of future score compared to performance on a single test (Maki, Shields, Wheeler, & Zacchilli, 2005).

With regard to judgement of learning as an dependent variable, right before the pre-test and post-test for domain knowledge all learners were asked to estimate their score on both domain knowledge tests. Based on the comparison with their performance on the domain knowledge, a pre-test and post-test judgement of learning score was calculated.

In view of generating automated functional validity feedback (for both experimental conditions - see: Cues for calibration), learners' estimation of their future score was gathered through the judgement of learning cues proceeding each self-test. For both purposes, learners were asked to score their estimation of future score on an ordinal scale with ten point interval that range from 0/10 (0%) to 10/10 (100%). Their actual score was measured through a single test. Scores were recalculated to an identical ten point interval also ranging from 0/10 (0%) to 10/10 (100%). Following this measurement, the absolute accuracy index was calculated. The formula of this index can be found below:

$$\text{Absolute Accuracy Index } = \frac{1}{N}\sum_{i=1}^{N}(c_i - p_i)^2$$

In this formula, c_i corresponds to the estimation of future score and p_i corresponds to the actual score. Each deviation score between learners' estimation of future score and actual score is squared so it ranges from zero to one, where a score of zero corresponds to perfect calibration accuracy and a score of one corresponds to no accuracy. Smaller deviations correspond to better accuracy.

2.5.5.Metacognitive skilfulness

In line with the literature on cues for calibration, we investigated learners' metacognitive skilfulness to get more profound insights in the cues' effect on learners' self-regulated learning. This was done by using an online aptitude measurement developed by Veenman (e.g., Veenman, Bavelaar, De Wolf, & Van Haaren, 2014; Veenman, Wilhelm, & Beishuizen, 2004), called "The otter task". This measurement is a computerized learning-by-discovery task in Authorware. The otter-task requires learners to experiment with five independent variables in order to discover their (combined) effects on the growth of the otter population. The five variables were habitat, environmental pollution, public entrance, setting out new otter couples, and feeding fish in wintertime. Independent variables could have no effect on the otter population (public entrance), a main effect (habitat; pollution), and interact with another variable (habitat x setting out otter couples; pollution x feeding fish). For each experiment, participants could choose a value for the five variables by clicking on the pictograms on the left, and then order the computer to calculate the growth of the otter population. Results of experiments done were transferred to a storehouse where learners could scroll up and down to consult earlier results. After a minimum of fifteen experiments, an exit button occurs which allows the learners to leave "The otter task", nonetheless they are free to continue. All actions done by the learners are logged in a text file. This log file is scored for metacognitive skilfulness through ten log file indicators, namely: (1) number of experiments, (2) think time, (3) scroll down, (4) scroll up, (5) transition with one altered variable, (6) mean number of changes, (7) number of unique experiments, (8) variation of variables, (9) systematic changes, and (10) complete variation of variables. All learners' values per log file indicator were standardized into z-scores. Finally, mean z-scores were calculated over the ten log file indicators as an overall measure of metacognitive skilfulness (for a full account of the methodology, see: Veenman et al. (2014)). This calculation resulted in an individual metacognitive skilfulness score per learner, comparing learners' individual score with the sample score.

2.5.6.The quality of the instruments

Traditional reliability analysis (Cronbach's alpha) was used in order to investigate the quality of the measurement instruments. Table 2 depicts the Cronbach's alpha values of the different scales. Given the threshold of .70 as proposed by Nunnally and Bernstein (1994), all instruments seem to be in reach of this threshold.

Table 2.

Pre and post reliability analysis per construct.

Latent variable	***Construct***	***α pre***	***α post***
Cognition	Domain knowledge (DK) (10 items)	.65	.66
Goal orientation	Mastery goal approach (MGA) (6 items)	.79	.75
	Mastery avoidance approach (MAA) (4 items)	.73	.80
	Performance goal approach (PGA) (5 items)	.86	.86
	Performance avoidance approach (PAA) (6 items)	.74	.77
Academic self-concept	Learning effort (LE) (10 items)	.79	.80
	Learning confidence (LC) (7 items)	.74	.72
Metacognition	Metacognitive skilfulness (MS) (10 log file indicators)	.84	

2.6. Procedure

Learners' were randomly assigned to three separate but identical learning environments, either the control group, the functional validity feedback experimental condition (F-condition), or the functional and cognitive validity feedback experimental condition (FC-condition). All (n=151) learners attending the module were invited to complete the otter task (during four available timeslots) prior to their first login in the online learning module. The learners got 60 minutes to complete this task. The online pre-test questionnaire, the pre-test judgement of learning question, and prior domain knowledge test were administered at the start of the online learning module and obligatory to activate the content of the online learning module. Learners got 28 days' time to complete the online learning module, learners in the experimental conditions received cues for calibration (as described earlier) during that time; learners in the control condition did not. After the completion of the intervention, learners in the three conditions completed the online post-test questionnaire, the post-test judgement of learning question, and the domain knowledge test. The learners did not receive any other form of instruction on the module content during the time period between the pre-test and the post-test. For the matching of the different datasets anonymized student IDs were used.

2.7. Analysis

First, to be able to determine the effect of learners' metacognitive skilfulness as an independent variable, we quartered the learners based on their metacognitive skilfulness score. This was by done by ordering all learners' scores from the lowest to the highest, followed by dividing them in 4 groups. Each group represented 25% of the sample. Learners were assigned a quartile number indication in which quartile their score was situated (1 = low to 4 = high). In this way, a new categorical variable metacognitive skilfulness quartile membership (PMSQ) was created and will be used as an independent variable throughout the analyses.

Second, descriptive statistics were calculated presenting the number of subjects involved (*N*), the minimum (*Min*) and maximum (*Max*) scores, the mean scores (*M*), the standard error (*SE*), the standard deviation (*SD*), and the variance (σ^2). This was done for learners' metacognitive skilfulness quartile membership (PMSQ) and each of the pre-test variables (domain knowledge (DK), mastery goal approach (MGA), mastery avoidance approach (MAA), performance goal approach (PGA), performance avoidance approach (PAA), learning effort (LE), learning confidence (LC), and judgement of learning (JOL)). Also the post-test variables were investigated for correlations, to identify the need for multivariate or univariate tests.

Third and final, through a one-way multivariate analysis of variance (MANOVA) with condition as independent variable, followed by univariate analyses of variance (ANOVAs) the three conditions' comparability among each other was checked for learners' prior domain knowledge, goal-orientation, academic self-concept, and metacognitive skilfulness quartile membership.

2.7.1.Investigation of learning behaviour

The event sequence analysis consisted of two major steps (e.g., Cicchinelli et al., 2018; Zhou, 2016). First frequent event sub-sequences were identified using exploratory sequence analysis. Secondly, discriminant frequent event sub-sequences were identified by using an explanatory approach as well. The latter analysis was based on the condition learners were in. This to identify what sub-sequences (dependent variable) occurred significantly more in which condition (independent variable). A similar approach was adopted for metacognitive skilfulness (PMSQ) and for the interaction between condition and metacognitive skilfulness. The learners' behavioural data was imported in R-statistics and analysed using the TraMineR package (Gabadinho, Ritschard, Mueller, & Studer, 2011).A similar approach to identify frequent event sub-sequences was used as Jovanović, Gašević, Dawson, Pardo, and Mirriahi (2017) and Van Laer and Elen (2016). Both studies emphasize the importance of two parameters when identifying frequent event sub-sequences. The first one is the time constraint (Studer, Mueller, Ritschard, & Gabadinho, 2010). As we followed a data-driven approach while investigating the ecological order of events, we chose to set this parameter on one. This indicates that only events that actually occurred following each other are included. Events further apart in time are not considered. The second one is the relative threshold number of times (pMinSupport) a sub-sequence occurs among the different learners (Müller et al., 2010). In this study, this parameter was arbitrarily set on .25 to assure frequent sub-sequences occurred at least in 25% of the learners.

Discriminant frequent event sub-sequence were identified in line with Kim and Shute (2015) and with Grover et al. (2017). The significant discriminating ability of the sub-sequences was first based on differences between conditions learners were in, secondly on metacognitive skilfulness, and finally on the interaction of the condition learners were in and learners' metacognitive skilfulness quartile membership (PMSQ). To be able to calculate the discriminating abilities of a frequent sub-sequence two arguments are needed (a) a sub-sequence (subseq) object containing the sub-sequences considered for discriminating the groups and (b) the variable that defines the groups (groups) (Garza, 2016). A chi-square test is used to investigate the significance of the relationship between the observed and expected occurrence of a frequent sub-sequence for each value of the measured variables (Studer et al., 2010). Finally, the effect sizes are calculated using the Cramer's V. The Cramer's V expresses the relationship between a certain discriminating frequent sub-sequence and the learners' characteristics and is reported in a value between zero and one. The closer to one the higher the relation. Cohen (1988) refers to small (≤.30), medium (≥.30 and ≤50), and large (≥.50) effect sizes.

2.7.2.Investigation of learning outcomes

In order to (1) examine the effect of the instructional intervention on learners' learning outcomes and (2) examine the interaction effect of instructional intervention and learners' metacognitive skilfulness on learners' learning outcomes, a two-way multivariate analysis of covariance (MANCOVA) test with pre-test and post-test data will be used. The MANCOVA can be seen as an extension of the multivariate analysis of variance (MANOVA) incorporating pre-test covariates. These covariates are related to the dependent post-test variables under investigation and reduce the error variance between pre-test and post-test results (Dimitrov & Rumrill, 2003). A MANCOVA is used to determine whether there are any statistically significant differences between the adjusted means of three or more independent (unrelated) groups, having controlled for the pre-test covariates.

In this study a MANCOVA will be done with as dependent variables learners' post-test scores on: domain knowledge (DK), mastery goal approach (MGA), mastery avoidance approach (MAA), performance goal approach (PGA), performance avoidance approach (PAA), learning effort (LE) learning confidence (LC), and judgement of learning (JOL) . Condition and learners' metacognitive skilfulness (PMSQ) will be used as independent variables. Learners' pre-test scores on domain knowledge (DK), mastery goal approach (MGA), mastery avoidance approach (MAA), performance goal approach (PGA), performance avoidance approach (PAA), learning effort (LE), learning confidence (LC), and judgement of learning (JOL) are used as covariates. The main effects, followed by the interaction effects, and univariate tests will be reported. Additionally, pairwise comparison with a Bonferroni correction will further investigate the found effects. Nonetheless, before conducting the MANCOVA test, the variables were tested for normality (Shapiro–Wilks' test), sphericity (Mauchly's Test of Sphericity), and homogeneity of variances (Levene's test). Figure 2, visualizes the MANCOVA.

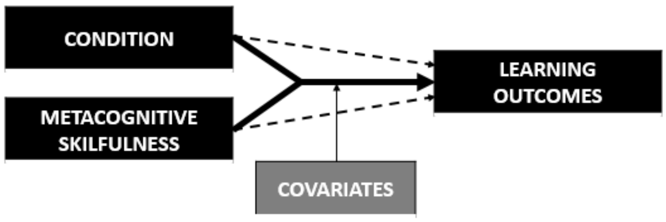

Figure 2. Visual representation of the MANCOVA analysis.

3. Results

3.1. Descriptive statistics

151 learners were included in the analysis. There were 134 women (88.74%) and 17 men (11.26%) aged between 19 and 58 years (*M*=21.87, *SD*= 6.84). Learners were assigned to a quartile, based on their score on metacognitive skilfulness. The first quartile (n = 38) represented Z-scores between -1.26 and -.61 (*M* = -.91, *SD* = .208), the second quartile (n = 38) between -.60 and -.15 (*M* = -.39, *SD* = .157), the third quartile (n = 39) between -.14 and .36 (*M* = .11, *SD* = .160), and the fourth quartile (n = 37) between .39 and 5.38 (*M* = 1.36, *SD* = 1.121). Descriptive statistics of learners' pre-test scores of the dependent variables can be found in table 3.

Table 3.

Descriptive statistics of the pre-test scores.

Pre-test variables	*N*	*Min*	*Max*	*M*	*S.E.*	*SD*	σ^2
Domain knowledge (PDK)	151	0.00	10.00	4.48	0.188	2.306	5.318
Mastery goal approach (MGA)	151	2.00	5.00	4.07	0.047	0.575	0.331
Mastery avoidance approach (MAA)	151	1.00	5.00	3.68	0.062	0.759	0.577
Performance goal approach (PGA)	151	1.00	4.20	2.37	0.060	0.751	0.564
Performance avoidance approach (PAA)	151	1.80	5.00	3.79	0.051	0.633	0.401
Learning confidence (LC)	151	1.57	4.43	3.02	0.040	0.498	0.249
Learning effort (LE)	151	1.90	4.90	3.45	0.041	0.510	0.261
Judgement of learning (JOL)	151	0.00	0.49	0.07	0.007	0.087	0.008

To identify if multivariate or univariate tests for further investigation would be most appropriate, relationships between the dependent variables were checked. Correlation analysis showed weak (.20-.39) to moderate (.40-.59) correlations (Evans, 1996) between the different post-test variables, namely between PGA and LC (r(149) = .36, p < .001), PAA and MGA (r(149) = .16, p = .043), PAA and MAA (r(149) = .51, p < .001), PAA and LC (r(149) = -.16, p < .01), and MAA and LE (r(149) = .38, p < .001). No other correlations were found.

3.2. Pre-test comparison of the experimental and control conditions

All 151 learners (control = 49, F-condition = 48, and FC-condition = 54) participated in the otter task, the prior judgement of learning question, the prior domain-knowledge test, and the pre-test questionnaire. The pre-test scores correlated weakly too moderately, so a multivariate analysis of variance (MANOVA) was applied to compare learners' pre-test scores (as dependent variables) for the three condition (independent variable). The MANOVA showed no significant differences (F (16, 282) = 1. 59, Wilk's Λ = .933, p = .71) for experimental and control condition. Nonetheless, the univariate tests showed a significant differences for domain knowledge (DK) (F (2, 151) = 3.42, p = .035) among the different conditions. Learners in the FC-condition seemed to score significantly (p = .012) higher (*MD* = 1.15) than learners' in the F-condition. No difference was found for the other variables.

3.3. The effect of condition and metacognitive skilfulness on learners' learning behaviour

3.3.1.Condition

The learners (n=149) included in the event sequence analysis generated 54434 events over the timespan of 28 days. A total of 249 frequent sub-sequences were extracted (TimeGap=1; pMinSupport=.25). 18 significant discriminant sub-sequences (pValueLimit=.05) were identified. Sub-sequences contained between two and seven events. Three conditions were compared (control condition, F-condition, and FC-condition) through chi-square tests.

Learners in the control condition made significantly the most use of sub-sequences consisting of Self-test events followed by other Self-test events (between χ(2) = 83.848, p < .001, V= .56 and χ(1) = 98.706, p < .001, V= .66). Whereas for the control condition the standardized residuals scores were between 6.50 and 7.00, for the F-condition they were between -2.97 and -3.35, and for the FC-condition between -3.37 and -3.51. Examination of the Cramer's V scores indicate, according to Cohen (1988), large effects (≥.50). Learners in the control condition also made significantly the most use of sub-sequences consisting of Self-test events followed by Exercises events. Here chi-square tests (between χ(2) = 11.964, p < .001, V= .08 and χ(1) = 16.092, p < .001, V= .11) showed smaller effect sizes

(≤.30). The standardized residuals for the control condition were between 2.22 and 2.50, for the F-condition between -.04 and -.42, and for the FC-condition between -1.96 and -2.33. Learners in the F-condition (*SR* between .51 and 1.72) used significantly more sub-sequences related to Exercise events followed by other Exercise events (χ^2-tests between $\chi(2) = 6.004$, $p < .001$, $V = .04$ and $\chi(1) = 6.546$, $p < .001$, $V = .05$). The control condition's standardized residuals score were between -1.23 and .82 and the FC-condition's between -1.35 and -.15. Finally, learners in the FC-condition did never demonstrated significantly more sub-sequences to any types of events for all significant discriminant sub-sequences.

3.3.2.Metacognitive skilfulness

In line with the investigation of the effect of condition on learners' learning behaviour, the effect of metacognitive skilfulness (PMSQ) was also studied. Significant discriminant sub-sequences, based on learners' metacognitive skilfulness (PMSQ) (quartile 1 to quartile 4) were identified. The same analysis was applied as when addressing the effect of the condition on learners' learning behaviour. Only 3 significant discriminant sub-sequences (pValueLimit=.05) were identified. Sub-sequence contained between four and six events.

The three discrimant sub-sequences, all showed Text events (downloading of one of the two articles learners needed to read) followed by Forum events. Learners belonging to the lowest quartile (Q1) used these sub-sequences significantly more events (between $\chi(3) = 7.814$, $p = .050$, $V = .23$ and $\chi(3) = 8.069$, $p = .044$, $V = .82$) than learners belonging to quartile 2 (*SR* between .22 and .36), quartile 3 (*SR* between -.02 and -.07), or quartile 4 (*SR* between -1.89 and -1.99). Examination of the Cramer's V scores indicate according to Cohen (1988) small effects of learners' metacognitive skilfulness on learners' learning behaviour (≥.50).

3.3.3.Condition and metacognitive skilfulness

Finally, to investigate on the interaction of condition and metacognitive skilfulness (PMSQ) on learners' learning behaviour, significant discriminant sub-sequences based on condition and metacognitive skilfulness (PMSQ) (1 = low to 4 = high) were identified. 10 significant discriminant sub-sequences (pValueLimit=.05) were identified and compared among 12 groups (3 conditions x 4 quartiles). Sub-sequence contained between three and six events.

Learners in the control condition belonging to the lowest quartile (Q1) of metacognitive skilfulness made significantly the most use of sub-sequences consisting of Self-test events followed by other Self-test events (between $\chi(11) = 86.47$, $p < .001$, $V = .76$ and $\chi(11) = 101.280$, $p < .001$, $V = .82$). Examination of the Cramer's V scores indicate according to Cohen (1988) large effects based on condition and PMSQ (≥.50). Here standardized residuals score were between 4.41 and 4.68. Their counterpart belonging to different quartiles of PMSQ, but to the same condition, used fewer such sub-sequences. For learners belonging to the second quartile (Q2), standardized residuals score were between 2.50 and 3.75, for learners belonging to the third quartile (Q3), between 2.94 and 3.41, and for the highest quartile (Q4), between 2.36 and 2.64. The same observation could be made for learners belonging to the lowest quartile (Q1) in the F-condition (*SR* between -.97 and -1.64) and the FC-condition (*SR* between -1.75 and -1.88). Learners belonging to higher quartiles in the F-condition or the FC-condition made less use of sub-sequences consisting of Self-test events followed by other Self-test events. Learners in the FC-condition belonging to the second quartile (Q2) used the least of these sub-sequences (SR between -1.89 and -2.03).

Learners in the F condition belonging to highest quartile (Q4) made significantly more use of sub-sequences consisting of Self-test events followed by Exercises events. Here Chi squared tests (between $\chi(11) = 20.774$, $p = .036$, $V = .37$ and $\chi(11) = 23.431$, $p = .015$, $V = .40$) showed medium effect sizes (between .30 and .50). The standardized residuals were between .83 and 1.16, where they were for their counterparts belonging to the third quartile (Q3), between -.83 and -1.27, for learners belonging to the second quartile (Q2), between .38 and .75, and for the lowest quartile (Q1), between -.97 and -1.26. The same observation with regard to sub-sequences consisting of Self-test events followed by Exercises events could be made for learners belonging to highest quartile (Q4) of metacognitive skilfulness in the control condition (*SR* between 1.25 and 1.71), but not for learners belonging to this quartile in the FC-condition (*SR* between -1.75 and -1.99). Learners belonging to lower quartiles in the control condition or the FC-condition made less use of sub-sequences consisting of Self-test events followed by Exercise events. Learners in the FC-condition belonging to the highest quartile (Q4) used the least of these sub-sequences (SR between -1.75 and -1.99).

Finally, learners in the highest quartile (Q4) and the FC-condition did never demonstrated significantly more sub-sequences to any types of events for all significant discriminant sub-sequences.

3.4. The effect of condition and metacognitive skilfulness on learners' learning outcomes

For the multivariate tests, the main effect of condition (F (16, 150) = .908, p = .562, Wilk's Λ = .831, η_p^2 = .09) was not significant, indicating that condition had no direct effect on the dependent variables under investigation. In contrast to this, the main effect of learners' pre-test metacognitive skilfulness (PMSQ) was significant (F (24, 218.124) = 1.987, p = .005, Wilk's Λ = .564, η_p^2 = .17), indicating that a different degree of skilfulness affects the dependent variables. Here, univariate tests showed that only learners' judgement of learning (F (3, 138) = 6.025, p = .001, η_p^2 = .18) significantly differed depending on learners' degree of skilfulness. Pairwise comparisons using a Bonferroni correction showed that learners in the third quartile (Q3) (*M* = .105) scored significantly less accurate, than learners in the lowest quartile (Q1) (*MD* = -.03, *SE* = .025, *p* = .023), the second quartile (Q2) (*MD* = -.02, *SE* = .025, *p* = .022), or in the highest quartile (Q4) (*MD* = -.03, *SE* = .025, *p* = .011). Figure 3 shows the mean post-test scores for judgement of learning per metacognitive skilfulness quartile.

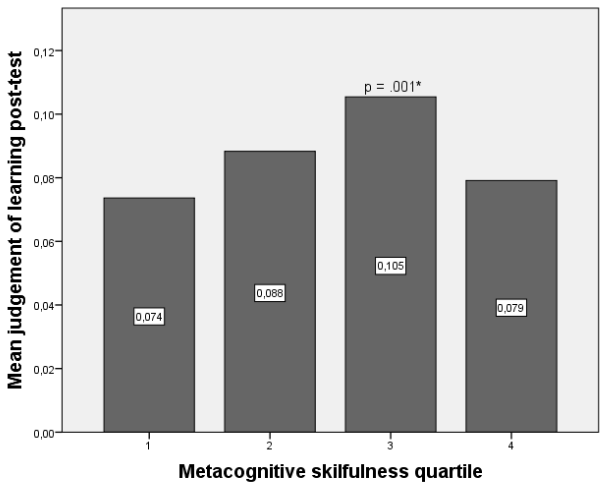

Figure 3. Mean judgement of learning post-test scores per metacognitive skilfulness quartile.

Results also reveal a significant interaction effect between condition and metacognitive skilfulness (*F* (48, 373.094) = 1.560, *p* = .025, Wilk's Λ = .591, η_p^2 = .09). The univariate tests showed learners' judgement of learning (*F* (6, 138) = 3.862, *p* = .002, η_p^2 = .13) and learning confidence (*F* (6, 138) = 1.474, *p* = .017, η_p^2 = .17) significantly differed depending on the condition learners were in and learners' degree of metacognitive skilfulness.

Post-hoc comparison using a Bonferroni correction showed that learners belonging to quartile 3 in the FC-condition (*M* = .25) scored significant less accurate than learners in the control condition (*MD* = -.11, *SE* = .051, *p* = .103) or F-condition (*MD* = -.17, *SE* = .042, *p* < .001) belonging to the same quartile, indicating that they were less accurate. The opposite was found for learners belonging to the highest quartile (Q4) in the FC-condition (*M* = .04). Learners belonging to the control condition (*MD* =.10, *SE* = .048, *p* = .033) or the F-condition (*MD* = .01, *SE* = .042, *p* = .753) and the same quartile, scored less accurate. The other conditions and quartiles configurations did not show any significant differences with regard to learners' judgement of learning. Figure 4 shows the estimate marginal means for judgement of learning per condition and metacognitive skilfulness quartile.

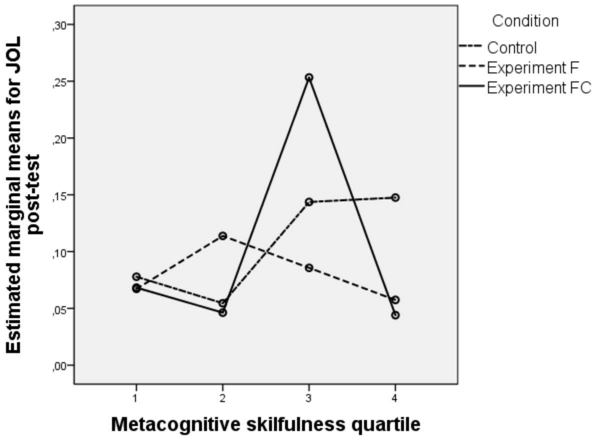

Figure 4. Estimate marginal means for judgement of learning per condition and metacognitive skilfulness quartile.

Finally, for learning confidence, post-hoc comparisons using a Bonferroni correction showed that learners with the lowest degree of metacognitive skilfulness (Q1) in the F-condition (*M* = 2.98) scored lowest on learning confidence. Learners belonging to the same quartile in the control condition (*MD* = .25, *SE* = .125, *p* = .050) or FC-condition (*MD* = .03, *SE* = .108, *p* = .804) scored higher on learning confidence. The opposite was found for learners belonging to the highest quartile (Q4) in the F-condition (*M* = 3.40) who scored highest on learning confidence. Learners belonging to the same quartile in the control condition (*MD* = -.37, *SE* = .137, *p* = .009) or FC-condition (*MD* = -.18, *SE* = .129, *p* = .178) scored lower on learning confidence. No other significances were found in the univariate tests for learning confidence. Figure 5 shows the estimate marginal means for learning confidence per condition and metacognitive skilfulness quartile.

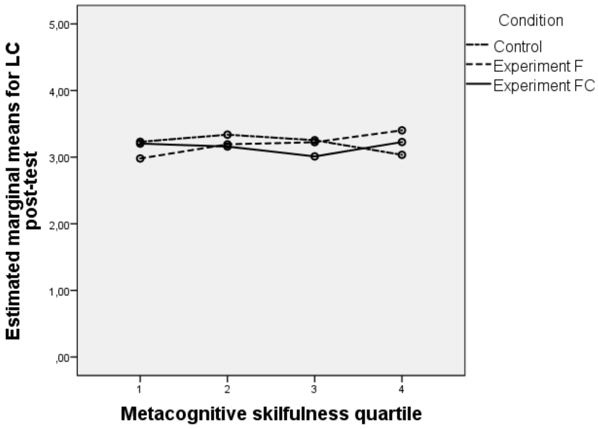

Figure 5. Mean judgement of learning post-test scores per condition and metacognitive skilfulness quartile.

4. Discussion

In what follows, first we relate the results to the hypotheses set in the introduction, secondly we explore the nature of the results and discuss them in terms of their theoretical and practical implications and provide recommendations for future research.

4.1. Findings

4.1.1. *Hypothesis 1: "Cues for calibration affect learners' learning behaviour."*

Cues for calibration clearly seem to affect learners' learning behaviour, so hypothesis one could be confirmed. Learners in the control condition made significantly more use of sub-sequences consisting of Self-test events followed by other Self-test events compared to the learners in the experimental conditions. Learners in the F-condition seemed to use such sub-sequences significantly less, and even less so by learners in the FC-condition. Similar findings were found in relation to sub-sequences consisting of Exercises events following Self-test events. Learners in the F-condition only used significantly more sub-sequences related to Exercise events following other Exercise events compared to the learners in the control condition, and learners in the FC-condition who used the sub-sequences the least. For all 18 significant discriminant sub-sequences learners in the FC-condition, never demonstrated significantly more use of sub-sequences compared to the other two conditions.

Additionally, independent from the condition also the effect of learners' metacognitive skilfulness was investigated. Here there seemed to be hardly any differences among learners. Learners with the lowest degree of metacognitive skilfulness (Q1) made significantly more use of sub-sequences related to Text events followed by Forum events. The higher learners' score was for metacognitive skilfulness, the fewer they used these sub-sequences.

4.1.2.Hypothesis 2: "Cues for calibration affect learners' learning behaviour differently for learners with different levels of metacognitive skilfulness."

Based on the following elements hypothesis 2 can be regarded to be confirmed. Learners in the control condition belonging to the lowest degree quartile (Q1) made the most use of sub-sequences consisting of Self-test events followed by other Self-test events compared to learners in the other groups.

Learners in the same condition but with other degrees of skilfulness used these sub-sequences less frequently. Learners belonging to the same quartile in the F-condition and the FC-condition exhibited the same behaviour. Learners belonging to different quartiles made less use of sub-sequences consisting of Self-test events followed by other Self-test events. Learners in the FC-condition belonging to the second quartile (Q2) used these sub-sequences the least.

Learners in the F condition belonging to the highest quartile (Q4) of metacognitive skilfulness made significantly more use of sub-sequences consisting of Self-test events followed by Exercises events. Learners belonging to the highest quartile (Q4) in the control condition exhibited the same behaviour. Learners belonging to different quartiles in the control condition or the FC-condition made less use of sub-sequences consisting of Self-test events followed by Exercise events. Learners in the FC-condition belonging to quartile 4 used the least of these sub-sequences.

Finally, learners with the highest degree (Q4) of metacognitive skilfulness in the FC-condition did never demonstrated significantly more sub-sequences to any types of events for all significant discriminant sub-sequences.

4.1.3.Hypothesis 3: "Cues for calibration positively affect learners' learning outcomes."

No significant main effect of condition on learners' learning outcomes could be found. Therefore, hypothesis 3 is falsified. However, learners' metacognitive skilfulness significantly affected learners' learning outcomes. Further univariate analyses have revealed this only to be the case for learners' judgement of learning. Learners with metacognitive skilfulness degrees between -.14 and .36 (Q3) judged their learning significantly less accurate on the post-test than learners with other levels of metacognitive skilfulness. Explorative analysis showed indications of a curvilinear relationship between metacognitive skilfulness and judgement of learning. Figure 6 shows this relationship.

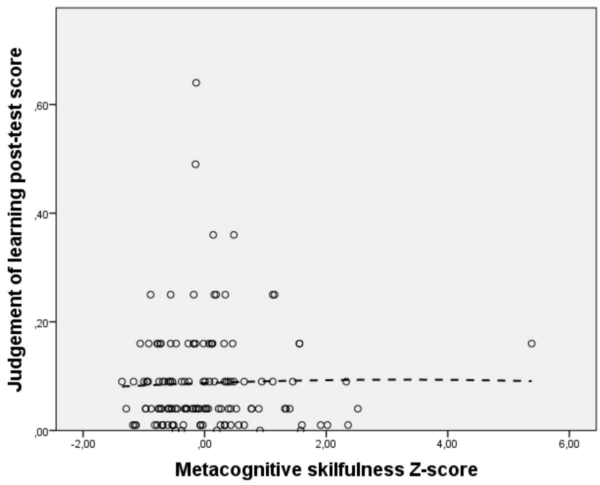

Figure 6. Curvilinear relationship between metacognitive skilfulness and judgement of learning.

4.1.4.Hypothesis 4: "Cues for calibration positively affect learners' learning outcomes most when learners have high levels of metacognitive skilfulness."

Finally, based on the following elements, hypothesis 4 can said to be confirmed although in an unexpected direction. Investigating the interaction effect of the cues for calibration (conditions) and learners' metacognitive skilfulness (PMSQ) on learners' learning behaviour, it became clear that certain constellations of condition and degrees of learners' metacognitive skilfulness significantly affected learners' learning outcomes. Nonetheless, this was the case only for two dependent variables. Univariate tests showed learners' judgement of learning and learning confidence to be affected by the interaction of both independent variables. With regard to judgement of learning, results showed that learners with metacognitive skilfulness degrees between -.14 and .36 (Q3) in the FC-condition scored less accurate, than learners in the control condition or F-condition belonging to the same quartile. For learners with high metacognitive skilfulness degrees between .39 and 5.38 (Q4) of the FC-condition, the results showed the opposite. These learners were more accurate than the others were.

In relation to learning confidence, results showed that learners with low degrees of metacognitive skilfulness between -1.26 and -.61 (Q1) in the F-condition scored lowest on learning confidences. Learners belonging to the same quartile in the control condition or FC-condition scored higher. The opposite result was found for learners with high degrees of metacognitive skilfulness between .39 and 5.38 (Q4) of the F-condition. These learners scored higher on learning confidence than learners in the same quartile for the other conditions.

4.2. Exploration of the unexpected findings

The current study yields three major findings. First, the more external feedback learners get and the higher their metacognitive skilfulness, the fewer sub-sequences learners use related to self-tests and exercises. Secondly, when learners have low degrees of metacognitive skilfulness and receive cues for calibration through functional validity feedback, they score significantly lower on learning confidence. The opposite is true in the F-condition for learners' with high degrees of metacognitive skilfulness. Learners are more confident. When both functional and cognitive validity feedback are provided, no differences are found for any learners. Third and final, learners with high degrees of metacognitive skilfulness who receive both functional and cognitive validity feedback are more accurate in judging their own learning than other learners. Nonetheless, this is not the case for learners with average degrees of metacognitive skilfulness. In conclusion, this study shows that cues for calibration, affected learners' learning behaviour and outcomes, and so self-regulated learning. However, the directions are unexpected. Below we provide a possible explanation.

4.2.1. *Learners' cue use*

In line with current research on the link between instructional interventions and learners' learning behaviour through log-file data (e.g., Rienties, Toetenel, & Bryan, 2015; Wolff, Zdrahal, Nikolov, & Pantucek, 2013), the type of calibration cue learners received influenced their learning behaviour. The observation that learners in the control condition made significantly more use of certain sub-sequences compared to the other conditions was rather striking as it contrasts with research reporting greater learner involvement with the learning environment and cues provided, when cues for calibration are provided (Szabo, Falkner, Knutas, & Dorodchi, 2018; Timmers, Walraven, & Veldkamp, 2015). However, from a self-regulated learning theory perspective a decrease in particular learning behaviour might be explained as follows. When learners are capable to identify the instructional requirements set, to comply with them, and so to be successful in achieving the learning outcomes targeted, they are self-regulated learners (e.g., Wolters, Won, & Hussain, 2017). In line with this reasoning the fewer actions needed to achieve the outcomes targeted the more effective one's self-regulated learning is (Winne & Hadwin, 1998). When cues for calibration are provided, including functional and cognitive validity feedback (FC-condition), self-regulated learners direct their behaviour towards the information that helps them to achieve the learning outcomes targeted (e.g., Butler & Winne, 1995; Dunlosky & Thiede, 2013; Geitz, Joosten-ten Brinke, & Kirschner, 2016; Rienties & Toetenel, 2016). When instead learners are only provided with indications about their calibration efforts (F-condition) they might act on this feedback and adapt their behaviour by for example making more exercises in an attempt to progress (e.g., Tempelaar, Rienties, & Giesbers, 2015). Finally, when in contrast learners are not provided with any information about their calibration efforts (control condition), it is solely up to them to gather this information, potentially resulting in feedback-seeking behaviour (e.g., Harrison, Könings, Schuwirth, Wass, & van der Vleuten, 2015). In conclusion, as a result of providing learners with functional and cognitive validity feedback, highly metacognitive skilful learners might be selective and only engage in specific goal-directed behaviour, whereas learners struggling with controlling their learning might rather perform a plenty-fold of undirected behaviours (e.g., Fonseca, Martí, Redondo, Navarro, & Sánchez, 2014; Van Laer & Elen, 2016).

4.2.2.Learners' cue interpretation

As cues of calibration are inevitably interpreted through the lens of one's self-perceptions, it is important to understand how learners interpret the information provided to them (Eva et al., 2012). One way of doing this, is through the observation of changes in learners' learning outcomes. The findings of the study presented show that learners receiving functional validity feedback and having low degrees of metacognitive skilfulness, scored significantly lower for learning confidence in contrast to their counterparts in the control condition. According to research on the effect of cues for calibration on learners' learning confidence (e.g., Van der Kleij, Feskens, & Eggen, 2015), when learners are confronted with functional validity feedback, learners' might interpret the feedback for example as an indicator of personal failing or looming problem, rather than as a cue for them to re-calibrate (Hattie & Timperley, 2007). Especially when learners have low degrees of learning confidence this might be decisive for their further use of cues, as they might relate cues with negative experiences (Levine & Donitsa-Schmidt, 1998). In line with this reasoning, research points out that learners provided with functional and cognitive validity feedback do not have this problem, as cognitive validity feedback directs them to appropriate action to overcome this feeling (e.g., Ridder et al., 2015). Functional validity feedback did not only affect learners with low degrees of metacognitive skilfulness negatively, functional validity feedback also led to increased learning confidence for learners with high degrees of metacognitive skilfulness. Nonetheless, in the light of calibration, learning confidence without any performance related increase might lead to overestimation of one's own capabilities and further down the road to a decrease in performance (e.g., Dunlosky & Rawson, 2012).

4.2.3.Cue's effectiveness for increased performance

In line with our findings, Hellrung and Hartig (2013) present, in their systematic literature review, a substantial body of literature reporting increased learners' judgement of learning evoked by the use of functional and cognitive validity feedback. However, this was only the case for learners with high degrees of metacognitive skilfulness. One possible explanation for this is that learners with high degrees of metacognitive skilfulness are more aware of the different underlying strategies potentially supporting calibration and re-calibration (e.g., Hacker, Bol, & Bahbahani, 2008). This would result in more accurate estimations of one's performance (Callender, Franco-Watkins, & Roberts, 2016). So the information learners receive on the accuracy of their perceived level of performance in relation to their actual level, helps them to re-calibrate (e.g., Muis et al., 2016; Winne & Jamieson-Noel, 2002). Although for a change in accuracy to occur, learners need insight into the cognitive processes needed to calibrate their learning (e.g., Alexander, 2013; Dunlosky & Thiede, 2013). The combination of functional and cognitive validity feedback proved to provide calibration and showed increased judgement of learning. Based on the results of accurate calibration, learners monitor their learning and select cognitive and metacognitive strategies (e.g., error correction strategies, revision activities, etc.) which may help to proceed them in the direction of the desired level of performance (Narciss, 2017). The finding that certain learners increased in judgement of learning but not in performance might relate to the latter. Even when learners can calibrate external and internal feedback, they might not possess or be able to recall the cognitive and/or metacognitive strategies needed to act in a way that will produce increased performance (e.g., Pintrich, 2002; Veenman, 1993). This would evoke sub-optimal self-regulated learning and hamper increased performance. Although we investigated the effect of learners' metacognitive skilfulness based on learners' domain general ability to control and apply cognitive and metacognitive processes, the cues for calibration provided might have lacked the potential to evoke the transfer of these processes to a domain specific context (Butler & Winne, 1995). This finding has been supported by prior findings (e.g., Ardasheva et al., 2017; Dinsmore & Fryer, 2018) indicating that cues on the use of cognitive and/or metacognitive strategies should strongly align with the content provided (e.g., Alexander, 2018; Tricot & Sweller, 2014).

5. Further directions and conclusions

The present study documents fine-grained insights into the relationship between learners' self-regulated learning and cues for calibration. To obtain these insights, we first investigated learners' calibration-cue use based on learners' individual differences. Secondly, we operationalized self-regulated learning through learners' learning behaviour and outcomes. Investigating both learning behaviour and outcomes provides insights on learners' self-regulated learning, as well as on the nature of cues' effects. The current study reveals that differences in learner behaviour were related to condition and learners' metacognitive skilfulness, thus establishing a link between learners' self-regulated learning and cues for calibration. Finally, in the discussion of the results, we unravelled the effect of the design and content of the cues for calibration provided and hypothesized that when cues for calibration are provided through functional and cognitive validity feedback, learners' calibration capabilities will increase. Yet for this to result in goal-directed self-regulated learning and so increased achievement, learners not only need to be supported in identifying and recalling the cognitive and metacognitive strategies needed, but also directed to how to apply the cognitive and metacognitive strategies in their context.

5.1. Further directions

To further enrich our understanding, some challenges need to be addressed. A first challenge is the sample-size. A total of 151 learners were involved in the study – 48 in the control condition, 49 in the F-condition, and 54 in the FC-condition, which according to Field (2013) is an appropriate rule of thumb for testing the effect of three conditions (≥ 30 participants per condition). Nonetheless, for testing the interaction between condition and metacognitive skilfulness, more learners per condition might be advisable and so the power of some of the presented statistics might be debatable. A second challenge relates to the use of a data-driven approach to analyse learners' learning behaviour and its arbitrary parameter setting. As theoretical insights can be derived from the results of data-driven trials, contributing to such an approach may prove more promising than, for example, recoding events as (covert) metacognitive strategies or activities. In further research, this data-driven approach might be explored by experimenting with different parameter settings or a combination of data-driven and theory-driven approaches might be taken. With regard to the latter, this could be achieved for example by recoding events or sub-sequences based on a theoretical framework unrelated to self-regulation theory (for example a tool-use scheme). This would make the sub-sequences identified more meaningful and so interpretable. The third and final challenge relates to the relation between learners' self-regulated learning and the effect of cues. In this study the design of the cues and learners' metacognitive skilfulness were related to learners' learning behaviour. To be able to identify meaningful learning behaviour in the light of learning outcomes, future research might want to model the path of how different types of learners use the cues for calibration provide, leading them to certain learning outcomes for example through Hidden Markov Modelling.

5.2. Conclusions

Given the current lack of certainty regarding the effects of cues for calibration on learners' self-regulated learning, teachers and instructional designers remain dependent on inconsistent conceptual claims that cues for calibration may improve self-regulated learning. Studies such as the one presented here help both researchers and practitioners to distinguish between the effect of cues for calibration and how learners react to them. Establishing more fine-grained links between learners' characteristics, learning behaviour, and learning outcomes could help us propel the investigation of the effect of cues in intervention research.

6. Acknowledgements

Deleted for review.

7. Bibliography

Abbott, A. (1995). Sequence analysis: new methods for old ideas. *Annual review of sociology, 21*(1), 93-113. doi: 10.1146/annurev.so.21.080195.000521

Alexander, P. A. (2013). Calibration: What is it and why it matters? An introduction to the special issue on calibrating calibration. *Learning and instruction, 24*, 1-3.

Alexander, P. A. (2018). Looking down the road: Future directions for research on depth and regulation of strategic processing. *British Journal of Educational Psychology, 88*(1), 152-166.

Anderson, L. W. (2005). Objectives, evaluation, and the improvement of education. *Studies in educational evaluation, 31*(2-3), 102-113.

Ardasheva, Y., Wang, Z., Adesope, O. O., & Valentine, J. C. (2017). Exploring Effectiveness and Moderators of Language Learning Strategy Instruction on Second Language and Self-Regulated Learning Outcomes. *Review of Educational Research, 87*(3), 544-582. doi: 10.3102/0034654316689135

Ariel, R., & Karpicke, J. D. (2018). Improving self-regulated learning with a retrieval practice intervention. *Journal of Experimental Psychology: Applied, 24*(1), 43.

Arrastia-Chisholm, M. C., Torres, K. M., & Tackett, S. (2017). Using Reflection to Increase Self-Regulation Among Pre-Service Teachers. *Fostering Reflective Teaching Practice in Pre-Service Education*, 148.

Azevedo, R., Martin, S. A., Taub, M., Mudrick, N. V., Millar, G. C., & Grafsgaard, J. F. (2016). *Are Pedagogical Agents' External Regulation Effective in Fostering Learning with Intelligent*

Tutoring Systems? Paper presented at the International Conference on Intelligent Tutoring Systems.
Balzer, W. K., & Doherty, M. E. (1989). Effects of cognitive feedback on performance. *Psychological bulletin, 106*(3), 410.
Bannert, M., Molenaar, I., Azevedo, R., Järvelä, S., & Gašević, D. (2017). *Relevance of learning analytics to measure and support students' learning in adaptive educational technologies.* Paper presented at the Proceedings of the Seventh International Learning Analytics & Knowledge Conference.
Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short-and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52*, 293-306.
Besser, E. D. (2016). *Exploring the role of feedback and its impact within a Digital Badge system from multiple perspectives: A case study of preservice teachers.* Purdue University.
Boekaerts, M. (1999). Self-regulated learning: Where we are today. *International journal of educational research, 31*(6), 445–457. doi: 10.1016/S0883-0355(99)00014-2
Boelens, R., Van Laer, S., De Wever, B., & Elen, J. (2015). Blended learning in adult education: towards a definition of blended learning (pp. 1-3). Brussels: Genth University.
Bol, L., & Hacker, D. J. (2001). A comparison of the effects of practice tests and traditional review on performance and calibration. *The Journal of Experimental Education, 69*(2), 133-151.
Bonk, C. (2017). *Best Practices for Online and Blended Learning: Introducing the R2D2 and TEC-VARIETY Models*. Paper presented at the Fall Faculty Development Conference, Indianapolis.
Broadbent, J., & Poon, W. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The internet and higher education, 27*, 1-13.
Bui, T., & Loebbecke, C. (1996). Supporting cognitive feedback using system dynamics: a demand model of the global system of mobile telecommunication. *Decision Support Systems, 17*(2), 83-98.
Butler, D. L., & Winne, P. H. (1995). Feedback and Self-Regulated Learning: A Theoretical Synthesis. *Review of Educational Research, 65*(3), 245-281. doi: 10.2307/1170684
Callender, A. A., Franco-Watkins, A. M., & Roberts, A. S. (2016). Improving metacognition in the classroom through instruction, training, and feedback. *Metacognition and learning, 11*(2), 215-235.
Chyung, S. Y. (1996). *The effects of conceptual tempo and computer-assisted instructional environment on college students' self-regulated learning and academic achievement.* Texas Tech University.
Cicchinelli, A., Veas, E., Pardo, A., Pammer-Schindler, V., Fessl, A., Barreiros, C., & Lindstädt, S. (2018). Finding traces of self-regulated learning in activity streams.
Clarebout, G., & Elen, J. (2009). The complexity of tool use in computer-based learning environments. *Instructional Science, 37*(5), 475. doi: 10.1007/s11251-008-9068-3
Cleary, T. J., Callan, G. L., & Zimmerman, B. J. (2012). Assessing self-regulation as a cyclical, context-specific phenomenon: Overview and analysis of SRL microanalytic protocols. *Education Research International, 2012*. doi: 10.1155/2012/428639
Cohen, J. (1988). Statistical power analysis for the behavioral sciences . Hilsdale. *NJ: Lawrence Earlbaum Associates, 2*.
Collazo, N. A. J., Elen, J., & Clarebout, G. (2015). The multiple effects of combined tools in computer-based learning environments. *Computers in Human Behavior, 51*, 82-95. doi: 10.1016/j.chb.2015.04.050
Crippen, K. J., Biesinger, K. D., Muis, K. R., & Orgill, M. (2009). The role of goal orientation and self-efficacy in learning from web-based worked examples. *Journal of Interactive Learning Research, 20*(4), 385.
Daniela, P. (2015). The relationship between self-regulation, motivation and performance at secondary school students. *Procedia-Social and Behavioral Sciences, 191*, 2549-2553. doi: 10.1016/j.sbspro.2015.04.410

Delgadillo, J., Overend, K., Lucock, M., Groom, M., Kirby, N., McMillan, D., . . . de Jong, K. (2017). Improving the efficiency of psychological treatment using outcome feedback technology. *Behaviour research and therapy, 99*, 89-97.
Deschacht, N., & Goeman, K. (2015). The effect of blended learning on course persistence and performance of adult learners: A difference-in-differences analysis. *Computers & Education, 87*, 83-89. doi: 10.1016/j.compedu.2015.03.020
DiFrancesca, D., Nietfeld, J. L., & Cao, L. (2016). A comparison of high and low achieving students on self-regulated learning variables. *Learning and individual differences, 45*, 228-236.
Dimitrov, D. M., & Rumrill, P. D. (2003). Pretest-posttest designs and measurement of change. *Work, 20*(2), 159-165.
Dinsmore, D. L., & Fryer, L. K. (2018). The intersection between depth and the regulation of strategy use. *British Journal of Educational Psychology, 88*(1), 1-8.
Dompnier, B., Darnon, C., & Butera, F. (2013). When performance-approach goals predict academic achievement and when they do not: A social value approach. *British Journal of Social Psychology, 52*(3), 587-596.
Drysdale, J. S., Graham, C. R., Spring, K. J., & Halverson, L. R. (2013). An analysis of research trends in dissertations and theses studying blended learning. *The internet and higher education, 17*, 90-100. doi: 10.1016/j.iheduc.2012.11.003
Duffy, M. C., & Azevedo, R. (2015). Motivation matters: Interactions between achievement goals and agent scaffolding for self-regulated learning within an intelligent tutoring system. *Computers in Human Behavior, 52*, 338-348.
Dunlosky, J., & Rawson, K. A. (2012). Overconfidence produces underachievement: Inaccurate self evaluations undermine students' learning and retention. *Learning and instruction, 22*(4), 271-280.
Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: Why understanding students' judgments can improve their achievement. *Learning and instruction, 24*, 58-61.
Earley, P. C., Northcraft, G. B., Lee, C., & Lituchy, T. R. (1990). Impact of process and outcome feedback on the relation of goal setting to task performance. *Academy of Management Journal, 33*(1), 87-105.
Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual review of psychology, 53*(1), 109-132. doi: 10.1146/annurev.psych.53.100901.135153
Elliot, A. J., & Church, M. A. (1997). A hierarchical model of approach and avoidance achievement motivation. *Journal of personality and social psychology, 72*(1), 218. doi: 10.1037//0022-3514.72.1.218
Elliot, A. J., & Dweck, C. S. (2013). *Handbook of competence and motivation*: Guilford Publications.
Elliot, A. J., & McGregor, H. A. (2001). A 2×2 achievement goal framework. *Journal of personality and social psychology, 80*(3), 501.
Ellis, S. (2012). Learning from errors: The role of after-event reviews *Human fallibility* (pp. 215-232): Springer.
Ellis, S., Carette, B., Anseel, F., & Lievens, F. (2014). Systematic reflection: Implications for learning from failures and successes. *Current Directions in Psychological Science, 23*(1), 67-72.
Endedijk, M. D., Brekelmans, M., Verloop, N., Sleegers, P. J., & Vermunt, J. D. (2014). Individual differences in student teachers' self-regulated learning: An examination of regulation configurations in relation to conceptions of learning to teach. *Learning and individual differences, 30*, 155-162. doi: 10.1016/j.lindif.2013.12.005
Entwistle, N. J., & Peterson, E. R. (2004). Conceptions of learning and knowledge in higher education: Relationships with study behaviour and influences of learning environments. *International journal of educational research, 41*(6), 407-428.
Ernst, B., & Steinhauser, M. (2017). Top-down control over feedback processing: The probability of valid feedback affects feedback-related brain activity. *Brain and cognition, 115*, 33-40.

Eva, K. W., Armson, H., Holmboe, E., Lockyer, J., Loney, E., Mann, K., & Sargeant, J. (2012). Factors influencing responsiveness to feedback: on the interplay between fear, confidence, and reasoning processes. *Advances in health sciences education, 17*(1), 15-26.

Evans, P. (1996). Government action, social capital and development: reviewing the evidence on synergy. *World development, 24*(6), 1119-1132.

Field, A. (2013). *Discovering statistics using IBM SPSS statistics*: Sage.

Fonseca, D., Martí, N., Redondo, E., Navarro, I., & Sánchez, A. (2014). Relationship between student profile, tool use, participation, and academic performance with the use of Augmented Reality technology for visualized architecture models. *Computers in Human Behavior, 31*, 434-445.

Frysak, J. (2017). *Feedback Mechanisms for Decision Support Systems: A Literature Review.* Paper presented at the World Conference on Information Systems and Technologies.

Gabadinho, A., Ritschard, G., Mueller, N. S., & Studer, M. (2011). Analyzing and visualizing state sequences in R with TraMineR. *Journal of Statistical Software, 40*(4), 1-37.

Garza, M. M. M. (2016). *Coevolution of Theory and Data Analytics of Digital Game-Based Learning.* Vanderbilt University.

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

Geitz, G., Joosten-ten Brinke, D., & Kirschner, P. A. (2016). Changing learning behaviour: Self-efficacy and goal orientation in PBL groups in higher education. *International journal of educational research, 75*, 146-158.

Gielen, S., Peeters, E., Dochy, F., Onghena, P., & Struyven, K. (2010). Improving the effectiveness of peer feedback for learning. *Learning and instruction, 20*(4), 304-315.

Graham, C. R., Henrie, C. R., & Gibbons, A. S. (2014). Developing models and theory for blended learning research. *Blended learning: Research perspectives, 2*, 13-33.

Greene, J. A., & Azevedo, R. (2007). A Theoretical Review of Winne and Hadwin's Model of Self-Regulated Learning: New Perspectives and Directions. *Review of Educational Research, 77*(3), 334-372. doi: 10.3102/003465430303953

Grover, S., Basu, S., Bienkowski, M., Eagle, M., Diana, N., & Stamper, J. (2017). A Framework for Using Hypothesis-Driven Approaches to Support Data-Driven Learning Analytics in Measuring Computational Thinking in Block-Based Programming Environments. *ACM Transactions on Computing Education (TOCE), 17*(3), 14. doi: 10.1145/3105910

Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: The effects of incentives, reflection, and explanatory style. *Metacognition and learning, 3*(2), 101-121.

Halverson, L. R., Graham, C. R., Spring, K. J., Drysdale, J. S., & Henrie, C. R. (2014). A thematic analysis of the most highly cited scholarship in the first decade of blended learning research. *The internet and higher education, 20*, 20-34.

Harackiewicz, J. M., Barron, K. E., Pintrich, P. R., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of educational psychology, 94*(3), 638-645. doi: 10.1037//0022-0663.94.3.638

Harrison, C. J., Könings, K. D., Schuwirth, L., Wass, V., & van der Vleuten, C. (2015). Barriers to the uptake and use of feedback in the context of summative assessment. *Advances in health sciences education, 20*(1), 229-245.

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

Hellrung, K., & Hartig, J. (2013). Understanding and using feedback–A review of empirical studies concerning feedback from external evaluations to teachers. *Educational Research Review, 9*, 174-190.

Jovanović, J., Gašević, D., Dawson, S., Pardo, A., & Mirriahi, N. (2017). Learning analytics to unveil learning strategies in a flipped classroom. *The internet and higher education, 33*, 74-85. doi: 10.1016/j.iheduc.2017.02.001

Kassab, S. E., Al-Shafei, A. I., Salem, A. H., & Otoom, S. (2015). Relationships between the quality of blended learning experience, self-regulated learning, and academic achievement of medical students: a path analysis. *Advances in medical education and practice, 6*, 27.

Kim, Y. J., & Shute, V. J. (2015). The interplay of game elements with psychometric qualities, learning, and enjoyment in game-based assessment. *Computers & Education, 87*, 340-356. doi: 10.1016/j.compedu.2015.07.009

Kitsantas, A., Steen, S., & Huie, F. (2017). The role of self-regulated strategies and goal orientation in predicting achievement of elementary school children. *International Electronic Journal of Elementary Education, 2*(1), 65-81.

Kizilcec, R. F., Pérez-Sanagustín, M., & Maldonado, J. J. (2017). Self-regulated learning strategies predict learner behavior and goal attainment in Massive Open Online Courses. *Computers & Education, 104*, 18-33. doi: 10.1016/j.compedu.2016.10.001

Klassen, R. (2002). A question of calibration: A review of the self-efficacy beliefs of students with learning disabilities. *Learning disability quarterly, 25*(2), 88-102.

Kluger, A. N., & DeNisi, A. (1996). The effects of feedback interventions on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological bulletin, 119*(2), 254.

Kuo, Y.-C., Walker, A. E., Schroder, K. E., & Belland, B. R. (2014). Interaction, Internet self-efficacy, and self-regulated learning as predictors of student satisfaction in online education courses. *The internet and higher education, 20*, 35-50. doi: 10.1016/j.iheduc.2013.10.001

Lesgold, A., Rubinson, H., Feltovich, P., Glaser, R., Klopfer, D., & Wang, Y. (1988). Expertise in a complex skill: Diagnosing x-ray pictures.

Levine, T., & Donitsa-Schmidt, S. (1998). Computer use, confidence, attitudes, and knowledge: A causal analysis. *Computers in Human Behavior, 14*(1), 125-146.

Lin, B., Coburn, S. S., & Eisenberg, N. (2016). Self-regulation and reading achievement. *The cognitive development of reading and reading comprehension*, 67-86.

Liu, W., Wang, C., & Parkins, E. (2005). A longitudinal study of students' academic self-concept in a streamed setting: The Singapore context. *British Journal of Educational Psychology, 75*(4), 567-586. doi: 10.1348/000709905X42239

Liu, W. C., & Wang, C. K. J. (2005). Academic self-concept: A cross-sectional study of grade and gender differences in a Singapore secondary school. *Asia Pacific Education Review, 6*(1), 20-27. doi: 10.1007/bf03024964

Liu, Z., Dev, H., Dontcheva, M., & Hoffman, M. (2016). Mining, Pruning and Visualizing Frequent Patterns for Temporal Event Sequence Analysis.

Lust, G. (2012). Opening the Black Box. Students' Tool-use within a Technology-Enhanced learning environment: An Ecological-Valid Approach.

Lust, G., Vandewaetere, M., Ceulemans, E., Elen, J., & Clarebout, G. (2011). Tool-use in a blended undergraduate course: In Search of user profiles. *Computers & Education, 57*(3), 2135-2144. doi: 10.1016/j.compedu.2011.05.010

Maki, R. H., Shields, M., Wheeler, A. E., & Zacchilli, T. L. (2005). Individual Differences in Absolute and Relative Metacomprehension Accuracy. *Journal of educational psychology, 97*(4), 723.

Mayer, R. E. (2004). Should there be a three-strikes rule against pure discovery learning? *American psychologist, 59*(1), 14.

McCutcheon, K., Lohan, M., Traynor, M., & Martin, D. (2015). A systematic review evaluating the impact of online or blended learning vs. face-to-face learning of clinical skills in undergraduate nurse education. *Journal of advanced nursing, 71*(2), 255-270. doi: 10.1111/jan.12509

Mega, C., Ronconi, L., & De Beni, R. (2014). What makes a good student? How emotions, self-regulated learning, and motivation contribute to academic achievement. *Journal of educational psychology, 106*(1), 121.

Midgley, C. (2014). *Goals, goal structures, and patterns of adaptive learning*: Routledge.

Midgley, C., Kaplan, A., & Middleton, M. (2001). Performance-approach goals: Good for what, for whom, under what circumstances, and at what cost? *Journal of educational psychology, 93*(1), 77.

Mory, E. H. (2004). Feedback research revisited. *Handbook of research on educational communications and technology, 2*, 745-783.

Muis, K. R., Winne, P. H., & Ranellucci, J. (2016). The Role of Calibration Bias and Performance Feedback in Achievement Goal Regulation. *International Education Research, 4*, 14-36.

Müller, N. S., Studer, M., Gabadinho, A., & Ritschard, G. (2010). *Analyse de séquences d'événements avec TraMineR.* Paper presented at the EGC.

Nadolski, R. J., & Hummel, H. G. (2017). Retrospective cognitive feedback for progress monitoring in serious games. *British Journal of Educational Technology, 48*(6), 1368-1379.

Narciss, S. (2017). Conditions and effects of feedback viewed through the lens of the Interactive Tutoring Feedback Model *Scaling up assessment for learning in higher education* (pp. 173-189): Springer.

Nelson, T., Narens, L., & Bower, G. (1990). The psychology of learning and motivation. *Metamemory: A theoretical framework and new findings*.

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: A model and seven principles of good feedback practice. *Studies in Higher Education, 31*(2), 199-218.

Oliver, M., & Trigwell, K. (2005). Can 'blended learning'be redeemed. *E-learning, 2*(1), 17–26. doi: 10.2304/elea.2005.2.1.17

Pajares, F., & Schunk, D. H. (2001). Self-beliefs and school success: Self-efficacy, self-concept, and school achievement. *Perception, 11*, 239-266.

Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: when measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research, 60*(6), 723-735. doi: 10.1080/00313831.2015.1066436

Paulson Gjerde, K., Padgett, M. Y., & Skinner, D. (2017). The Impact of Process vs. Outcome Feedback on Student Performance and Perceptions. *Journal of Learning in Higher Education, 13*(1), 73-82.

Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory into Practice, 41*(4), 219-+. doi: 10.1207/s15430421tip4104_3

Popelka, E. (2015). Improving the accuracy of middle school students' self-assessment, peer assessment, and mathematics achievement.

Puustinen, M., & Pulkkinen, L. (2001). Models of Self-regulated Learning: A review. *Scandinavian Journal of Educational Research, 45*(3), 269-286. doi: 10.1080/00313830120074206

Rashid, T., & Asghar, H. M. (2016). Technology use, self-directed learning, student engagement and academic performance: Examining the interrelations. *Computers in Human Behavior, 63*, 604-612. doi: 10.1016/j.chb.2016.05.084

Ridder, J., McGaghie, W. C., Stokking, K. M., & Cate, O. T. (2015). Variables that affect the process and outcome of feedback, relevant for medical training: a meta-review. *Medical education, 49*(7), 658-673.

Rienties, B., & Toetenel, L. (2016). The impact of learning design on student behaviour, satisfaction and performance: A cross-institutional comparison across 151 modules. *Computers in Human Behavior, 60*, 333-341.

Rienties, B., Toetenel, L., & Bryan, A. (2015). *Scaling up learning design: impact of learning design activities on LMS behavior and performance.* Paper presented at the Proceedings of the Fifth International Conference on Learning Analytics And Knowledge.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition and learning, 4*(1), 33-45.

Sedrakyan, G., & Snoeck, M. (2016). *Cognitive feedback and behavioral feedforward automation perspectives for modeling and validation in a learning context.* Paper presented at the International Conference on Model-Driven Engineering and Software Development.

Segedy, J. R. (2014). *Adaptive scaffolds in open-ended computer-based learning environments*: Vanderbilt University.
Senko, C., Durik, A. M., Patel, L., Lovejoy, C. M., & Valentiner, D. (2013). Performance-approach goal effects on achievement under low versus high challenge conditions. *Learning and instruction, 23*, 60-68. doi: 10.1016/j.learninstruc.2012.05.006
Siadaty, M., Gašević, D., & Hatala, M. (2016). Trace-based micro-analytic measurement of self-regulated learning processes. *Journal of Learning Analytics, 3*(1).
Song, H. S., Kalet, A. L., & Plass, J. L. (2016). Interplay of prior knowledge, self-regulation and motivation in complex multimedia learning environments. *Journal of Computer Assisted Learning, 32*(1), 31-50.
Sonnenberg, C., & Bannert, M. (2015). Discovering the effects of metacognitive prompts on the sequential structure of SRL-processes using process mining techniques. *Journal of Learning Analytics, 2*(1), 72-100.
Spring, K. J., & Graham, C. R. (2017). Thematic Patterns in International Blended Learning Literature, Research, Practices, and Terminology. *Online Learning, 21*(4).
Stolp, S., & Zabrucky, K. M. (2017). Contributions of metacognitive and self-regulated learning theories to investigations of calibration of comprehension. *International Electronic Journal of Elementary Education, 2*(1), 7-31.
Stone, N. J. (2000). Exploring the relationship between calibration and self-regulated learning. *Educational psychology review, 12*(4), 437-475.
Studer, M., Mueller, N. S., Ritschard, G., & Gabadinho, A. (2010). *Classer, discriminer et visualiser des séquences d'événements.* Paper presented at the EGC.
Suthers, D., & Verbert, K. (2013). *Learning analytics as a middle space.* Paper presented at the Proceedings of the Third International Conference on Learning Analytics and Knowledge.
Szabo, C., Falkner, N., Knutas, A., & Dorodchi, M. (2018). *Understanding the Effects of Lecturer Intervention on Computer Science Student Behaviour.* Paper presented at the Proceedings of the 2017 ITiCSE Conference on Working Group Reports.
Tempelaar, D. T., Rienties, B., & Giesbers, B. (2015). In search for the most informative data for feedback generation: Learning Analytics in a data-rich context. *Computers in Human Behavior, 47*, 157-167.
Thiede, K. W., Anderson, M. C. M., & Therriault, D. (2003). Accuracy of metacognitive monitoring affects learning of texts. *Journal of educational psychology, 95*(1), 66-73. doi: 10.1037/0022-0663.95.1.66
Timmers, C. F., Walraven, A., & Veldkamp, B. P. (2015). The effect of regulation feedback in a computer-based formative assessment on information problem solving. *Computers & Education, 87*, 1-9.
Tricot, A., & Sweller, J. (2014). Domain-specific knowledge and why teaching generic skills does not work. *Educational psychology review, 26*(2), 265-283.
Van der Kleij, F. M., Feskens, R. C., & Eggen, T. J. (2015). Effects of feedback in a computer-based learning environment on students' learning outcomes: A meta-analysis. *Review of Educational Research, 85*(4), 475-511.
Van Laer, S., & Elen, J. (2016). Adults' Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. *Technology, Knowledge and Learning*, 1-31.
Van Popta, E., Kral, M., Camp, G., Martens, R. L., & Simons, P. R.-J. (2017). Exploring the value of peer feedback in online learning for the provider. *Educational Research Review, 20*, 24-34.
Veenman, M. V. (1993). *Intellectual ability and metacognitive skill: Determinants of discovery learning in computerized learning environments*: Faculteit Psychologie, Universiteit van Amsterdam.
Veenman, M. V., & Alexander, P. (2011). Learning to self-monitor and self-regulate. *Handbook of research on learning and instruction*, 197-218.
Veenman, M. V., Bavelaar, L., De Wolf, L., & Van Haaren, M. G. (2014). The on-line assessment of metacognitive skills in a computerized learning environment. *Learning and individual differences, 29*, 123-130. doi: 10.1016/j.lindif.2013.01.003

Veenman, M. V., Elshout, J. J., & Meijer, J. (1997). The generality vs domain-specificity of metacognitive skills in novice learning across domains. *Learning and instruction, 7*(2), 187-209. doi: 10.1016/S0959-4752(96)00025-4

Veenman, M. V., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and instruction, 14*(1), 89-109.

Vohs, K. D., & Baumeister, R. F. (2016). *Handbook of self-regulation: Research, theory, and applications*: Guilford Publications.

Wiliam, D., & Thompson, M. (2007). Integrating assessment with learning: What will it take to make it work? : Lawrence Erlbaum Associates.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. *Metacognition in educational theory and practice, 93*, 27–30.

Winne, P. H., & Jamieson-Noel, D. (2002). Exploring students' calibration of self reports about study tactics and achievement. *Contemporary Educational Psychology, 27*(4), 551-572.

Wolff, A., Zdrahal, Z., Nikolov, A., & Pantucek, M. (2013). *Improving retention: predicting at-risk students by analysing clicking behaviour in a virtual learning environment.* Paper presented at the Proceedings of the third international conference on learning analytics and knowledge.

Wolters, C. A., Pintrich, P. R., & Karabenick, S. A. (2005). Assessing academic self-regulated learning. *What Do Children Need to Flourish?*, 251-270.

Wolters, C. A., Won, S., & Hussain, M. (2017). Examining the relations of time management and procrastination within a model of self-regulated learning. *Metacognition and learning, 12*(3), 381-399.

Yang, C., Potts, R., & Shanks, D. R. (2017). The forward testing effect on self-regulated study time allocation and metamemory monitoring. *Journal of Experimental Psychology: Applied, 23*(3), 263.

Zhou, M. (2016). Data pre-processing of student e-learning logs *Information Science and Applications (ICISA) 2016* (pp. 1007-1012): Springer.

Zimmerman, B. J. (2013). From cognitive modeling to self-regulation: A social cognitive career path. *Educational psychologist, 48*(3), 135-147. doi: 10.1080/00461520.2013.794676

Zimmerman, B. J., & Schunk, D. H. (2001). *Self-regulated learning and academic achievement: Theoretical perspectives*: Routledge.

Zimmerman, B. J., Schunk, D. H., & DiBenedetto, M. K. (2015). A personal agency view of self-regulated learning. *Self-Concept, Motivation and Identity: Underpinning Success with Research and Practice*, 83.

How can teachers support student interaction during computer-supported collaborative learning? An exploratory case study in a higher education setting for hands-on learners

Abstract

This study focused on how teacher support might influence students' interactions during collaboration. As little research on computer-supported collaborative learning (CSCL) is conducted in contexts that include students with a background in vocational or technical secondary education, an exploratory case study was carried out with 1 teacher and 10 students with prior craft knowledge aiming to become vocational teachers. The purpose was to make visible, with empirical examples, how teacher support might contribute to more productive student interaction during CSCL. General understanding on the quality of student interaction during collaboration was obtained through qualitative content analysis, while a detailed interpretative analysis of the dialogues between the teacher and the groups made it possible to trace different interaction patterns. The results suggest that it was hard for both students and the teacher to empower productive interactions. Two interaction patterns showed to be effective: the pattern in which the teacher takes the role of the devil's advocate, and the pattern in which the teacher provides gradual assistance. The latter finding suggests that not only the type of questions, but also the order of the questions asked by the teacher, is important.

Introduction

Since the 1990s, computer-supported collaborative learning (CSCL) is increasingly receiving attention in educational research, often with the aim to foster productive interactions between group members, which in turn should produce desired learning outcomes (Dillenbourg, Järvelä, & Fischer, 2009; Lazonder, 2005). However, previous research has shown that effective collaboration, in which students go beyond the comparison of information and focus on jointly building new knowledge, often does not occur spontaneously (Dillenbourg, 2002; Mayordomo & Onrubia, 2015; Onrubia & Engel, 2012). Research therefore suggests that providing additional support is necessary for optimizing student interaction during collaboration (Dillenbourg, 2002). For example, one approach to support students' interaction is providing just-in-time teacher support. This means that the teacher provides calibrated support to the groups during the collaboration process (Kaendler, Wiedmann, Rummel, & Spada, 2014; Onrubia & Engel, 2012; Puntambekar & Hübscher, 2005).

Although a large amount of research has been focusing on supportive measures to enhance the effectiveness of collaborative learning (Dillenbourg, 2002), research in contexts that include students with a background in vocational or technical secondary education is under-represented within the field of CSCL (Schwendimann, De Wever, Hämäläinen, & Cattaneo, 2017). These students, referred to here as hands-on learners, typically prefer to learn through hands-on experiences and demonstrations, and are less familiar with problem-solving and CSCL (Hämäläinen, De Wever, Malin, & Cincinnato, 2015; Hämäläinen, De Wever, Nissinen, & Cincinnato, 2017). Moreover, previous research in this context has shown that these students often lack strategies to get the most out of the collaboration, or to decide together upon how to tackle a collaborative task (Boelens & De Wever, 2017). This indicates a need to identify beneficial ways to support hands-on learners in CSCL. In the present study, we therefore examine the influence of just-in-time teacher support on hands-on learners' interaction during collaboration. The aim is to illustrate, with empirical examples, how teacher support can foster hands-on learners' interaction during CSCL, and consequently lead to more effective collaboration. The remainder of this introduction proceeds as follows: first a brief overview of the literature is given that points out what effective collaboration or productive interaction entails, and afterwards different ways to promote productive interaction during CSCL through teacher support are discussed.

Promoting productive interaction during collaboration

Previous work has stated that productive interaction is not merely sharing ideas or knowledge with each other, but also treating each other's opinions critically and negotiating about the content (De Wever, Van Keer, Schellens, & Valcke, 2009; Gunawardena, Lowe, & Anderson, 1997; Onrubia & Engel, 2012). Gunawardena et al. (1997) argue that this process of *collaborative knowledge construction* evolves through five levels. The first two levels represent the lower mental functions: sharing and comparing information (level 1), and identifying areas of disagreement (level 2). The latter three levels represent higher mental functions, including: negotiating meaning and co-construction of knowledge (level 3), testing and modifying the proposed synthesis that resulted from co-construction (level 4), and applying the newly co-constructed knowledge (level 5) (De Wever et al., 2009; Gunawardena et al., 1997).

However, numerous studies have pointed out the difficulty of reaching the more advanced levels of collaborative knowledge construction (e.g. Dillenbourg, 2002; Mayordomo & Onrubia, 2015; Onrubia & Engel, 2012). As a result, research has called out to involve the teacher more closely in the collaboration process to support productive student interaction on a just-in-time basis (Dillenbourg et al., 2009; Hämäläinen & De Wever, 2013; Hmelo-Silver, Duncan, & Chinn, 2007; Mayordomo & Onrubia, 2015). In particular, the teacher has a central role to monitor and diagnose student interaction during collaboration, and to provide groups with adaptive support (Kaendler et al., 2014; Puntambekar & Hübscher, 2005). This teacher support may have the purpose of both facilitating the learning or interaction process, and providing content knowledge based on students' needs (Hmelo-Silver et al., 2007). For instance, the teacher can provide just-in-time support by prompting, advising, asking thought-provoking questions, or giving additional explanations to students during collaboration (Kaendler et al., 2014). As such, the teacher becomes an active participant in the group's discussion.

Previous research has identified different types of support to facilitate interaction during collaboration. For instance, Fahy et al. (2000) distinguishes five kinds of contributions: vertical questioning, horizontal questioning, statements, reflections, and scaffolding. As shown in Table 1, the intention of the teacher, or the expected response from (one of) the group (members), is different for the five categories. Other research has observed two types of teacher support: a *knowledge-providing approach*, in which teachers actively introduced new information, and a *joint problem-solving approach*, in which teachers asked specifying questions and continued shared problem solving (Hämäläinen & De Wever, 2013). These two types of support had a different effect on student interaction (Hämäläinen & De Wever, 2013). In the former approach, students' reaction was to provide knowledge, while in the latter approach, students' reaction was more directed to joint problem-solving (Hämäläinen & De Wever, 2013). In short, although some research has been carried out on teacher support to enhance students' interaction during collaboration, there is still very little understanding of how this kind of support is organized in detail. There is a need for specific guidelines regarding teacher support to promote hands-on learners' interaction during CSCL tasks.

Table 1. *Types of support to facilitate interaction as described by Fahy et al. (2000).*

Type of support	**Intention**
Vertical questioning	Ask questions to acquire information from others
Horizontal questioning	Ask questions to initiate a dialogue or discussion
Statements	Provide information or corrections to others
Reflections	Provide insight in internal conflicts, reasoning processes, doubts, beliefs
Scaffolding	Invite others to comment, by for instance naming others

As of yet, however, the studies that are focusing on CSCL in vocational education and training settings (e.g. Hämäläinen & De Wever, 2013), often take place in a setting that addresses concrete professional tasks, while the learning tasks in the present study are more in line with learning in academic, higher education settings. To sum up, the students in this study are similar to those in vocational education and training research, while the instructional context in this study is closer related to the higher education context. Situated against this background, the present study deals with the effect of teacher support on student interaction during CSCL. Two research questions are formulated, of which the first one aims to provide an overall view of what happened during the group work, and the second aims to provide detailed insight in the kind of teacher support:

1. Which levels of collaborative knowledge construction are attained by the groups, and what is the influence of teacher support on these levels of collaborative knowledge construction?
2. What interaction patterns emerge from the interplay between teacher support and how students respond to that support during collaboration?

Method

Research setting and participants

The present study is part of a research project on the design of blended learning environments for hands-on learners. This study is situated in a teacher training program, in which hands-on learners or students with a degree of vocational or technical secondary education (e.g. bakery, electricity, hairdressing) aim to become teachers in vocational programs (see also Boelens, Voet, & De Wever, 2018). Participants in this study were ten students (two male and eight female) and their teacher (female). The students were divided into two dyads and two triads, and decided themselves about the composition of the groups. The average age of the students was 27 years (*SD*=6.17, range=21-39). The students indicated that they had little to no experience with CSCL. The age of the teacher was 43 years and she was involved in the design and construction of the learning task. The teacher had limited experience in fostering productive interaction during CSCL.

Design of the CSCL task

To facilitate productive student interaction, the group work was organized by the star group work coordination approach (Onrubia & Engel, 2012). This means that, prior to starting working together as a group, students first had to produce an individual contribution (i.e. gather and summarize information) to the group task (Onrubia & Engel, 2012). Also, students received clear instructions and examples of how they could engage in productive discussions with each other in a constructive way. This was done by encouraging them to take on several roles (e.g. moderator and summarizer) in the discussion (see e.g. De Wever, Van Keer, Schellens, & Valcke, 2010) and by providing them specific examples on how they could formulate and support their opinions. An example of the latter, was this specific guide: (a) formulate your opinion (e.g. “I prefer the second definition”), (b) elaborate on your opinion (e.g. “this definition is the most complete because the three criteria (i.e. retardation, didactic resistance, no clear cause) for the diagnosis of dyslexia are present”), and (c) provide an example (e.g. “‘despite good approach’ refers to didactic resistance”).

The purpose of the CSCL task was that the four groups all studied another learning or developmental disability, and afterwards presented their work to each other. Students were required to gather and process information about this specific disability, and the group had to formulate several supportive measures to cope with a learning or development disability in their classroom practice. A *process worksheet* that structured the CSCL task by introducing 7 main steps with in total 20 underlying subtasks, guided the groups towards the end product (van Merriënboer & Kirschner, 2013) (see Appendix A for an example). After an introductory face-to-face meeting, the groups had three weeks to complete the first five subtasks of their CSCL task using the online environment. After these three weeks, the groups had a face-to-face meeting to complete the remaining 15 subtasks. Students collaborated in a shared Google document, and were instructed to use the Google forum to interact with each other during the online part of the task.

The teacher observed the group discussions both in the online and the face-to-face environment. On a just-in-time basis, the teacher diagnosed and supported student interaction with the purpose of fostering students to go beyond sharing and comparing of information, and thus guiding the interaction toward a higher level of collaborative knowledge construction. In particular, the teacher asked questions to stimulate students to clarify, elaborate, and reflect on their ideas. The teacher could make use of specific guidelines and prompts to help her to diagnose and support the group discussions.

Data collection and analysis

Direct measures were used to investigate students' levels of collaborative knowledge construction and the support provided by the teacher, by logging the students' and teacher's postings in the forum (69 separate messages), and recording the face-to-face interaction during collaboration with audio recorders. These recordings lasted on average 1:59:37 hours (*SD*=0:12:39, range=1:45:43-2:12:37) and were transcribed. In this way, we were able to capture and analyze all interaction (both online and face-to-face) related to the group work.

Levels of collaborative knowledge construction

Objective units that are defined by the original author or speaker (De Wever et al., 2009; Gunawardena et al., 1997; Henri, 1991) were chosen as units of analysis. To analyze students' online contributions, messages were selected as unit of analysis (n=67), while for analyzing students' face-to-face contributions, turn takes were selected as unit of analysis (n=3293). The coding and analysis procedure consisted of three consecutive stages. The first stage served to identify the contributions in which students are interacting about the content of the task, i.e. content-related interaction. Content-related interaction included units in which the content of the task was discussed, such as sharing content-related information, questions, comments, requests, information sources, or discussion of that content (Onrubia & Engel, 2012; Strijbos, Martens, Prins, & Jochems, 2006). Contributions that were related to the coordination of the group, social issues, technical issues, or other parts of the course were excluded from analysis. Two coders analyzed 464 (14%) units together, and all other units (n=2896) were coded independently. To determine the level of correspondence among these coders, Krippendorff's alpha was calculated ($\alpha = 0.69$).

In the second stage, the units labeled as content-related interaction (n=1541, 46% of all interaction) were analyzed to determine the levels of collaborative knowledge construction. For this, the interaction analysis model of Gunawardena et al. (1997) and the corresponding coding scheme defined by Gunawardena et al. (1997) and refined by De Wever et al. (2009) was used to conduct a qualitative content analysis. Two independent coders carried out the coding activity. After working with coding examples for each level of knowledge construction, 254 (20%) units were coded together to discuss and elaborate on the coding process. All other units (n=1287) were coded independently. The reliability of coding the level of collaborative knowledge construction for each unit was checked by Krippendorff's alpha (α=0.65).

In the third stage, a way was sought to make the data more accessible and meaningful to produce results. In this respect, the data was aggregated at subtask level (n=20 subtasks) to extract for every subtask the highest level of collaborative knowledge construction. As such, we identified the subtasks in which students reached the more advanced levels of collaborative knowledge construction. The reliability of coding the level of collaborative knowledge construction for each subtask was checked by Krippendorff's alpha (α=0.69). All disparities were discussed by the two independent coders until agreement was reached on all codes.

Teacher support

To analyze the teacher's online contributions, messages were selected as unit of analysis (n=2), while for analyzing the teacher's face-to-face contributions, turn takes were selected as unit of analysis (n=205). The coding and analysis procedure consisted of three consecutive stages. The first stage served to identify the contributions in which the teacher support had the purpose of facilitating the interaction process or providing content knowledge. Contributions that were related to the coordination of the group (e.g. statements with reference to time, signaling that a group is not performing according to the task division), social issues, technical issues, or the course in general were excluded from analysis. After coding each unit, the coders indicated for each subtask whether or not the teacher provided support that was aimed to facilitate the interaction process or to provide content knowledge. This was the case for 11 out of the 20 subtasks. Both coders independently analyzed the data and agreed with each other about all subtasks. In this way, the analysis of the teacher support could be aligned with the levels of knowledge construction attained by the students.

In the second stage, the dialogue between the teacher and the students in these 11 subtasks was further analyzed. For each subtask, it was investigated whether the teacher support had an influence on the levels of collaborative knowledge construction. For example, when the group's discussion was situated in level 2 or higher, both coders independently analyzed whether this was due to the support of the teacher. This was done based on how the students responded (one or more contributions) to the support of the teacher (one or more contributions). The two coders analyzed three subtasks together, and then coded the other eight subtasks independently. Since both coders agreed with each other about all subtasks, reliability was ensured and no further discussion was needed.

The third coding stage also focused on the 11 subtasks in which the teacher provided support, and aimed to identify the different interaction patterns that emerged from the interplay between teacher support and how students responded to that support. Previous coding schemes have focused on student interaction (e.g. De Wever et al., 2009) or teacher support (Onrubia & Engel, 2012) independently from each other. However, to analyze the interplay between student interaction and teacher support in the present study, a coding scheme that concentrates specifically on patterns that occur during these dialogues was needed. As such, we developed a coding scheme that was informed by the interaction patterns defined by Fahy et al. (2000). Further elaboration of these patterns and one new pattern were deduced from the empirical material collected for this study by the first author. In this way, four patterns were distinguished based on Fahy et al. (2000): pattern of (1) gradual assistance, (2) the devil's advocate, (3) appointing group members who are less actively participating, and (4) question and answer. A fifth pattern, sharing general advice, was deduced from the data.

Results

The results section consists of two parts. In the first part, we aim to present an overall picture of students' levels of collaborative knowledge construction and the provided support by the teacher based on the data at sub-task level. In the second part, we zoom in on the interplay between students' interaction and teacher support, through a fine-grained interpretative analysis of the dialogues that took place.

Levels of collaborative knowledge construction and teacher support

To shed light on students' interaction levels within the four groups, Figure 1 plots their highest reached level of collaborative knowledge construction for each subtask. The subtasks are presented on the horizontal axis, while the highest achieved level of collaborative knowledge construction is indicated on the vertical axis. Overall, Figure 1 shows that most of the discussions were situated at the less advanced levels (level 1 and 2) of collaborative knowledge construction. However, what stands out in the results is that there is an important difference between the four groups in terms of the achieved levels of collaborative knowledge construction. Group 1 and 4 achieved more advanced levels of collaborative knowledge construction (i.e. level 3 and higher), while the discussions in group 2 and 3 were mostly or exclusively situated in the less advanced levels of collaborative knowledge construction (i.e. level 1 and 2).

The teacher provided support that aimed to facilitate the interaction process or to provide content knowledge in 11 subtasks. Regarding the influence of the teacher support, only in two subtasks the group's discussion was facilitated to a higher level of collaborative knowledge construction. Figure 1 shows for each group when the teacher provided support and whether this resulted in a higher level of collaborative knowledge construction. The semicircle arrows indicate that the teacher support had no influence on the level of collaborative knowledge construction, while the upwards arrows show to which level of collaborative knowledge construction the discussion was stimulated by the teacher.

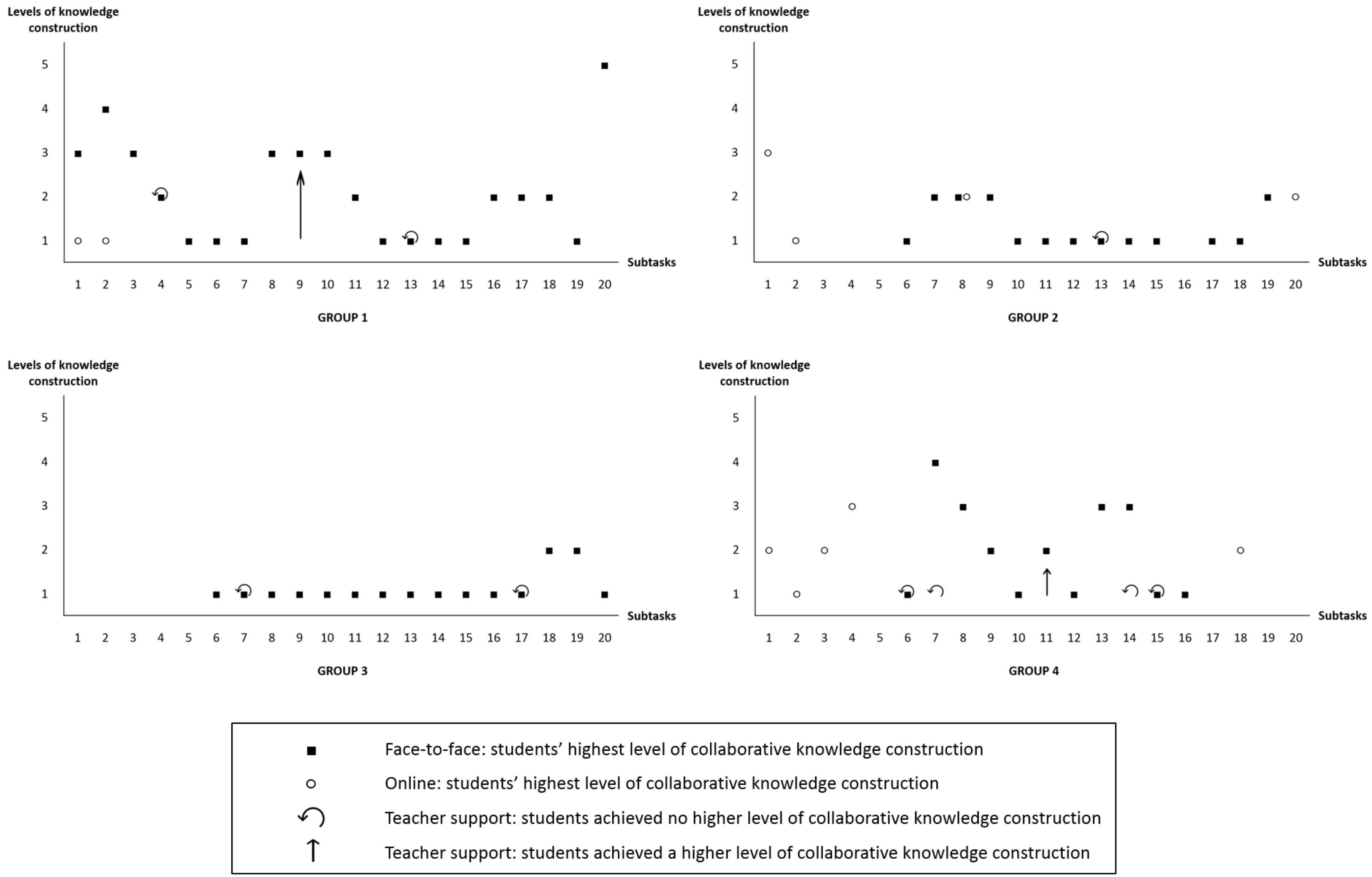

Figure 1. Levels of collaborative knowledge construction (Y-axis) attained by the groups, and influence of teacher support (indicated by arrows) over the 20 subtasks (X-axis).

Note. When no level is indicated for a subtask, there was no information available about how the students interacted with each other.

Interaction patterns

To better understand what actually happened when the teacher provided support, these 11 dialogues were studied in detail. Based on the interaction analysis, five interaction patterns could be discerned regarding the interplay between students' interaction and teacher support, of which two showed to be effective to stimulate students to reach a higher level of collaborative knowledge construction. In this section, these patterns are illustrated (excerpts of the interaction between the students and the teacher can be found in Appendix B). The names in the next paragraphs are pseudonyms that refer to the students in the study.

Pattern A: Gradual support

In this pattern, the teacher gradually stimulated students' discussion to a higher level of collaborative knowledge construction. This pattern occurred only once, in students' interaction during collaboration of group 1 (Esra, Lucas, and Victor, see Appendix C for more background information on their collaboration process). An excerpt of the interaction between the teacher and the students is provided in Appendix B. As can be seen from the excerpt, the support was initially started by the teacher, at a moment where students identified disagreement with each other and were searching for a way to reach consensus. At the request of the teacher, the group members shared and compared their information again, and identified areas of disagreement. In addition, they started to negotiate about the relative weight to be assigned to the different arguments or sources. The teacher first posed concrete and specific questions to ensure that all students displayed the information they individually retrieved (e.g. what information did you find in this book?), and stated where they retrieved that information (e.g. what book did you use?). Afterwards, she stimulated the students to decide how much weight they should assign to the type of source or argument (e.g. do you think a forum is a reliable source? What would be a reliable source?). In short, the teacher first asked open questions related to level 1 of collaborative knowledge construction, then she shifted to questions related to level 2, and finally the teacher provided support related to level 3. The teacher brought no new information to formulate a solution. In this way, the teacher managed to take the discussion to a higher level of collaborative knowledge construction.

Pattern B: The devil's advocate

The second pattern implied that the teacher expressed doubts about a certain solution to elicit students to restate their position. This pattern occurred once, in students' interaction during collaboration of group 4 (Olivia, Mia, and Anna, see Appendix C for a description of their collaboration process). As can be seen from the excerpt in Appendix B, the teacher was listening to the students' discussion, and spontaneously provided support. At that specific moment, the group was processing the task rather fast. The teacher provoked their thinking by questioning one of their earlier retrieved solutions (e.g. is it not more than 1%?). This stimulated the group to go back to their individual contributions and reformulate what they had found in group. Mia restated their position, and advanced an additional argument in its support by referencing to what they found in the literature (e.g. we found that 40 to 60 thousand people have ASD) and arguing that it is about the population in Flanders. Although the teacher tried to provoke more arguments and negotiation, the group provided no concrete arguments why this solution is the right one, and focused on the next subtask. The teacher no longer asked any further questions. In short, the teacher managed to stimulate the group to restate their position and provide arguments to support their position (level 2 of collaborative knowledge construction). To do this, the teacher expressed disagreement or doubts about a certain solution, without bringing new information or arguments to formulate a solution.

Pattern C: Sharing general advice

In each group, the teacher provided once general advice to stimulate students' interaction to a higher level of collaborative knowledge construction. To illustrate this pattern, an excerpt of the interaction of group 2 (Sophie and Emma, see Appendix C for a description of their collaboration process) is presented in Appendix B. The excerpt shows that the teacher reminded the students to display and discuss about their individual opinions. The teacher did not refer to a specific aspect of the task or did not ask a concrete question to elicit individual ideas or opinions. Consequently, this pattern showed no direct effect on the levels of collaborative knowledge construction reached by the groups.

Pattern D: Appointing group members who are less actively participating

The teacher attempted four times to involve a group member who was not actively participating in the group work at a certain point in the time. This occurred in three of the four groups (group 1, 3, and 4). The teacher attempted to involve an individual group member by, for instance, asking explicitly for his/her opinion about a certain topic, to elicit the comparison of information and identification of areas of (dis)agreement. However, the teacher often used closed questions to involve group members who were less actively participating what resulted in short and concise answers, or did not react when a student was not answering immediately. Consequently, the interventions did not lead to more productive interaction. For example, Appendix B presents an excerpt of the interaction of group 4.

Pattern E: Question and answer

In this pattern, the teacher tried to provoke new information, knowledge or ideas of the group members by asking concrete and open questions about the content. This pattern occurred only once, in students' interaction during collaboration of group 3 (Charlotte and Eva, see Appendix C for a short description of their collaboration process). In Appendix B, an excerpt of the interaction between the students and the teacher is presented. It can be seen from the data that the attempt to elicit a productive discussion failed, as the students provided mainly short answers, and agreed with each other without further negotiation or argumentation.

Discussion

The two major findings of this study can be summarized as follows. First, although two of the four groups reached more advanced levels of collaborative knowledge construction in a few occasions, discussions in all groups generally were rather limited to sharing and comparing information and ideas, and teacher support seemed to have little influence on students' interaction. Second, five different interaction patterns were identified that provide more detailed insight in the interplay between teacher support and student interaction, and two of those seemed to be more effective than the others. In what follows, we discuss these findings in detail.

First, discussions between the group members were often restricted to sharing and comparing information and ideas, and also teacher support had only a limited effect as to fostering more advanced levels of collaborative knowledge construction. The finding that it seems to be hard for students to reach more advanced levels of collaborative knowledge construction, is in line with previous research in vocational education settings (Hämäläinen & Oksanen, 2012), and even in higher education settings (De Wever et al., 2010). Not surprising, previous research in the field already called for an active role of the teacher during the collaboration process (Dillenbourg, 2002; Hämäläinen & Oksanen, 2012; Kaendler et al., 2014; Onrubia & Engel, 2012). However, contrary to the findings of Hämäläinen and Oksanen (2012), our results pointed out that it is not evident that teacher support leads to more productive interaction during collaboration. A possible explanation for this result might be the fact that one teacher had to monitor and support four groups at the same time. In particular, the teacher was not involved in the whole collaboration process of each group, as was the case in the study of Hämäläinen and Oksanen (2012). This might made it harder to provide adaptive support on a just-in-time basis. Our results once again demonstrate that monitoring and supporting students' interaction during collaboration is a complex and challenging task (Hämäläinen & Oksanen, 2012; Kaendler et al., 2014). As such, there is still an important task for professional development initiatives on teacher support to enhance students' collaboration process on a just-in-time basis. Finally, another possible explanation might be that although the teacher in our study could rely on guidelines to shape the support during collaboration, the teacher did not have the time to develop her competency to monitor and support students' interaction as this was the first task in which she made use of the guidelines. In addition, also students in this study had less or no previous experience with working jointly on a CSCL task. More practice moments in the future, for both the teacher and the students, might be needed to learn how to empower productive interaction.

Particularly interesting is the second major finding, as five different interaction patterns were identified, of which two showed to be effective. The question is, however, why some of these patterns were effective, while others were not. In what follows, we first discuss two relatively frequent occurring patterns that showed no direct effect on student interaction. In addition, we discuss the three remaining patterns, that only occurred once during the collaboration, and of which two seemed to be effective to foster more productive interaction.

Two patterns that showed to be less effective to foster productive student interaction, occurred several times during collaboration: sharing general advice and appointing group members who are less actively participating in the group work. First, the teacher provided in each group general advice, what never (immediately) resulted in more productive interaction. This result may be explained by the fact that this kind of support is too general and not situated at a specific aspect of the CSCL task. Another possible explanation might be that the teacher posed closed questions, such as “have you reached consensus?” In this specific case, the question “how did you reach consensus?” might elicit more reasoning and productive interactions. Second, the teacher intervened four times to involve certain group members in the discussion. This did not provoke the desired response, as some students did not react on her question, or others simply answered that they agree with the other group members. Subsequently, the teacher made no attempt to provoke further interaction. These results strengthen our earlier stated claim that professional development initiatives are needed for passing on good practices to teachers that show how adaptive support during collaboration (e.g. posing questions in a specific way) can provoke student reasoning and productive interaction (see e.g. *the Socratic dialogue,* as described by Collins, 1976).

Each of the three following patterns occurred only once during collaboration: gradual assistance, the devil’s advocate, and question and answer. While the first two patterns showed an effect on student interaction, the question and answer pattern did not. The three patterns have in common that the teacher asked mainly open questions related to a specific part of the CSCL task with the intention to provoke joint problem-solving. However, there are three remarkable differences between both patterns provoking a more productive discussion and the pattern that did not.

First, a difference was found in involving the group members in the discussion. In both cases where the teacher provokes more productive interaction, the teacher manages to give the floor to two or three different group members. By way of contrast, in the question and answer pattern, mostly one of the group members responds to the teacher’s questions, while the other student only provides confirmative statements. Consequently, this may explain why this pattern did not lead to more productive interaction, as this entails group members who not only share their individual ideas, but also treat each other’s opinions critically (Onrubia & Engel, 2012).

A second difference is that in both effective patterns, the teacher especially asked questions related to the subject-matter without providing new knowledge, while in the less effective pattern, she does provide information to the group. This might explain in part the effectiveness of teacher support, as previous research showed that a teacher who goes beyond the provision of information, and asks students questions that require elaborating on their opinions, might be more powerful to elicit negotiation of meaning (Vosniadou, Ioannides, Dimitrakopoulou, & Papademetriou, 2001). This is also consistent with findings of previous work (Hämäläinen & De Wever, 2013), that has shown that when the teacher asks questions and contribute to problem-solving instead of provide knowledge, students are more engaged in joint problem-solving. For instance, when the teacher takes the role of the devil's advocate, this may lead to *conflict-oriented consensus building* in which students may be pushed to provide better and more arguments for their position (Weinberger & Fischer, 2006). In contrast to *quick consensus building,* or accepting others' contributions to be able to continue (Weinberger & Fischer, 2006), the former style of reaching consensus helps students to reach more advanced levels of collaborative knowledge construction (Weinberger & Fischer, 2006; Weinberger, Stegmann, & Fischer, 2007).

A third difference between both patterns that provoked productive interaction and the pattern that did not, might be related to the group dynamics. Both groups in which the teacher managed to stimulate a more productive dialogue, were visibly motivated and had also in general more productive interactions. The other group seemed less motivated to produce qualitative discussions, and their dialogue was exclusively situated at less advanced levels of collaborative knowledge construction. A likely explanation might be that the same kind of support has a different effect for particular individuals or groups depending upon their specific competences (Cronbach & Snow, 1969), and that, for instance, those groups that already performed better benefit more from teacher support (Otto & Kistner, 2017). This might be caused by the group itself, or, another possible explanation might be that it is easier for the teacher to provide adaptive support to groups that already have a rather productive discussion. However, we have little evidence for this, and future studies on this topic are therefore recommended.

Limitations and suggestions for further research

A limitation of this study is that some of the identified patterns did not emerge repeatedly and were not observed in all groups' interaction. Two patterns, sharing general advice and appointing group members who are less actively participating, occurred more often and in (almost) all groups, which makes our point stronger that they are probably not effective. The other three patterns, i.e. gradual assistance, the devil's advocate, and question and answer, however, were observed only once. Although it was not our aim to provide proof, but rather to make visible what kind of teacher support worked well, more occurrences of these patterns could strengthen our findings. Nevertheless, this study answers the current need for exploring ways to foster hands-on learners' joint problem-solving competencies (Hämäläinen et al., 2017). As such, we provide specific advice for teachers involved in these settings in the next section. However, to develop a full picture of how the teacher can optimize students' interaction during collaboration, further studies will be needed with, for instance, a more longitudinal focus, investigating the evolution of student interaction and teacher support during consecutive CSCL tasks. Another important issue for future research is including a larger sample with more students and more teachers, to make more substantiated claims about the effectiveness of interaction patterns.

Implications

The aim of the present study was to make visible, with empirical examples, how teacher support might contribute to more productive student interaction during collaboration. General understanding on the quality of student interaction during collaboration was obtained through qualitative content analysis. In addition, a detailed interpretative analysis of the interplay between teacher support and student interaction made it possible to trace five different interaction patterns and to understand how the teacher could contribute to more productive student interaction during collaboration. The results hold two important implications to both theory and practice.

First, based on the findings of the present study, some recommendations can be made to support hands-on learners in CSCL tasks. More specifically, when teachers aim to foster more productive interaction during collaboration, the following five guidelines might be helpful: (1) provide support related to a specific part of the task, (2) give the floor to all the group members to elaborate on opinions and ideas, (3) focus support on asking questions that require reasoning and elaboration of students' own opinions, (4) provide gradual assistance, i.e. start with the less advanced levels of collaborative knowledge construction to gradually lift the discussion to higher levels of collaborative knowledge construction, and (5) take the role of the devil's advocate by questioning statements and directions suggested by the group.

Second, although the present study was explorative, the five interaction patterns that were discerned contribute to the field of study that aims to find more knowledge on how teachers can provide support to promote productive interaction during collaboration (see e.g. Hämäläinen & De Wever, 2013; Kaendler et al., 2014; Onrubia & Engel, 2012). The findings reported here point out that not only the kind of questions, but also the order of the questions, determined students' collaboration process. These results can inform new targeted interventions aimed at promoting productive interaction between students during collaboration.

Acknowledgment

This work was supported by the Agency for Science and Technology (Project Number: SBO 140029). The funding source had no involvement in the study design.

References

Boelens, R., & De Wever, B. (2017). Collaborating on a shared document: Vocational and technical students' approaches and experiences. In *Proceedings of the European Conference on e-Learning, ECEL* (pp. 56–64). Retrieved from https://search.proquest.com/docview/1968935738?accountid=11077%0A

Boelens, R., Voet, M., & De Wever, B. (2018). The design of blended learning in response to student diversity in higher education: Instructors' views and use of differentiated instruction in blended learning. *Computers & Education*, *120*, 197–212. http://doi.org/10.1016/j.compedu.2018.02.009

Collins, A. (1976). Processes in acquiring knowledge. In R. C. Anderson, R. J. Spiro, & W. E. Montague (Eds.), *Schooling and the acquisition of knowledge* (pp. 339–363). Hillsdale, NJ: Erlbaum.

Cronbach, L. J., & Snow, R. E. (1969). *Individual Differences in Learning Ability as a function of Instructional Variables*.

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2009). Structuring asynchronous discussion groups: The impact of role assignment and self-assessment on students' levels of knowledge construction through social negotiation. *Journal of Computer Assisted Learning*, *25*(2), 177–188. http://doi.org/10.1111/j.1365-2729.2008.00292.x

De Wever, B., Van Keer, H., Schellens, T., & Valcke, M. (2010). Roles as a structuring tool in online discussion groups: The differential impact of different roles on social knowledge construction. *Computers in Human Behavior, 26*(4), 516–523. http://doi.org/10.1016/j.chb.2009.08.008

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. *Three Worlds of CSCL: Can We Support CSCL?*, 61–91. http://doi.org/10.1007/s11165-004-8795-y

Dillenbourg, P., Järvelä, S., & Fischer, F. (2009). The evolution of research on computer-supported collaborative learning. From design to orchestration. In N. Balacheff, S. Ludvigsen, T. de Jong, A. Lazonder, & S. Barnes (Eds.), *Technology-Enhanced Learning* (pp. 3–19). Dordrecht: Springer Netherlands. http://doi.org/10.1007/978-1-4020-9827-7

Fahy, P. J., Crawford, G., Ally, M., Cookson, P., Keller, V., & Prosser, F. (2000). The development and testing of a tool for analysis of computer-mediated conferencing transcripts. *The Alberta Journal of Educational Research*, *46*(1), 85–88.

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, *17*(4), 397–431. http://doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG

Hämäläinen, R., & De Wever, B. (2013). Vocational education approach: New TEL settings—new prospects for teachers' instructional activities? *International Journal of Computer-Supported Collaborative Learning*, *8*(3), 271–291. http://doi.org/10.1007/s11412-013-9176-1

Hämäläinen, R., De Wever, B., Malin, A., & Cincinnato, S. (2015). Education and working life: VET adults' problem-solving skills in technology-rich environments. *Computers and Education, 88,* 38–47. http://doi.org/10.1016/j.compedu.2015.04.013

Hämäläinen, R., De Wever, B., Nissinen, K., & Cincinnato, S. (2017). Understanding adults' strong problem-solving skills based on PIAAC. *Journal of Workplace Learning*, *29*(7/8), 537–553. http://doi.org/10.1108/JWL-05-2016-0032

Hämäläinen, R., & Oksanen, K. (2012). Challenge of supporting vocational learning: Empowering collaboration in a scripted 3D game - How does teachers' real-time orchestration make a difference? *Computers and Education*, *59*(2), 281–293. http://doi.org/10.1016/j.compedu.2012.01.002

Henri, F. (1991). Computer conferencing and content analysis. In A. R. Kaye (Ed.), *Collaborative learning through computer conferencing: The Najaden papers* (pp. 117–136). London: Springer-Verlag. http://doi.org/10.1007/978-3-642-77684-7

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. http://doi.org/10.1080/00461520701263368

Kaendler, C., Wiedmann, M., Rummel, N., & Spada, H. (2014). Teacher competencies for the implementation of collaborative learning in the classroom: A framework and research review. *Educational Psychology Review*, *27*(3), 505–536. http://doi.org/10.1007/s10648-014-9288-9

Lazonder, A. W. (2005). Do two heads search better than one? Effects of student collaboration on web search behaviour and search outcomes. *British Journal of Educational Technology*, *36*(3), 465–475. http://doi.org/10.1111/j.1467-8535.2005.00478.x

Mayordomo, R. M., & Onrubia, J. (2015). Work coordination and collaborative knowledge construction in a small group collaborative virtual task. *The Internet and Higher Education*, *25*, 96–104. http://doi.org/10.1016/j.iheduc.2015.02.003

Onrubia, J., & Engel, A. (2012). The role of teacher assistance on the effects of a macro-script in collaborative writing tasks. *International Journal of Computer-Supported Collaborative Learning*, *7*(1), 161–186. http://doi.org/10.1007/s11412-011-9125-9

Otto, B., & Kistner, S. (2017). Is there a Matthew effect in self-regulated learning and mathematical strategy application? - Assessing the effects of a training program with standardized learning diaries. *Learning and Individual Differences*, *55*, 75–86. http://doi.org/10.1016/j.lindif.2017.03.005

Puntambekar, S., & Hübscher, R. (2005). Tools for Scaffolding Students in a Complex Learning Environment : What Have We Gained and What Have We Missed? *Educational Psychologist*, *40*(1), 1–12. Retrieved from https://pdfs.semanticscholar.org/553a/2d5453f69c00da4c6ebda5aeabb20210da42.pdf

Schwendimann, B. A., De Wever, B., Hämäläinen, R., & Cattaneo, A. A. P. (2017). CSCL and

Vocational Education: A Bond Worthy of Investigation? In B. K. Smith, M. Borge, E. Mercier, & K. Y. Lim (Eds.), *Making a Difference: Prioritizing Equity and Access in CSCL, 12th International Conference on Computer Supported Collaborative Learning (CSCL) 2017* (pp. 557–560). Philadelphia: International Society of the Learning Sciences. Retrieved from https://repository.isls.org/handle/1/282

Strijbos, J. W., Martens, R. L., Prins, F. J., & Jochems, W. M. G. (2006). Content analysis: What are they talking about? *Computers and Education*, *46*(1), 29–48. http://doi.org/10.1016/j.compedu.2005.04.002

van Merriënboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning: a systematic approach to four-component instructional design* (Second edi). New York: Routledge.

Vosniadou, S., Ioannides, C., Dimitrakopoulou, A., & Papademetriou, E. (2001). Designing learning environments to promote conceptual change in science. *Learning and Instruction*, *11*(4–5), 381–419. http://doi.org/10.1016/S0959-4752(00)00038-4

Weinberger, A., & Fischer, F. (2006). A framework to analyze argumentative knowledge construction in computer-supported collaborative learning. *Computers and Education*, *46*(1), 71–95. http://doi.org/10.1016/j.compedu.2005.04.003

Weinberger, A., Stegmann, K., & Fischer, F. (2007). Knowledge convergence in collaborative learning: Concepts and assessment. *Learning and Instruction*, *17*(4), 416–426. http://doi.org/10.1016/j.learninstruc.2007.03.007

Appendix A. Example of a process worksheet

Step 1: How to define dyslexia?

Subtask 1. Share the individual retrieved definitions of all group members, and note them below. Which definition do you prefer, and explain why you prefer this definition?

Step 2: How to diagnose dyslexia?

Subtask 2. Who should you refer to as a teacher in the case of a suspicion of a pupil with dyslexia?

Subtask 3. What are the diagnostic criteria for dyslexia? Explain in your own words.

Subtask 4. Which diagnostic criteria do you recognize in this case? Provide the criterion and explain why this is an example of the criterion. [case description]

Subtask 5. What are – according to you – the most important advantages and disadvantages of the diagnosis dyslexia? Provide two advantages and two disadvantages. Explain in your own words.

Step 3: What are the signs and symptoms of dyslexia?

Subtask 6. Provide five examples of behavior that may indicate dyslexia.

Subtask 7. Some people might try to hide their learning disability. Why would they do that? Explain in your own words.

Step 4: How many people are affected by dyslexia?

Subtask 8. How many people are affected by dyslexia in Flanders?

Subtask 9. Is dyslexia more often diagnosed in boys than girls?

Subtask 10. Is dyslexia hereditary? Why (not)? Explain in your own words.

Subtask 11. How big is the chance that there will be a pupil with dyslexia in your classroom? Explain your answer.

Step 5: What are the consequences of dyslexia for pupils?

Subtask 12. What are the consequences of dyslexia for the daily life? Provide three examples.

Subtask 13. What are the consequences of dyslexia for your specific school subjects? Provide three examples and write down the specific school subject (i.e. a school subject related to one of the group member's field)

Subtask 14. What are the consequences of dyslexia for learning in general? Provide three examples.

Subtask 15. What are the consequences of dyslexia for the social-emotional functioning of the person? Provide three examples.

Subtask 16. Do you know people with dyslexia? Does this person experience difficulties in the daily life or at school? If yes, which ones? If no, why not?

Step 6: How best to treat dyslexia on school-level?

Subtask 17. The school will provide a structural approach for pupils with dyslexia. This case describes such an approach. [case description] How would you as a teacher involve the pupil and his/her parents in such an individual supportive plan? Why would you do it that way? Provide arguments.

Step 7: How best to treat dyslexia on classroom-level?

Subtask 18. What supportive measures can you implement on the classroom-level? This means that all pupils can make use of these supportive measures, not only the pupils with a diagnosis of dyslexia. Search for 5 examples that are applicable in your own field. Write down the supportive measure and indicate for which school subject this applies.

Subtask 19. What supportive measures can you implement on the pupil-level? Search for each STICORDI (stimulation, compensation, remediation, and dispensation) measure five examples that are applicable in your own field.

Subtask 20. Think about what you knew about dyslexia before performing this task. Did you have prejudices regarding dyslexia? Explain.

Appendix B: Excerpts of the dialogues

1. Pattern A: Gradual support

Teacher	**Are you proceeding well?**
Esra	No, we are arguing
Teacher	**That's good! If it is a qualitative discussion… What is the discussion about?**
Esra	Well, does dyscalculia occurs more often in boys than in girls? Victor and Lucas found...
Lucas	I found different things: more often in boys, more often in girls, or as much as in boys as in girls
Esra	Well, they both found that it occurs as much as in boys as in girls. I had written that I did not find a reliable source, but that I found that it occurs more often in boys. My source stated that a certain part of the brain develops more slowly in the brains of boys or girls, and therefore it occurs less.
…	
Lucas	And the website I consulted was really a reliable source, but it stated something else than in my book, which was also a reliable source.
Victor	I also used a book as source
Teacher	**Which book did you use?**

Lucas	The "dyscalculie survivalgids", I received it from the school of my daughter
Victor	And I have read the book: "Ik hoor er ook bij"
Teacher	**And what information did you find in these books? As much as in boys as in girls?**
Esra	Yes, both sources stated that it occurs as much as in boys as in girls.
…	
Lucas	But which information do we now include in the task?
Esra	That it occurs just as often in boys as in girls?
Teacher	**What can you say? I would say the truth**
Lucas	What is the truth?
Esra	That I have read on a forum that it occurs more often in boys than in girls
Teacher	**And do you think a forum is a reliable source?**
Esra	It was a forum where parents who had children with dyscalculia shared information
Lucas	I wouldn't say that is reliable
Esra	Indeed, they go to different doctors and are provided with different information
Teacher	**And what would be a reliable source to shape your answer to the question?**
Lucas	The university?
Victor	Well, my source was a text written by a professor at Ghent University
Esra	Ok, then, do you agree if I delete my answer and take your answer, that there is scientific research that states that most of the researchers agree that dyscalculia occurs as often in boys as in girls?

2. Pattern B: The devil's advocate

Mia	How much was it, how many percent in Flanders?
Olivia	1% of the people living in Flanders
Teacher	**1%? Is it not more than 1%?**
Mia	We found 40.000 to 60.000 people
Teacher	**Yes, would that be more than 1 in 100?**
Mia	It is about the population in Flanders, we all three retrieved the same information
Olivia	Yes, I found that
Mia	And I found that too, it is about Flanders, so yes
…	
Teacher	**And is this about people with a diagnosis? Or an estimate?**
Mia	An estimate
Anna	Yes, because there isn't much research available about this
Teacher	**This is a difficult thing, because it appears that there are many people who have ASD, but are not diagnosed, and there are also people with a diagnosis who do not have it…**
Olivia	(provides no answer to the teacher) You also have to mention your source here

3. Pattern C: Sharing general advice

Teacher	**You consulted the same sources?**
Sophie	Yes
Emma	Most of the time, because there were not many good websites
Teacher	**Ok, don't forget to elaborate on what you have found and to discuss about your opinions**
Emma	We do that all the time

4. Pattern D: Appointing group members who are less actively participating

Teacher	**Yes, and what was Anna her answer? Yours is very detailed and applied to one person**
Olivia	I thought we needed to provide such an example…
Anna	(ignores the question) In another subtask we have to provide examples of behavior, so maybe we can use her example there?
Teacher	**What is your opinion Anna? Do you agree?**

Anna	Yes, we all have the same opinion

5. Pattern E: Question and answer

Teacher	**What could you do more?**
Charlotte	A conversation
Teacher	**And when would you propose a conversation?**
Charlotte	At the parents' evening
Teacher	**Yes, for instance, or maybe, when the guidance plan is introduced?**
Charlotte	Yes, there should also be a conversation at the startup
Eva	Ok, I will write that down
	…
Teacher	**And would it be interesting to involve the parents and the pupil to adjust the plan?**
Eva	Regular meetings should be organized
Charlotte	Yes
Eva	Also with the pupil? Or only the parents?
Teacher	**What should you do?**
Eva	Yes
Teacher	**Why do you think that?**
Eva	The pupil can say what works or works not for him/her, I think
Charlotte	Yes

Appendix C: Background information on the groups' collaboration process (based on observations, interviews, and an analysis of the revision history of the document)

1. Background information about the collaboration process of group 1

In group 1, three students collaborated on the topic dyscalculia. The group seemed very motivated to deliver a good end product. Esra spontaneously took the role as leader during the group work: she led the discussion and took notes in the group document. The other two group members actively participated during the group discussion. However, Lucas participated spontaneously in the group discussion, while Victor was more hesitant. The analysis of the group discussion (see Figure 1) showed that the group members often shared and compared their ideas and information (level 1), identified (in)consistencies between ideas and information (level 2), and negotiated about their arguments or co-constructed knowledge (level 3).

2. Background information about the collaboration process of group 4

In group 4, three group members collaborated on the topic ASD (autism spectrum disorder). Olivia spontaneously took the role as leader during the group work: she mostly led the discussion and took notes in the group document. Mia actively participated during the discussions, while Anna was rather quiet. In particular, Anna also had her laptop in front of her to make adaptations related to the lay-out of the group document, but she also did other things that were not related to the group work. Olivia and Mia tried to involve Anna by frequently asking her opinion. The group, and especially Olivia and Mia, seemed motivated to successfully complete the group work. This resulted in group discussions wherein students often shared and compared their ideas and information (level 1), identified (in)consistencies between ideas and information (level 2), and sometimes negotiated about their arguments or co-constructed knowledge (level 3) (see Figure 1).

3. Background information about the collaboration process of group 2

Group 2 consisted of two group members working on the topic NLD (non-verbal learning disorder). Both students seemed motivated to deliver a good end product. Regarding note taking in the group document, students alternated roles: Emma took notes during the online collaboration, while Sophie took notes during the face-to-face collaboration. There was no leader in the group. During the collaboration process, both students displayed their individually retrieved information, and together they decided about the final solution. The students themselves indicated that they often agreed with each other and thus had not a lot to discuss about, as they had found quite similar information. This is also reflected in the analysis of their interaction, as this group's interaction is especially situated at the less advanced levels of collaborative knowledge construction, i.e. level 1 and level 2 (see Figure 1).

4. Background information about the collaboration process of group 3

In group 3, two students collaborated on the topic ADHD (attention deficit hyperactivity disorder). The group spent a lot of time on off-topic talk. Eva took the lead, while Charlotte was busy doing other things (i.e. preparing the next lesson). They seemed less motivated than the other groups to invest time in the task. Consequently, little time was spent on discussing about the content, and Eva took the group task for her own account. They went over the task quickly, and their discussion was especially situated in level 1, sharing and comparing information (see Figure 1).

Conjecture mapping to support hands-on adult learners in open-ended tasks

Abstract

This case reports on a teacher education course that aimed to support vocationally educated adults, referred to as hands-on learners, to accomplish open-ended tasks. Conjecture mapping was used to identify the most salient design features, and to test if, how, and why these course features supported learners. Inspired by ethnographic approaches, sustained engagement and multiple data sources were used to explain the effects of the course design on participants' behavior and perceptions: student and teacher interviews, observations, and artifacts. The results reveal that almost all of the proposed design features stimulated the participants toward the intended enactment processes, which in turn yielded the intended learning outcomes. For instance, worked examples (i.e., design feature) not only engendered the production of artifacts that meet high standards (i.e., enactment process) because they clarify the task requirements, but also fostered a safe structure (i.e., enactment process) by providing an overall picture of the task. The conjecture map resulting from this study provides a theoretical frame to describe, explain, and predict how specific course design features support hands-on adult learners in open-ended tasks, and assists those who aim to implement open-ended tasks in similar contexts.

1. Introduction

Higher and adult education institutions are not only responsible for ensuring that students develop conceptual and procedural knowledge, but also for developing lifelong learners who are able to respond to the changing needs of working life and empowered to engage in opportunities for continuing improvement (Damşa & Nerland, 2016; Hämäläinen, De Wever, Malin, & Cincinnato, 2015; Nerland, 2012). One of the most important skills these students need today is the ability to tackle open-ended tasks (Hämäläinen et al., 2015; Könings, Brand-Gruwel, & van Merriënboer, 2005). Tackling open-ended tasks entails solving realistic and relevant problems (van Merriënboer & Kirschner, 2013), by generating, critically examining, sharing, and using knowledge (Damşa & Nerland, 2016), while several diverse answers or solutions are possible (Hannafin & Hill, 2007).

Developing the ability to tackle open-ended tasks can be enabled by problem-centered instruction (e.g., Merrill, 2007), complex learning tasks (e.g., van Merriënboer, Clark, & de Croock, 2002), inquiry-based learning tasks (e.g., Edelson, Gordin, & Pea, 1999; Hmelo-Silver, Duncan, & Chinn, 2007), or resource-based learning tasks (e.g., Hannafin & Hill, 2007). All these approaches have several characteristics in common. First, they place high value on authentic, realistic, and relevant learning (Edelson et al., 1999; Hmelo-Silver et al., 2007; Merrill, 2007; van Merriënboer & Kirschner, 2013). Second, students are prompted to take an active role in their own learning process, being tasked to critically examine sources, meaningfully use sources, develop data-supported explanations, and communicate their ideas, opinions, and knowledge (Hannafin & Hill, 2007; Hmelo-Silver et al., 2007). Third, the teacher fulfills the important role of facilitating and supporting the learning process (Merrill, 2007). And fourth, these approaches often encourage collaborative learning as a way to stimulate active knowledge construction processes (Hmelo-Silver et al., 2007; Könings et al., 2005).

While the abovementioned instructional approaches can support development of the ability to tackle open-ended tasks in general, they may work differently with specific groups of students. In particular, students lacking experience in carrying out open-ended tasks often find it too difficult without appropriate guidance and support (van Merriënboer, 2013). One population of such students is that of adult learners with a background in vocational or technical secondary education (Hämäläinen et al., 2015). These students, referred to here as hands-on adult learners, are often accustomed to focusing mainly on the acquisition of practical skills. Typically, these students learn through hands-on experiences, demonstrations, and practice, and are less familiar with tackling open-ended tasks and problem-solving skills (Hämäläinen et al., 2015; Smith, 2001). Consequently, previous research has shown that hands-on adult learners might encounter several obstacles in learning environments that include open-ended tasks. Regarding cognition, for example, teachers have indicated that hands-on adult learners often struggle to analyze task demands, identify core information, or structure and summarize information (Biemans et al., 2016; Boelens, Voet, & De Wever, 2018). Concerning behavior, it might be challenging for hands-on learners to devote sufficient time to perform the tasks (Boelens et al., 2018) or to manage effort to complete a task or course (Biemans et al., 2016). With regard to hands-on learners' motivation, research has found that they often have low self-efficacy beliefs (Dubeau, Plante, & Frenay, 2017; Tsai & Shen, 2009), which can have negative effects on performing or completing (open-ended) tasks which they do not feel competent enough to perform.

To accommodate these cognitive, behavioral and affective struggles, a number of studies have argued that, especially for students unaccustomed to open-ended tasks, appropriate guidance and support is required (Brand-Gruwel, Wopereis, & Vermetten, 2005; Damşa & Nerland, 2016; Nadolski, Kirschner, & van Merriënboer, 2005; van Merriënboer & Kirschner, 2013). For instance, research has found that students performed better and more efficiently when the open-ended task was split in a (limited) number of phases (Nadolski, Kirschner, & van Merriënboer, 2006). In addition, students receiving driving questions to carry out learning tasks also performed better, although not more efficiently (Nadolski et al., 2006). Another frequently mentioned supportive design feature is the use of worked examples (e.g., van Gog, Paas, & van Merriënboer, 2008), which seem to have efficiency benefits for students (McLaren, Van Gog, Ganoe, Karabinos, & Yaron, 2016).

Focus of the study

Prior research has thus revealed several course design features that can have an influence on outcomes. However, these design features are often studied in isolation from each other and the focus is mainly on whether a specific feature was effective to reach the outcomes. As such, it remains unclear how course design features function in concert with each other in a designed environment. Consequently, teachers are left with little guidance as to which features of learning designs should be prioritized when trying to find an appropriate balance between support and autonomy when designing open-ended tasks that address the needs of hands-on adult learners. The present study seeks to address this gap by identifying the salient features of a course design and the enactment processes they engender, making it possible for hands-on adult learners to reach the learning outcomes.

The current study was conducted in a course on psycho-pedagogical competence for hands-on adult learners enrolled in a teacher training program to become teachers in vocational secondary education. The main learning objectives of the course were that students will be able to recognize the most common learning and developmental disabilities, to state the consequences for pupils, and to develop classroom practices that address the needs of pupils with learning or developmental disabilities (further details are provided under Methods).

To achieve the study's aim, conjecture mapping was used as a tool to generate and test the theoretical basis of the course design, with the goal of affirming, rejecting, or refining initial hypotheses (McKenney & Reeves, 2018; Sandoval, 2014). A conjecture map reveals causal relationships between design features, the processes they engender during enactment, and the resulting outcomes (Figure 1 presents an overview of the conjecture map elements used in this study). Design features (in this case, characteristics of the pre-service teacher course for hands-on learners) can be described in terms of: (1) materials and resources (MR), which are the physical artefacts that are part of the intervention; (2) activity and task structures (AT), which describe the main events through which the intervention will be carried out; and (3) participation and practices (PP), including norms and expectations for how actors are to engage during those events (McKenney & Reeves, 2018; Sandoval, 2014). Design features are purposefully selected. Providing the theoretical basis for their selection, *design conjectures (DC)* articulate the processes that specific design features (should) engender. For example, "authentic tasks cause learner engagement because they are intrinsically motivating" conjectures that the design feature (authentic tasks) will yield a specific process during enactment (engagement) and offers motivation as an explanation. Enactment processes (i.e., called mediating processes by Sandoval (2014)) result from design features, and contribute to learning outcomes. They can be reified in three ways (Salomon, 1996; Sandoval, 2014): (1) participant artifacts (PA), such as the products that students generate from their activities; (2) observable interactions (OI), which are students' interactions that emerge from the design; and (3) participant experiences (PE), including how the learning environment is experienced by students. Typically, these processes are purposefully planned. Analogous to design conjectures, *process conjectures (PC)* (i.e., called theoretical conjectures by Sandoval (2014)) articulate why the enactment processes (should) yield certain outcomes. For example, "learner engagement supports disciplinary understanding by keeping students on-task" conjectures that the enactment process (engagement) will yield a specific outcome (disciplinary understanding), and offers time-on-task as an explanation. Finally, the enactment processes mediate learning outcomes (Sandoval, 2014), which can be cognitive (C) (e.g., content-knowledge), behavioral (B) (e.g., persistence), or affective (A) (e.g., interest).

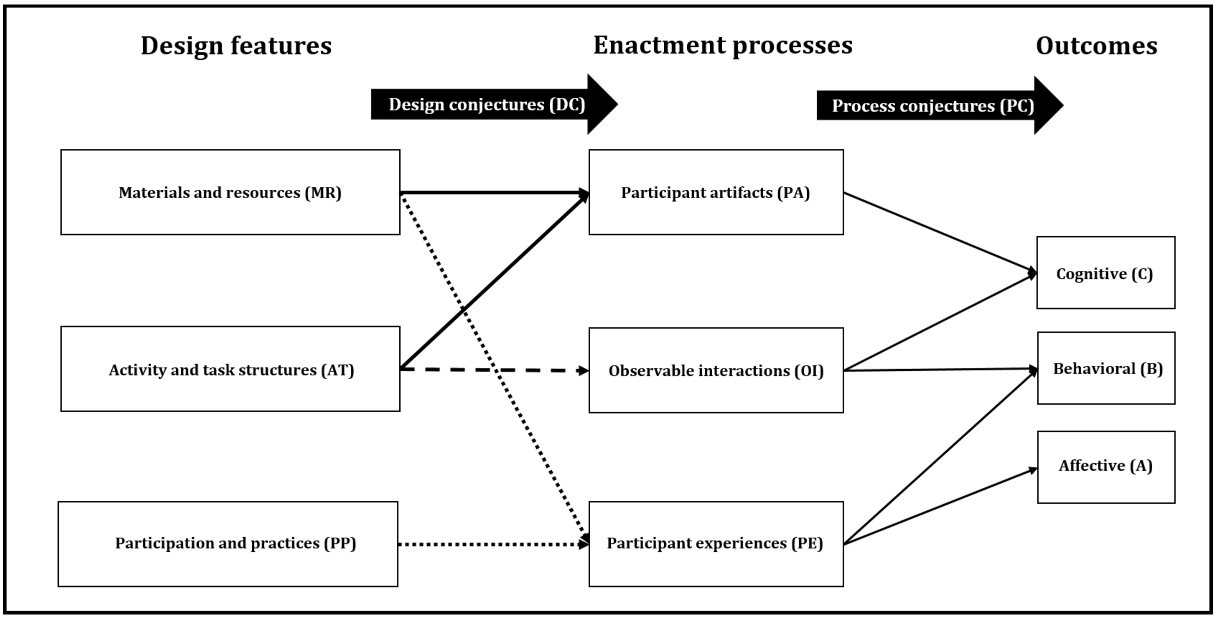

Figure 1. Schematic overview of this study's conjecture map (adapted from Sandoval (2014)).

Theoretical framework

The goal of the present study was to identify the salient features of a balanced course design that would yield the kinds of processes during enactment that are needed to benefit the learning outcomes of hands-on adult learners. We therefore take the desired enactment processes as starting points for building the conjecture map. For each enactment process, we describe what it is, why it is important for hands-on learners' learning outcomes (i.e., process conjecture), and how and why specific design features (materials/resources, activity/task structures, or participation/practices) could contribute to that enactment process (i.e., design conjectures). At the end of this section, a detailed representation of the relations between the course design features, the enactment processes, and the outcomes, is presented in the resulting conjecture map (Figure 2).

Enactment processes: participant artifacts

Produce artifacts that meet high standards

The first enactment process is students' production of artifacts that meet high standards. The creation of an artifact entails, for example, that students have to gather and process information about a case study and produce a written report to show their performance. Creating artifacts affords opportunities to engage with the learning content and make sense of the instruction (Sandoval, 2014). Further, producing artifacts that meet high standards requires students to master the subject matter (Pintrich, 2000). This enactment process thus stimulates cognitive processing and can help achieve the cognitive outcomes, namely reaching the learning objectives related to the course. In sum, the first process conjecture states (PC-PA-1): *if students produce artifacts that meet high standards, they will achieve the cognitive learning objectives*.

Creating artifacts that meet external expectations can be facilitated by three design features, namely: worked examples, process worksheets, and the explicit task of assessing the reliability of the sources used. *Worked examples*, which are specific demonstrations of the task students are required to do, provide students with acceptable solutions and useful solution steps (van Merriënboer & Kirschner, 2013). This explicates the task demands and provides a valid standard against which students can compare their own work (Nicol & Macfarlane-Dick, 2006). As such, worked examples can lead to producing artifacts that meet high standards because they help students to become familiar with the task requirements and standards (Nicol & Macfarlane-Dick, 2006; van Merriënboer & Kirschner, 2013; Vandewaetere et al., 2014). Providing *process worksheets* guide students through the different steps they need to take to perform the task (Nadolski, Kirschner, Merriënboer, & Hummel, 2001), as it segments the open-ended task in problem-solving steps and accompanying guiding questions (Nadolski et al., 2006; van Merriënboer & Kirschner, 2013). These questions can, for instance, guide students to select the most relevant information, or cluster the information in a specific way (Brand-Gruwel & Gerjets, 2008; Nadolski et al., 2006). Accordingly, process worksheets help ensure that students search for and present information in a structured way, which makes it more likely to meet the high standards. Finally, the explicit task of *assessing the reliability of sources* entails that students scan their information sources to decide whether the retrieved information is useful, relevant, and reliable (Brand-Gruwel, Wopereis, & Walraven, 2009). This is essential to gather and process correct knowledge, and thus to develop products that meet high standards. To sum up, the first design conjecture is (DC-PA-1): *the use of worked examples (MR), and/or process worksheets (MR), and/or the explicit task to assess the reliability of the sources (AT) will help students to produce artifacts that meet high standards*.

Summarize key ideas

The second enactment process related to participant artifacts is summarizing the key ideas. This includes, for example, that students reduce the learning material substantially by capturing the key information (King, 1992) in a summary of the case study on which they have worked. Actively retrieving and recalling information by means of creating an accurate summary might help students to reach the learning objectives. To be more specific, as long as the summary is accurate, this task supports knowledge consolidation (i.e., strengthen students' knowledge and remember the content) (Bråten & Samuelstuen, 2007; King, 1992; Roediger & Butler, 2011). Stimulating students to engage in this enactment process to process the learning material is especially important for hands-on learners, as both teachers in this context and previous research have indicated that these students often have not yet developed their own effective strategies for processing the content (Boelens et al., 2018; King, 1992). We therefore conjecture (PC-PA-2): *if students accurately summarize the key ideas of the topic, they will achieve the cognitive learning objectives*.

This enactment process can be elicited by a pre-structured summarizing task (e.g., students compile a leaflet for their fellow students) in which they are required to capture the core information about their case study. The pre-structured summarizing task helps students to summarize the key ideas because students have to actively run through and reconsider the content in order to extract, recapitulate, and present the main points (Weinstein, Husman, & Dierking, 2000). As such, the second design conjecture is (DC-PA-2): *a pre-structured summarizing task (AT) will help students to accurately summarize the key ideas*.

Enactment processes: observable interactions

Negotiation

Negotiation during collaboration implies that all group members are engaged with the learning content by discussing about ideas, opinions, and (the meaning of) concepts to reach consensus and construct shared understanding (Dillenbourg, 2002). This enactment process emerges when group members not only share and compare their information and ideas, but also treat each other's opinions critically through explaining, arguing, or questioning one another (Gunawardena, Lowe, & Anderson, 1997; Onrubia & Engel, 2012). Negotiation is important during collaboration to support the cognitive outcomes, namely achieving the learning objectives, as these processes of explaining, critiquing, or questioning lead to more advanced comprehension of the content or the co-construction of new knowledge (Barron, 2003; Dillenbourg, 2002; Onrubia & Engel, 2009). Previous research in vocational education settings has shown that this enactment process often does not occur spontaneously during collaboration (Hämäläinen & Oksanen, 2012). In addition, hands-on adult learners are often not familiar with collaborative (open-ended) tasks (Hämäläinen et al., 2015; Smith, 2001), and may thus need explicit guidance to develop effective ways to jointly build knowledge. The process conjecture states (PC-OI-1): *if students engage in negotiation during collaboration, they will achieve the cognitive learning objectives*.

Negotiation during collaboration can be facilitated by four supportive features, namely: a pre-structured summarizing task, an individual idea generation task, the explicit task of assessing the reliability of sources, and discursive practices that are characterized by collaboration, support, and easy access to each other's expertise. The *pre-structured summarizing task* mentioned previously could also elicit negotiation if done collectively, as the group has to reconsider their previous work and decide together upon the most essential information to be enclosed in the overview. In addition, the *individual idea generation task* implies that, prior to working together as a group, students first have to explore and generate ideas individually (Onrubia & Engel, 2009, 2012). This can facilitate negotiation during collaboration because students will be inclined to share their individual ideas with one another, after which they need to reach consensus by comparing and contrasting ideas, expressing novel ideas, and discussing about the meaning of concepts. The explicit task to *assess the reliability of sources* requires students to scan each information source based on a predefined set of criteria. This activity can foster negotiation, as it provides students with a specific line of reasoning or argument when they are engaged in a discussion. For instance, when students encounter disagreements, they can identify and compare the reliability of the information sources to reach consensus based on sound arguments (Gunawardena et al., 1997). A last feature of the course design that aims to foster negotiation is *discursive practices characterized by collaboration, support and easy access to each other's expertise*. This design feature is important, as negotiation is only possible when students are willing to share their ideas and opinions with each other, and have a feeling of shared responsibility (Arvaja & Pöysä-Tarhonen, 2013; Van Oers & Hännikäinen, 2001). For example, a learning climate in which there is no threshold to approach fellow students or the teacher and in which it is allowed to make mistakes can foster negotiation. Altogether, the third design conjecture states (DC-OI-1): *a pre-structured summarizing task (AT), and/or an individual preparatory task (AT), and/or the explicit task to assess the reliability of the information sources*

(AT), and/or a learning environment that is characterized by collaboration, support and easy access to each other's expertise (PP) will help students to engage in negotiation during collaboration.

Pace aligning actions

Another salient enacting process related to observable interactions is that students align the pace of their progress with each other. For example, before starting the group work, it is important that students both generated individual ideas and finished their individual preparatory task. Aligning the pace of their progress is important to ensure that all students persist during the open-ended task. Student persistence involves students' attempts to continue to participate in a course or task, despite the presence of possible obstacles (Pintrich, 2004; Rovai, 2003). It is an important behavioral outcome, that is related to *doing* (Reeves, 2011), and students aligning the pace of their progress support this outcome by diminishing possible obstacles such as lack of self-discipline, bad time planning, or allocating insufficient time for performing the tasks (Pintrich, 2004; Rovai, 2003). This enactment process is especially important for hands-on adult learners as research has indicated that these students might show unstructured learning behavior and often lack a sense of time, resulting in poor time planning skills and devoting insufficient time to the tasks (Biemans et al., 2016; Boelens et al., 2018; Jossberger, Brand-Gruwel, van de Wiel, & Boshuizen, 2015). Applied to the current case, the process conjecture is (PC-OI-2): *if students engage to align pace with each other, they will persist throughout the course*.

This enactment process can be facilitated by dividing the open-ended task into subtasks with accompanying *intermediate deadlines, and reminding* students of the upcoming deadlines and their learning progress. Reminders can be addressed to individual students, student groups, or the whole class. The provision of intermediate deadlines and reminders is important, as these might decrease students' concerns about the planning of the course and the allocation of sufficient study time for the subtasks (Karoğlu, Kiraz, & Ozden, 2014), and keep them engaged throughout the course design (X. Liu, Bonk, Magjuka, Lee, & Su, 2005), both of which could contribute to students aligning the pace of their progress with one another. Consequently, the fourth design conjecture is (DC-OI-2): *intermediate deadlines and reminders (AT) will prompt students to align their pace with each other*.

Enactment processes: participant experiences

Students experience a safe structure

The first enactment process related to participant experiences, is that students experience a safe structure in terms of: course and task details; sequencing and structuring of the (sub)tasks; and perceived workload and task difficulty (Nadolski et al., 2006; Smith, 2001). This holds for both behavioral and affective outcomes. First, regarding the behavioral outcomes, research has suggested that ensuring that all tasks, requirements, and procedures are clearly formulated and communicated fosters student persistence by avoiding frustrations about the consistency and the clarity of the course design (Rovai, 2003; Workman & Stenard, 1996). Second, experiencing a safe structure might also be associated with affective outcomes, namely motivation. Students' motivational beliefs about themselves in relation to the task (Pintrich, 2000) include: perceptions of task difficulty or feasibility; beliefs about the importance, relevance, and utility of the task; interest or liking of the task; and positive and negative affective reactions to the self or the task (Pintrich, 2000, 2004; Ryan & Deci, 2000). Motivational beliefs are concerned with *valuing* (Reeves, 2011), as to be motivated is to be engaged to reach the goals of the course design (Ryan & Deci, 2000). Experiencing a safe structure can support students' motivational beliefs positively because they feel more empowered in a safe structure to tackle challenging tasks. The experience of a safe structure is an important process for hands-on adult learners, as they often lack experience in open-ended tasks and might feel not competent enough to perform these tasks (Dubeau et al., 2017). In addition, research has shown that these students prefer learning in a structured environment, with a well-organized course structure and clear expectations (Jossberger, Brand-Gruwel, van de Wiel, & Boshuizen, 2018; Smith, 2000, 2001). As such, the process conjecture is (PC-PE-1): *if students experience a safe structure, they will persist throughout the course and/or they will hold positive motivational beliefs*.

The experience of a safe structure can be facilitated by three design features, namely: worked examples, process worksheets, and the provision of intermediate deadlines and reminders. *Worked examples* ensure that the students experience a high level of support, since they can consult a complete elaboration of the open-ended task (van Merriënboer & Kirschner, 2013). Studying a worked example ensures that students do not have to search for their own strategies or method to complete the open-ended task (van Merriënboer, 2013). In addition, when students encounter difficulties during task performance, they can go back to the worked example to see how an expert approached a certain part. This can lead to perceptions of a safe task structure. *Process worksheets* ensure that the open-ended task is clearly structured and sufficiently detailed (Nadolski et al., 2006; van Merriënboer & Kirschner, 2013). This fosters students' experience of a safe structure by dividing the task in subsequent steps and accompanying guiding questions, which direct students' attention to search for the most relevant information, to cluster this information according to the imposed structure, and to report the sources they consult. At last, *intermediate deadlines and reminders* emphasize the division of the open-ended task in subtasks. In particular, this design feature provides students with a timeline in which fixed deadlines are set, and frequently reminds students about upcoming deadlines and their learning progress. This influences students' perceptions of a safe structure by decreasing concerns about the course structure and avoiding that students are overwhelmed by the amount of work (Karoğlu et al., 2014). Accordingly, the fifth design conjecture states (DC-PE-1): *the use of worked examples (MR), and/or process worksheets (MR), and/or the provision of intermediate deadlines and reminders (AT) will help students to experience a safe structure*.

Students experience autonomy

The second enactment process related to participant experiences, is the extent to which students experience autonomy while engaged in the course design. Experiencing autonomy refers to the psychological freedom that students feel and exert, to control and take responsibility for their own behavior and the learning activities they engage in (Deci & Ryan, 2000; Haerens, Vansteenkiste, Aelterman, & Van den Berghe, 2016). This is an important process that influences motivational beliefs such as positive affective reactions and liking the task because students are intrinsically motivated and satisfied (Deci & Ryan, 2000; Ryan & Deci, 2000). In the case of hands-on adult learners, research has indicated that although these students prefer a lot of structure and teacher control, they also want authentic learning tasks that challenge them (Jossberger et al., 2018). Accordingly, this process conjecture is (PC-PE-2): *if students experience autonomy when engaged in the open-ended task, they will hold positive motivational beliefs*.

In education, student autonomy can be fostered when students are expected to *take charge of their own learning process*. In this case, this design feature entailed both online learning activities, and the autonomy in the open-ended task. On the one hand, students have to take charge of their own learning by means of online learning activities, wherein students have control over when, what, where, and how long to engage in a learning activity (Barnard, Lan, To, Paton, & Lai, 2009). On the other hand, in open-ended tasks, students are required to take an active role in the learning process by acquiring and processing new content on their own, and presenting their work to others (Drexler, 2010). This influences students' experience of autonomy by emphasizing their own responsibility to perform the tasks successfully. As such, the sixth design conjecture assumes (DC-PE-2): *students who are expected to take charge of their own learning process (PP) will experience autonomy*.

Students experience a sense of community

The third enactment process related to participant experiences is students experiencing a sense of community. This process is characterized by group cohesion, which is a sense of shared identity or belonging to the class group (Arbaugh et al., 2008; Garrison, 2007; Green, Preston, & Sabates, 2003), open communication in a safe learning environment, and the development of personal relationships while acknowledging each other's individual personality (Arbaugh et al., 2008;

Garrison, 2007). This enactment process is important with regard to the behavioral and affective outcomes. First, research has suggested that personal interaction among students and between students and the teacher, and the creation of a safe learning community can stimulate student persistence (Rovai, 2003). In particular, a sense of community provides students who are unfamiliar with open-ended tasks with a more secure feeling (Ausburn, 2004) and can reduce certain obstacles or thoughts of dropout from the course (Pintrich, 2004). Next, Deci and Ryan (2000) have stated that feeling secure and connected to others in the classroom is important for students to be intrinsically motivated, what causes positive affective reactions to the self or the task. This is an important enactment process for hands-on adult learners as research in this context has found that vocational learners have a strong preference for learning in an environment with welcoming and supportive relationships between the students themselves, and the students and the teacher (Smith, 2000). The process conjecture is (PC-PE-3): *if students experience a sense of community, they will persist throughout the course and/or they will hold positive motivational beliefs*.

Two design features that aim to foster the experience of a sense of community are: intermediate deadlines and reminders, and discursive practices that are characterized by collaboration, support, and easy access to each other's expertise. The provision of *intermediate deadlines and reminders* shows a high level of teacher commitment regarding the students. In particular, students may feel supported when the teacher reminds them of upcoming deadlines and students' learning progress, but may also have the feeling that the teacher is available and accessible. In short, this can cause a sense of community by the commitment and the proximity of the teacher (Ausburn, 2004; McDonald, 2014). The second design feature includes *environments characterized by collaboration, support, and easy access to each other's expertise.* This design feature mentioned previously is also important to create a sense of belonging in a community, as this is only possible when both the teacher and the students show commitment towards each other (Tomlinson & Imbeau, 2013). Accordingly, the seventh design conjecture is (DC-PE-3): *the provision of intermediate deadlines and reminders (AT) and/or an environment that is characterized by collaboration, support, and easy access to each other's expertise (PP) will prompt students to experience a sense of community*.

Summary

Based on the above outlined theoretical framework, seven design conjectures are articulated based on the association between the features of the course design and the enactment processes. In addition, also seven process conjectures are formulated based on the association between the enactment processes and the learning outcomes. Table 1 presents the design and process conjectures for each enactment process.

Table 1

Design and process conjectures in this study

Enactment process	Design conjecture	Process conjecture
Produce artifacts that meet high standards	***DC-PA-1***: The use of worked examples, and/or process worksheets, and/or the explicit task to assess the reliability of the sources will help students to produce artifacts that meet high standards	***PC-PA-1***: If students produce artifacts that meet high standards, they will achieve the cognitive learning objectives
Summarize key ideas	***DC-PA-2***: A pre-structured summarizing task will help students to accurately summarize the key ideas	***PC-PA-2***: If students accurately summarize the key ideas of the topic, they will achieve the cognitive learning objectives
Negotiation	***DC-OI-1***: A pre-structured summarizing task, and/or an individual preparatory task, and/or the explicit task to assess the reliability of the information sources, and/or a learning environment that is characterized by collaboration, support and easy access to each other's expertise will help students to engage in negotiation during collaboration	***PC-OI-1***: If students engage in negotiation during collaboration, they will achieve the cognitive learning objectives
Pace aligning actions	***DC-OI-2***: Intermediate deadlines and reminders will prompt students to align their pace with each other	***PC-OI-2***: If students engage to align pace with each other, they will persist throughout the course
Students experience a safe structure	***DC-PE-1***: The use of worked examples, and/or process worksheets, and/or the provision of intermediate deadlines and reminders will help students to experience a safe structure	***PC-PE-1***: If students experience a safe structure, they will persist throughout the course and they will hold positive motivational beliefs
Students experience autonomy	***DC-PE-2***: Students who are expected to take charge of their own learning process will experience autonomy	***PC-PE-2***: If students experience autonomy when engaged in the open-ended task, they will hold positive motivational beliefs

Students experience a sense of community	***DC-PE-3***: The provision of intermediate deadlines and reminders and/or an environment that is characterized by collaboration, support, and easy access to each other's expertise will prompt students to experience a sense of community	***PC-PE-3***: If students experience a sense of community, they will persist throughout the course and they will hold positive motivational beliefs

A map showing how the course design features are related to the enactment processes and the learning outcomes is shown in Figure 2. The white zone represents the primary focus of this study. Regarding the enactment processes, we aim to be parsimonious and only connect the embodied design features with the most obvious enactment processes. Although these processes are discussed here in isolation from each other, we note that some of them may also influence each other reciprocally. However, such analyses extend beyond the scope of the present study.

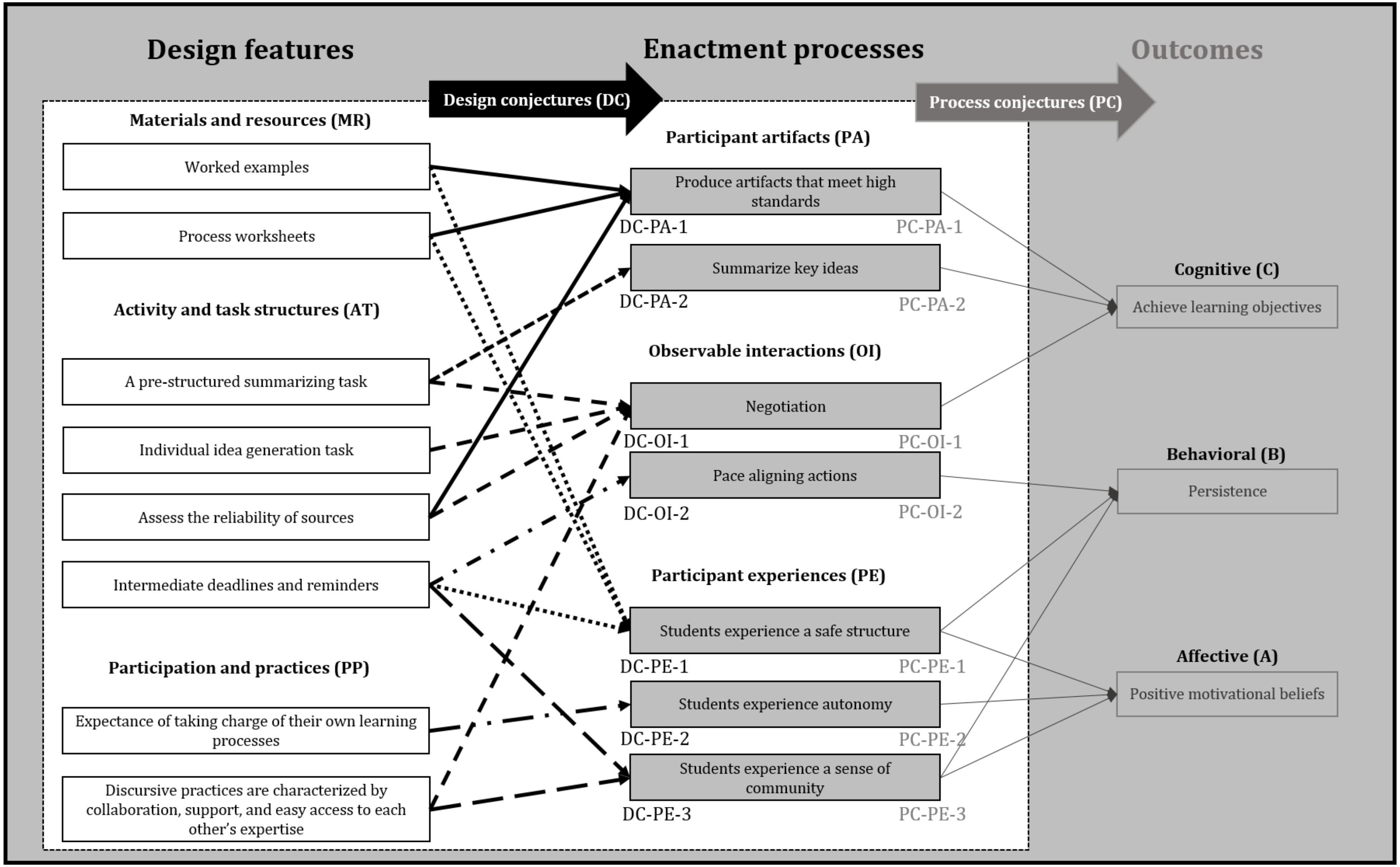

Figure 2. Design conjectures for supporting hands-on adult learners in open-ended tasks.

The present study

The course design under investigation was the second design cycle of a research project that focuses on the design, implementation, and evaluation of design features that are particularly supportive to hands-on adult learners when tackling open-ended tasks in blended learning environments. In particular, based on a formal evaluation of the first design cycle (Boelens & De Wever, 2016, 2017) and an informal evaluation of the course design that combined the perspectives of the students, the teacher, and the researcher (Könings et al., 2005), the teacher and the first author of this study (i.e., researcher) co-constructed and optimized the learning scenario. In this way, the course design under investigation was firmly informed by the teacher's expertise, international literature, and field testing (Könings et al., 2005; McKenney & Reeves, 2012). Situated against the background of our project, at this phase of development, the primary aim of the present study is to test the *design conjectures* of this course design. Secondarily, we examine the plausibility of the *process conjectures* to yield the intended outcomes, in light of the design conjectures, and to determine if they require refinement. The research questions are:

(1) To what extent are the enactment processes achieved, and how do the features of the course design contribute to this?
(2) How do the enactment processes appear to contribute to the learning outcomes?

Method

Context

The present study is part of a research project on the design of blended learning environments for hands-on adult learners. This study is situated in a teacher training program, in which hands-on learners with prior craft knowledge (e.g. bakery, electricity, hairdressing) retrain to become teachers in vocational programs (see also Boelens et al., 2018). These learners have participated in two of the four tracks of secondary education in Belgium: vocational or technical (not general or artistic). The intervention under investigation took place in the course *psycho-pedagogical competence* throughout 16 weeks. This course is part of the second year of a 2,5-year teacher training program. The main course objectives were: students are able to recognize the most common learning and developmental disabilities (i.e., dyslexia, dyscalculia, non-verbal learning disability, attention deficit hyperactivity disorder, autism spectrum disorder), students are able to state the consequences for the learning behavior of pupils, and students are able to develop classroom practices that address the needs of pupils with learning or developmental disabilities based on these insights.

Research design

A case study approach was adopted to understand participants' behavior and perceptions regarding the blended course design in an authentic context (Yin, 2009). To obtain emic (participant) as well as etic (researcher) views, the design of the case study was inspired by ethnographic approaches (Hammersley & Atkinson, 2007). To get to know the context and the learners, the first author actively participated in the teacher training program, which enabled her to follow one cohort of hands-on learners during two consecutive academic years (2015-2016 and 2016-2017). Each year, the researcher conducted participant-observation throughout the duration of one course, and asked questions to students and teachers through formal and informal interviews. In addition, the same researcher worked together closely with the teacher in the present study through two cycles of course design, development, and evaluation and revision (Roschelle & Penuel, 2006), the second of which forms the context of this study. This afforded the researcher a detailed understanding of what happened precisely during this process, as well as the teacher's perspectives on it.

Participants

The participants were both the hands-on learners (n=10) and the teacher who re-designed and implemented the blended course design. The average age of the students was 27 years (SD=6.17, range=21-39), and the female teacher was 43 years old. Participation was voluntary and both teacher and students gave their informed consent, having been made fully aware of the nature and purpose of the research. The response rate was 100%. Appendix A provides additional information about the hands-on learners who participated in this study (i.e., gender, age, diploma, the group in which they collaborated, and each one's pre-course perceptions about open-ended tasks).

Data collection

As is common in ethnographic studies, multiple data sources were used to explain the richness and complexity of the effects of the course design on participants' behavior and perceptions (Cohen, Manion, & Morrison, 2007; Hammersley & Atkinson, 2007). First, two student interviews and two teacher interviews were conducted to get a detailed understanding of individual experiences and interpretations that are of importance to the present study's research questions (Cohen, Manion, & Morrison, 2007). While ethnographers typically employ an unstructured data collection (Hammersley & Atkinson, 2007), the four semi-structured interview protocols used did have an open orientation. That is, the interview protocols were not aligned with the arrows in the conjecture map presented in this study. Instead, the protocols centered on all educational activities during the course design, and open questions were formulated to investigate participants' experiences without steering them, searching for relevant information on all activities and design features. For each course design feature underlying *material/resources* and *activity/task structures*, participants were asked to elaborate on their actions during the learning scenario, and questioned about (1) the perceived effectiveness of the feature, and (2) why this feature did or did not contribute to student learning. The two features underlying *participation/practices* were more implicitly addressed in the interviews (e.g., "Did you miss the explicit guidance of the teacher when gathering knowledge? Why (not)?"). The interview protocols were thus organized around open-ended questions, which allowed the interviewer to ask for more details or to clarify misunderstandings (Cohen et al., 2007). All participants were interviewed twice by the first author of this study, once during and once after the implementation of the learning design. The teacher was also asked to describe her experiences in a diary throughout the course design, and she could make use of her notes during the interview. The duration of a student interview varied between 9 and 23 minutes, with an average duration of 16 minutes (*SD*=4.08), and the teacher interviews lasted respectively 52 and 76 minutes. All interviews were audio-recorded and transcribed. To avoid socially desirable biases, the interviewer ensured that all participants felt comfortable and secure to talk freely during the interview (Cohen et al., 2007), by explaining that there were no right or wrong answers, that all data would be treated confidentially, and that reporting would be anonymized.

Next to teacher and student interviews, all educational activities were observed as they unfolded (Brophy, 2006). In particular, all face-to-face meetings were observed and video-taped by the first author, while also the online activities (e.g., email conversations, forum posts) were observed and logged. During the observations, the researcher took notes to reduce the data (Brophy, 2006).

A last data source is the produced artifacts, including students' individual preparation, the group document, a leaflet of each group's topic, the lesson plan for their presentation, and student presentations. Additionally, the teacher's evaluation of the group document, based on the predefined rubric, was collected.

Table 2 presents an overview of each data source in relation to the enactment processes central to this study. In addition, the student and teacher interviews were used to test the design conjectures as well as the plausibility of the process conjectures.

Table 2

Data sources aimed to measure the enactment processes

	Student interviews	**Teacher interviews**	**Observations**	**Artifacts**
Artifacts that meet high standards		X		X
Summarize key ideas		X		
Negotiation	X	X	X	
Pace aligning actions	X	X		X
Students experience a safe structure	X	X		
Students experience autonomy	X	X		
Students experience a sense of community	X		X	

Data analysis

In order to analyze students' and the teacher's perceptions, the interview transcripts were coded using NVivo 11 (Miles & Huberman, 1994). First, the first author read and reread the interview transcripts to become familiar with the rich information the data entailed. Second, an analysis framework was compiled based on the theoretical framework and research questions. In this analysis framework, participants were listed in columns. The rows listed for each enactment process (1a) how the enactment process was achieved, and (1b) how the enactment process was related to the intended outcome(s) (i.e., process conjecture(s)), and (2) for each course design feature how this influenced the enactment process(es) (i.e., design conjecture(s)). As such, the patterns of response arising across the participants were put together in order to compare issues that each of them has raised (Cohen et al., 2007). Third, this analysis framework was used by the first author to analyze all transcripts. Each transcript was carefully read and reread to identify the relevant data.

Observations were used to provide insight in two enactment processes: negotiation and students experiencing a sense of community. First, direct observations of students' interaction during collaboration were analyzed to investigate whether they went beyond sharing and comparing information and negotiated about their opinions and ideas to construct knowledge. The findings of this analysis are reported in a previous study (Boelens & De Wever, under review).

Second, indications were searched in the field notes that the researcher took during the observations to explore whether students experienced a sense of community.

Artifacts were analyzed to investigate whether students produced artifacts that meet high standards, and whether students aligned their pace with each other. For the former, we analyzed the teacher's evaluation based on the predefined rubric to assess the group document. The rubric consisted of 17 criteria with 4 indicators for each criterion that described the specific expectations. For each criterion, a maximum of three points could be earned. For the latter, we analyzed whether each student met all intermediate deadlines. This was done by comparing the revision history of the documents, examining previous versions of evolving documents, and dates of saving during the course.

Results

For each of the seven enactment processes, we present three classifications to help answer the research questions, together with the data which yielded the classifications. First, a classification is given for the *enactment process* in relation to intentions (i.e. achieved, not achieved, or sometimes achieved), justified by description of what it looked like. Next, the related *process conjecture is* classified (plausible, not plausible, plausible but limited data, or insufficient data to comment on plausibility), in light of results concerning the outcomes. Third, the *design conjectures* are classified (supported, not supported, and supported but with limited data), explained by findings about if and how specific features of the course design influenced the enactment process.

Enactment processes: participant artifacts

Produce artifacts that meet high standards

Producing artifacts that meet high standards was a successfully enacted process. The students first performed an individual task in which they searched for information and generated their own ideas to gather knowledge about the topic. Later on, students collaborated in small groups to create a group document in which they brought together and negotiated about their retrieved information, ideas, and opinions to acquire and process knowledge about their topic. Finally, the groups presented their work to each other. The teacher indicated during the interview that all students were well prepared for the group work. In addition, she stated that the group documents differed from each other, but all met the (minimum required) standards. The differences in quality were, according to the teacher, attributable to the individual responsibility of the groups, as they decided for themselves how deep to process the information in the open-ended task. The teacher was also satisfied with students' presentations. Regarding the group document, the teacher's evaluation based on the rubric showed that students' mean score was 7.4 (out of 10) (*SD*=1.12), with a minimum score of 5.74 and a maximum score of 8.19.

The results suggest that PC-PA-1 is plausible: *Produced artifacts that meet high standards can lead to cognitive outcomes, because students process and master the content*. During the interviews, the majority of the students (n=6) indicated that by means of actively processing the content, they already mastered the content and remembered it better, as opposed to when the teacher or an expert tells everything in class. In addition, three students commented that they valued the fact that all students could learn from each other through the presentations, and Esra argued that she memorizes things better when presented by her fellow students. The teacher also reasoned during the interview that students remember the content better in this way than when an expert tells about the learning and developmental disabilities.

Three features of the course design were intended to support students in producing artifacts that meet high standards: worked examples, process worksheets, and the explicit task of assessing the reliability of the sources. During the student interviews, all students perceived the *worked examples* as helpful for their learning, because these helped them to familiarize themselves with the external expectations related to the open-ended task. Students consulted the worked examples mainly when they felt the need for more clarification. For instance, Olivia reported: "when we were unsure about something while performing the task, we consulted the worked example." These external expectations concerned both content-related expectations, for instance, Eva said: "two questions in the process worksheet were very similar to each other, and by analyzing the worked example, I could know what was actually intended", and editorial requirements, such as lay-out, length, and formulation of the responses. However, Lucas was the only student who was less enthusiastic about the worked examples. On the one hand, he did not recognize the added value of some of the worked examples, because the examples "covered a different topic", while on the other hand, he indicated to have trouble to retrieve documents on the learning platform, and, in this case, he was not inclined to seek and open the worked examples. In line with what the majority of the students reported, the teacher stated during the interview that the worked examples clarified the external expectations and requirements. As such, worked examples support the production of artifacts that meet high standards because students are aware of the task requirements.

Regarding the *process worksheets*, the student interviews revealed that all students preferred this structure above a less structured task. In particular, students stated that the process worksheet contributed to their learning, as the different steps with sub-questions helped them to select main ideas and cluster the information. Charlotte putted it as follows: "About some questions in the task I would not have written anything without the process worksheet. In my opinion, it was good that the process worksheets were there, otherwise the task might be less complete than the teacher would expect.” This view was echoed by Lucas who indicated that the process worksheet prevented the copying and pasting text from an existing source into the task, since the information needed to be carefully selected and clustered. From the interview with the teacher no information related to this part of the design conjecture was obtained. Nonetheless, process worksheets support the production of artifacts that meet high standards because students select and cluster main ideas in a structured way.

Turning now to the explicit task of *assessing the reliability of the sources*, the majority of the students (n=7) reported during the interviews that they took into account the proposed criteria to scan their sources. For instance, both Olivia and Emma said that this helped them to get a better view on the reliability of the source when the retrieved information seemed not very plausible, which helped them to direct their attention to objective or reliable information. Also, the teacher mentioned during the interview that the students generally used reliable sources to build their knowledge. In sum, the task of assessing the reliability of sources supports the production of artifacts that meet high standards because students gather and process reliable information. These results support DC-PA-1 with confidence: *The use of worked examples, and/or process worksheets, and/or the explicit task to assess the reliability of the sources can help students to produce artifacts that meet high standards*.

Summarize key ideas

The enactment process aimed to accurately summarize the key ideas was not achieved. The groups compiled a leaflet based on their group document with the intention to reconsider and recapitulate the content once again. Although the teacher emphasized the added value of the summarization of key ideas during the interview, she was not completely satisfied with its elaboration: "I should have safeguarded the quality of the content of the leaflet better. For instance, sometimes supportive measures (to cope with a learning or developmental disability in a classroom) were given that had little to do with the students' field."

Since this process was not enacted as was intended, we have insufficient data to comment on the plausibility of PC-PA-2: *Summarizing key ideas can lead to cognitive outcomes, because students consolidate their knowledge*. During the interviews, students particularly indicated that the leaflet will assist them to study for the exam, but they did not state that the creation of the leaflet helped them to, for instance, remember the content better. During the teacher interview, the teacher did not elaborate on the outcome of summarizing key ideas, as she was not satisfied with the enactment of this process.

A *pre-structured summarizing task* was intended to support students to accurately summarize the key information. Based on the student interviews, the majority of the students (n=8) perceived this task as an essential element that supported their learning, because this helped them to produce a clear overview of their own topic. As Sophie indicated: "When you are tackling an open-ended task for such a long time, you sometimes lose the overview. And now (by means of the summarizing task) you have an overall picture of the topic." By way of contrast, two students indicated that the summarizing task did not contribute to their learning. For instance, Mia said: "this was redundant, because the same information of the task came back in the leaflet." Moreover, the student interviews revealed information about how the groups approached the creation of the leaflet. On the one hand, Esra reported that they did not spend a lot of time and effort to compile the leaflet, but copied and pasted the earlier retrieved information. On the other hand, Olivia argued: "while we summarized the retrieved information, we reconsidered whether everything was well formulated, or whether we should do something different." During the teacher interview, the teacher concluded that the summarizing task itself does not guarantee the enactment of accurately summarizing key information. These findings provide no support for DC-PA-2: *the pre-structured summarizing task did not contribute to summarizing key ideas because it did not stimulate students to reconsider the content to extract and recapitulate the main points*.

Enactment processes: observable interactions

Negotiation

The process of negotiation during collaboration was sometimes achieved. The small groups first interacted with each other through an online forum, and later during a face-to-face meeting. All students indicated during the interviews that they jointly discussed about the information to compile the final group document. However, Eva indicated that she did most of the work alone because Charlotte was busy doing other things during the face-to-face moment, and Emma noted that she and Sophie often had similar opinions and therefore little negotiation. The teacher provided a similar picture during the interview. She noted that Eva and Charlotte quickly passed the task without negotiation, Sophie and Emma exchanged a lot but negotiated less, while the other two groups seemed to negotiate more about their opinions and ideas. Furthermore, the direct observation of students' interaction during collaboration has illustrated that these two latter groups jointly constructed knowledge in a few occasions, while the interaction in all groups generally was limited to sharing and comparing information and ideas without treating each other's opinion critically (Boelens & De Wever, submitted).

The results suggest that PC-OI-1 is plausible: *Negotiation during collaboration can lead to cognitive outcomes, because students have a more advanced comprehension of the content*. Six students indicated during the interviews that actively processing the content in small groups helped them to master the content better. In addition, the teacher commented during the interview that students processed the content more in depth because they did not share out different parts of the task and worked apart from each other, but were mutually engaged in task solutions.

Four features of the course design were intended to foster negotiation: the pre-structured summarizing task, the individual idea generation task, the explicit task to assess the reliability of sources, and discursive practices that are characterized by collaboration, support, and easy access to each other's expertise. Regarding the *pre-structured summarizing* task, it became clear during the student interviews that three groups jointly composed the leaflet, while in group 3 Eva composed the leaflet on her own. From the groups who jointly composed the leaflet, group 1 commented that it was simply copy-pasting the information, while group 4 stated that this was a last check to jointly discuss about the content of the group task. Thus, the groups approached the summarizing task in a different way, which not always resulted in the process of negotiation. From the teacher interview no information related to this part of the design conjecture was obtained. In sum, the pre-structured summarizing task did not contribute to negotiation because it did not stimulate students to collectively reconsider their work and decide upon the most essential information.

With respect to the *individual idea generation task,* the student interviews showed that all students perceived this approach advantageous for their learning. Olivia stated: "by means of the individual preparation, each student possessed the basic information before we started with the collaborative task". Eight of the students indicated that this was a good starting point to compare and discuss about each other's work. For instance, Sophie said: "it was interesting to compare our individual contributions, and to decide which information to include in the group task." Moreover, Victor reported that "we had to underpin our answers in the individual preparation, and we automatically did this too during the group work." In sum, student interviews revealed that this individual idea generation task helped them to share and compare their work, and served as a starting point for negotiation and jointly constructing knowledge. According to the teacher, this approach ensured that every student felt individual responsible for the group work and was prepared for the discussion, which led to a thoughtful discussion. In addition, she stated that the fact that there was a difference in nuance and new thinking exercises in the group task, ensured that it was not just copy-pasting information, and students were inclined to negotiate with each other. Therefore, the individual idea generation task caused negotiation because students discussed their individual ideas to reach consensus.

Turning to the explicit task to *assess the reliability of the sources*, whereas the majority of the students (n=7) indicated that they scanned their sources on reliability, none of the students stated that they used this as an argument in their discussion during the collaboration. In contrast, the teacher stated in the interview that she noted during the face-to-face conversations that students were very conscious of the reliability of their sources. In short, the task to assess the reliability of sources appeared to cause negotiation by providing an objective argument during collaboration.

The last feature of the course design intended to foster negotiation was the emergence of *discursive practices* that are characterized by collaboration, support, and easy access to each other's expertise. During the interviews with the students, both Victor and Esra argued that because they had a good relationship with the other group members, everyone shared his or her ideas and "you can more easily tell your opinion (Victor)". Moreover, for Lucas it was important to feel that he could rely on the support and expertise of the other group members, what enabled him to discuss and negotiate during collaboration. The interview with the teacher revealed that, in her opinion, the students were willing to invest time in the group work, resulting in qualitative dialogues during collaboration. Thus, discursive practices characterized by collaboration, support, and easy access to each other's expertise did support negotiation because students were willing to share and probe ideas.

To summarize, these findings only partly support DC-OI-1: *an individual idea generation task, and/or a learning environment characterized by collaboration, support and easy access to each other's expertise, can help students to engage in negotiation during collaboration*. In addition, the results appear to support the conjecture that *the explicit task to assess the reliability of the sources can help students to engage in negotiation*, but data are limited. Finally, one part of the conjecture, *a pre-structured summarizing task can help students to engage in negotiation*, was not supported by the results.

Pace aligning actions

Students enacted successfully pace aligning actions during the open-ended task that was divided in eight parts with intermediate deadlines. The student interviews, teacher interview, and direct observations revealed that all students met the deadlines for the individual task, except Charlotte missed the first deadline to accomplish the first two steps of the individual task. Consequently, her group could not proceed with the group work. However, she had quickly caught up with her backlog so that the group did not run into problems to align their actions and meet the predetermined deadlines. Afterwards, all students and groups met all deadlines and everyone stayed in pace with each other.

The results show that PC-OI-2 is plausible: *Pace aligning actions can result in student persistence, by reducing possible obstacles (e.g., bad time planning)*. During the student interviews, Esra said that "the guidance in the time has ensured that she successfully completed the open-ended task". Moreover, Olivia reported that “it would be hard to have the discipline to start with the open-ended task on time”, and Charlotte indicated that she often postpone doing schoolwork, so “it was actually good that we had to submit a part of the project every week or every two weeks, and not all at once.” In addition, the teacher stated during the interview that if it would be difficult for the students to stay in pace with each other and meet the deadlines, then there would be a lot of resistance against the open-ended task, or even worse, students would dropout from the program.

This enactment process was fostered by setting *intermediate deadlines and frequently reminding students* to upcoming deadlines and their learning progress. The student interviews revealed that all students were positive about the planning with intermediate deadlines. Moreover, seven students indicated that this was one of the success factors of the course design. The intermediate deadlines and reminders helped students to feel less insecure about the course planning (n=6) and empowered them to stay engaged during task execution (n=3). For instance, Lucas said: “This was fantastic, you know for sure that you are on track and you did not forget to post anything, it provides a secure feeling.” In addition, Sophie indicated that the motivation would not be that big without the intermediate deadlines, and Eva stated that “in this way, we produced step by step a nice final product”. Regarding the reminders sent by the teacher, the majority of the students (n=7) agreed that this was useful to know that a deadline is approaching or where they are situated in the learning process. However, three students (Emma, Sophie, Mia) did not feel the need for additional supervision provided by the teacher. Emma said: “on the one hand, I felt that this was redundant, however, on the other hand, it can be useful to receive a reminder when you still need to submit a task or to know where you are in the sequence of learning tasks”. The teacher indicated during the interview that it was good for the students to spread the work at a maximum, through intermediate deadlines and reminders. This ensured that it was feasible for them to stay in pace. These results provide support for DC-OI-2: *Intermediate deadlines and reminders prompt students to align their pace with each other because it avoids procrastination and decreases concerns about the planning*.

Enactment processes: participant experiences

Students experience a safe structure

This enactment process was achieved. All students indicated during the interviews that the open-ended task was very clear, had a good and detailed structure, and was well supervised by the teacher. For instance, students stated that the open-ended task was clearly formulated (Emma) and explained (Eva), and expectations were clear (Mia). In the teacher's opinion, students received sufficient supervision and the students experienced the clear structure underlying the open-ended task.

The results indicate that PC-PE-1 is plausible, but with limited data: *If students experience a safe structure this can result in student persistence by reducing insecurity and frustrations, and in positive motivational beliefs by feeling more empowered to tackle challenging tasks*. Concerning the behavioral outcome, the student interviews revealed that the clear structure helped them to gain confidence in how to approach the task, to be sure they did not forget something, and to ensure that they did not start too late with the execution of the tasks (Lucas and Sophie). Also Victor stated that the clear structure and the step-by-step approach ensured that it seemed less work than when you have to do everything at once. He argued that the clear structure helps to prevent himself from panicking, procrastinating, and dropping out. The teacher said during the interview that she felt that the students really needed that structure to keep up and feel confident.

With regard to the motivational outcomes, all students indicated during the interviews that they had no trouble motivating themselves for the open-ended task. Four students (Esra, Olivia, Sophie, Eva) indicated that the clear structure entailed a motivating element. For instance, Olivia stated that it helped her to keep track of her own progress: "The division in the separate steps gave a good feeling. Then you know, after this, we already completed step one and two. Yes, you see the result immediately." As such, the clear structure ensures that students feel empowered to tackle the open-ended task, by ensuring that students are not overwhelmed and stay motivated. No information about this process conjecture was obtained from the teacher interview.

Three features of the course design were intended to ensure that students experienced a safe structure: worked examples, process worksheets, and intermediate deadlines and reminders. Four students reported during the interviews that the *worked examples* were useful to get an overall picture of what you have to do (Sophie, Anna, Eva, Charlotte), and to estimate the workload or how much time the task will cost (Sophie). According to the teacher, worked examples were really useful, because students can look at them whenever they want to get a global view of the open-ended task. Briefly, worked examples enhance students' experience of a safe structure by providing an overall picture of the task.

Turning to the *process worksheets*, all students were satisfied about the structure this brought. Emma indicated that the clear guiding questions ensured that they sufficiently elaborated on the information and reported their sources. Next to this, Esra said: "It was good to see, for instance, that we are already at step three, or we have to take two more steps and then we are done." Emma also indicated that the structure helped her to organize and plan her learning: "I personally prefer a task that is well-structured, because it is easier to arrange the work. For example, for those pieces that I had to make at home, I could easily divide the work, or say that I'm going to stop (after a certain step), and then I can easily start working on it (the next step) next time." From the teacher interview no information about this part of the design conjecture was obtained. In sum, process worksheets enhance students' experience of a safe structure because they scaffold the search and learning process.

Nine of the ten students explicitly argued during the interviews that *the intermediate deadlines and especially the reminders* felt good to know whether they are processing well, or they still have to do a lot of work. Victor said: "in this way, you are sure that you have not lost sight of certain tasks." During the interview, the teacher indicated that she had the impression that by regularly reminding the students (both face-to-face and online) of upcoming deadlines and their learning progress, students had a clear structure to fall back on. To summarize, intermediate deadlines and reminders enhance students' experience of a safe structure by providing feasible steps to tackle open-ended tasks. These findings provide support for DC-PE-1: *the use of worked examples, and/or process worksheets, and/or the provision of intermediate deadlines and reminders help students to experience a safe structure.*

Students experience autonomy

The enactment process of students experiencing autonomy was achieved. The students had to perform specific learning activities in the online environment, and had an active role in acquiring and processing knowledge individually and in group. In eight of the interviews with students, it became clear that they experienced autonomy. They talked about the independence they perceived (Eva, Olivia), the opportunity to choose when to engage in the learning activities (Anna) and to work at your own pace (Sophie), the chance to showcase your own ideas and to present your work to each other (Victor, Emma, Lucas, Eva, Esra), and the opportunity to discuss in group about the content (Victor, Emma). Also the teacher indicated during the interview that students experienced autonomy.

The findings indicate that PC-PE-2 is plausible: *If students experience autonomy this can result in positive motivational beliefs, because student choice fosters intrinsic motivation*. During the student interviews, seven students emphasized that this autonomy was satisfying and contributed to their motivation. In particular, students used the following concepts: fun, motivating, interesting, and pleasant. Also, the teacher indicated during the interview that, in her opinion, students experienced the autonomy positively, and find the open-ended task more pleasant and motivating than receiving all information during a lesson from a teacher or expert.

The fact that students were expected to take charge of their own learning processes was aimed to ensure that students experienced autonomy. The interviews with the students showed that the autonomy students experienced was mainly due to the fact that they had both the individual and the group responsibility to gather and process information, and to present their work to each other (Victor, Emma, Eva, Esra, Sophie, Lucas, Olivia). In addition, two students (Anna, Sophie) explicitly referred to the increased responsibility to take charge of their own learning process during the online part of the open-ended task. During the teacher interview, the teacher also stated that students took charge of their own learning processes by having an active role in processing the content and presenting their own work to their peers, and by having the responsibility (in the online part) to decide when, where, and how long to engage in the learning activities. These results provide support for DC-PE-2: *students who are expected to take charge of their own learning will experience autonomy by taking responsibility and an active role in the process.*

Students experience a sense of community

This enactment process was achieved, namely: students experienced a sense of community. The students were expected to work together in small groups, to present their work to each other, and to support each other during the presentations, while the teacher created a safe and supportive learning climate. During the majority of the student interviews (n=9), it became clear that students experienced a sense of community. For instance, students said that they perceived the collaboration as successful because all group members showed commitment and dedication (n=4), open communication was possible (n=3), everyone said honestly what they thought (Anna), and Esra emphasized that all group members had a close relationship with each other. In addition, Charlotte stated that everyone was very open to each other during the presentations. Regarding the role of the teacher, Charlotte and Eva indicated that the teacher showed her commitment to bring the open-ended task to a good end. Regarding the observations, indications were found that students were experiencing a sense of a community. For instance, students asked each other for advice, they communicated immediately when there were ambiguities, and also for general issues regarding the teacher training program they consulted (the advice of) the teacher.

The results indicated that PC-PE-3 is plausible, but with limited data: *If students experience a sense of community, this can lead to student persistence because it reduces obstacles and insecurities, and increases positive motivational beliefs because students want to contribute to the group*. With respect to student persistence, Sophie said that she would not be sure if she would be up for completing the open-ended task when she would have to do it by herself, as it is a lot of work. In addition, Lucas reported that this process gave him a more secure feeling by sharing each other's ideas about how the task should be tackled. From the interview with the teacher no information related to this part of the process conjecture could be obtained. Regarding motivational outcomes, students argued that both the supportive relationship with the teacher (Eva and Charlotte), and the feeling of belonging and open communication in a small group (Sophie, Mia, Lucas) had a positive effect on their motivation. The teacher also stated that immediate positive feedback and quickly responding to students' questions motivates them.

Two features of the course design were intended to establish a sense of community: intermediate deadlines and reminders, and discursive practices characterized by collaboration, support, and easy access to each other's expertise. First, two students explicitly indicated that because the teacher provided *intermediate deadlines and sent frequent reminders* and offered them help, this gave them the feeling that the teacher was committed and willing to help when they experienced problems. Charlotte putted it as follows: "in this way the teacher showed that she wants to stimulate and support us [to bring the open-ended task to a good end]." From the interview with the teacher no information related to this part of the design conjecture was obtained. Still, intermediate deadlines and reminders prompt the experience of a sense of community because of the commitment and the proximity of the teacher.

Second, regarding the *discursive practices* that are characterized by collaboration, support, and easy access to each other's expertise, all students indicated that the teacher was easily accessible, both online and face-to-face, and quickly replied on their questions, concerns and frustrations. Five students explicitly stated that the teacher was one of the success factors of the course design, as she provided them with a lot of support and motivation (Anna, Sophie, Olivia, Eva, Charlotte). Next to this, students also supported each other during the project. For instance, all students experienced the collaboration as very positively. They reported that there was open communication, everyone showed commitment, they complemented each other and learned a lot from each other. The teacher also argued during the interview that when students encounter difficulties, she tries to respond to that. In this way, she shows that she wants to help them and pays attention to their problems, which is important to create a sense of community and belonging. In short, discursive practices characterized by collaboration, support, and easy access to each other's expertise support the experience of a sense of community because the students and teacher showed commitment toward each other. These results support DC-PE-3: *the provision of intermediate deadlines and reminders and/or an environment that is characterized by collaboration, support, and easy access to each other's expertise will prompt students to experience a sense of community.*

Discussion

Reflections on the findings

This study focused on salient design features of a teacher education course that aimed to foster hands-on adults' learning in an open-ended task. Conjecture mapping was used to explore if, how, and why the course design supported learners. While we also examined the influence of the enactment processes on the learning outcomes (i.e., process conjectures), the primary aim of the present study was to investigate the influence of course design features on specific enactment processes (i.e., design conjectures). Findings revealed that almost all of the proposed design features stimulated the hands-on adult learners toward the intended enactment processes, which in turn yielded the intended outcomes. However, the results show that there is still room for improvement with regard to two of the seven enactment processes, namely: summarizing key ideas and negotiating. Therefore, Figure 3 shows the retrospective theory of action (i.e., conjecture map), which underpinned this course and could be used as a basis for shaping other courses aiming to support hands-on adult learners in open-ended tasks. Thereafter, we discuss possible explanations for the findings related to summarizing key ideas and negotiating, and propose alternative design features, accordingly.

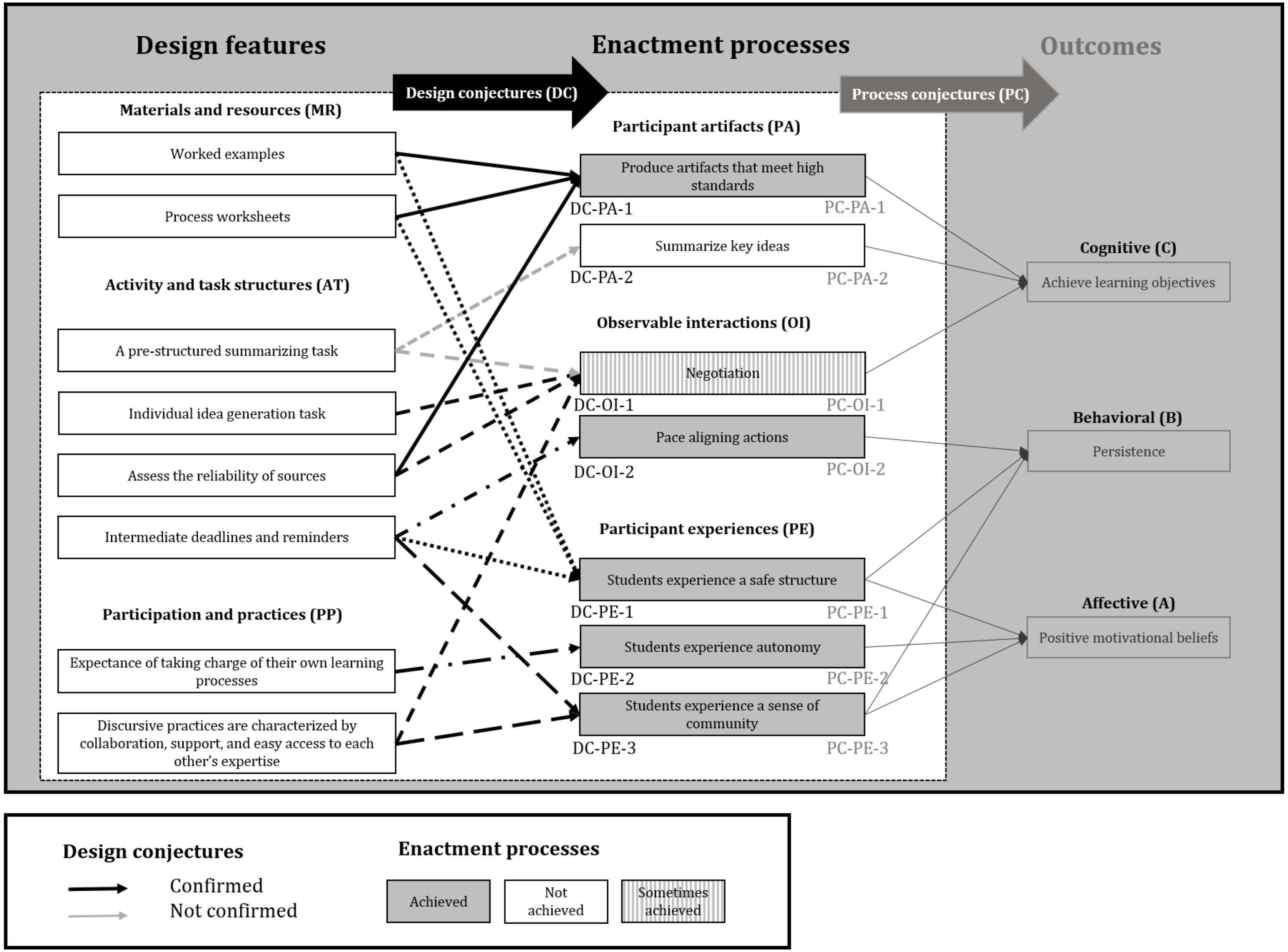

Figure 3. Retrospective conjecture map for supporting hands-on adult learners in open-ended tasks.

Regarding the first enactment process, *summarizing key ideas*, the results show that the pre-structured summarizing task, in its current form, did not help students to accurately summarize key ideas. In particular, the findings suggest that in some cases, this task leads to copying and pasting information instead of actively processing the content to extract the main points. A possible explanation for this might be that students do not possess effective or sufficient summarizing strategies to accurately summarize the key ideas based on a written text. As such, students might benefit from deliberate practice exercises in which they, for instance, learn to use their own words to summarize the material (King, 1992), which might empower them to more accurately summarize key ideas. In addition, the teacher can provide additional support in the way of formative feedback, after which students can modify their summarizing strategies to elicit more accurately summarizing key ideas (Hattie & Timperley, 2007). Another explanation could be that the summarizing task was too structured and did not provoke active processing of the content. In future attempts to improve the design, the summarizing task could be adapted, for instance by prompting students to focus more on keywords to ensure that students actively reconsider and process the content (De Bruin, Thiede, Camp, & Redford, 2011), which can lead to more accurate summarizing practices.

Concerning the second enactment process, *negotiation during collaboration*, the results show that this process was partly achieved, meaning that student interaction was mainly restricted to sharing and comparing information, and only two (of the four) groups sometimes jointly negotiated about the content and co-constructed knowledge. Findings related to the pre-structured summarizing task indicate that students did not always perform this task collectively, nor did the task encourage students to discuss the relevance of the information. While the other three design features (i.e., individual idea generation task, the explicit task to assess the reliability of sources, and a learning environment characterized by collaboration, support, and easy access to each other's expertise) seemed to prompt interaction, they were insufficient to obtain the intended process of negotiation. This result is likely related to the fact that most of the students were unfamiliar with (computer-supported) collaborative learning in an open-ended task. Consequently, more practice moments may be needed to familiarize them with how they can effectively negotiate with each other to construct knowledge. Another possible explanation might be that students are not naturally inclined to construct new knowledge through negotiation, as shown by previous research in vocational education settings (Hämäläinen & Oksanen, 2012). Further work should be undertaken to support the likely tradeoff decision concerning additional practice moments (if they are sufficient to elicit more negotiation), versus adding a different design feature (which might entail the risk of over-scripting collaborative learning; see Dillenbourg, 2002).

Reflections on the methods and suggestions for further research

A first consideration of the methods used concerns the fact that, although the whole class group participated in the study, the sample size was rather small (i.e., 1 teacher and 10 students). However, the aim of this study was in the first place to gather in-depth information about how the design worked in this type of context, and small-scale case studies are certainly appealing to achieve this aim (Boddy, 2016). Further, the triangulation of data sources made it possible to substantiate the results by examining the conjectures from different angles. In most of the cases, the data sources strengthened each other's findings, while in some cases this unraveled (small) variations between students. Now that this study has provided detailed explanations for why the design features contributed to the enactment processes and the learning outcomes, future research should explore their adaptation tolerance by studying variation in results with different manifestations of these features.

In line with this, the second consideration is that the course design articulated in this study was enacted in only one context: a teacher training program for adult learners who are transitioning from vocational and technical careers to secondary level teachers in their field. Although a single case can be highly instructive (Boddy, 2016), this also places limitations on the generalizability of the designed solution. However, the generalizability should not be restricted to only this particular context. On the contrary, we believe that these insights are valuable for teachers, educational designers, and researchers working in similar contexts where (formal) adult education programs involve students with prior craft knowledge. Replicating this type of study in different contexts could shed light on the generalizability of the design framework, while also providing deeper insight into how other student groups cope with open-ended tasks that embody these design features.

Third, by presenting a case of conjecture mapping, we were able to articulate a theoretically-grounded course design and to empirically test the design conjectures. In this way, this study addressed the current need for educational design research to more clearly articulate its core (pedagogical) building blocks (Graham, Henrie, & Gibbons, 2014). However, further research is needed to move beyond plausibility testing of the process conjectures, to ascertain if the outcomes were indeed caused by the proposed mechanisms, or if other mechanisms better explain them.

Implications

The results of the current study hold several important implications. First, the validated design framework can assist those who aim to implement open-ended tasks in similar contexts. The framework could be used as a whole, but also provides multiple entry points and thus flexibility of use. For instance, course developers may choose to design open-ended tasks that include the specific supportive features that are shown here to be important for students' learning processes and outcomes (e.g., individual idea generation task, worked examples). Alternatively, developers may consider the specific enactment processes they aim to engender, and implement the corresponding design features (e.g., stimulate a safe structure by implementing intermediate deadlines and reminders, and process worksheets). Additionally, developers focusing on student outcomes (e.g., student persistence), could direct their attention toward specific enactment processes related to that outcome (e.g., pace aligning actions) and the associated design features (e.g., intermediate deadlines and reminders). As such, the design solution proposed in this study provides guidelines, though no certainties, to develop open-ended tasks aiming to achieve cognitive, behavioral, and affective outcomes in an educational context with hands-on adult learners (Edelson, 2006).

Second, the findings contribute to theoretical understanding by identifying and testing how and why the salient course design features are related to enactment processes and learning outcomes in a context with hands-on adult learners (Whetten, 1989). To a great extent, the findings in this study confirm the existing theories described in relation to the individual design and process conjectures (e.g., worked examples support the production of artifacts that meet high standards). Moreover, by bringing these ideas together in a carefully aligned set of design and process conjectures, not only a design framework, but also as a theoretical framework is created. This theoretical framework describes, explains, and predicts how key design features can support the execution of and learning from open-ended tasks by hands-on adult learners. Given that the intended outcomes were obtained in this study (i.e., students achieved the learning objectives, persisted during the course, and held positive motivational beliefs), it seems warranted to recommend for further elaboration and critique of this framework (e.g., by others).

Third, what stands out in our results is that all learning outcomes were achieved by a combination of multiple enactment processes. This also implies that most of the design features indirectly lead to more than one learning outcome. For instance, while in the literature worked examples and process worksheets mainly stem from a cognitivist perspective and benefits are described in terms of performance and efficiency (see e.g., Nadolski, Kirschner, & van Merriënboer, 2006), these design features also seem to contribute to behavioral and affective outcomes through the process *students experience a safe structure*. This finding highlights the added value of looking at an educational intervention from a holistic perspective, and acknowledging the most important components (i.e., design features, mediating processes) to see how they function and interact with each other to achieve the learning outcomes (Sandoval, 2014).

Closing remarks

This study was undertaken to better understand how to support lifelong learners as they respond to the changing needs of their working lives. Specifically, a course was designed for professionals transitioning from vocational careers into the role of teaching at vocational education institutions. For this kind of hands-on adult learners, the study showed for instance that worked examples not only engender the production of artifacts that meet high standards because they clarify the task requirements, but also foster a safe structure by providing an overall picture of the task. In addition, deadlines and reminders support pace aligning actions by taking away concerns about the planning, as well as stimulate a sense of community because of the commitment and proximity of the teacher. Finally, discursive practices characterized by collaboration, support, and easy access to each other's expertise foster negotiation by feeling a shared responsibility to share ideas, and contributing to a sense of community as the students and the teacher show commitment toward each other. These individual findings, as well as the revised conjecture map as a whole, constitute modest yet important steps towards understanding how to support hands-on adult learners to accomplish open-ended tasks.

References

Arbaugh, J. B., Cleveland-Innes, M., Diaz, S. R., Garrison, D. R., Ice, P., Richardson, J. C., & Swan, K. P. (2008). Developing a community of inquiry instrument: Testing a measure of the Community of Inquiry framework using a multi-institutional sample. *The Internet and Higher Education*, *11*(3–4), 133–136. http://doi.org/10.1016/j.iheduc.2008.06.003

Arvaja, M., & Pöysä-Tarhonen, J. (2013). Tracing discursive processes of shared knowledge construction in a technology-enhanced higher education setting. *Interactive Learning Environments*, *21*(4), 321–337. http://doi.org/10.1080/10494820.2011.559171

Ausburn, L. J. (2004). Course design elements most valued by adult learners in blended online education environments: an American perspective. *Educational Media International*, *41*(4), 327–337. http://doi.org/10.1080/0952398042000314820

Barnard, L., Lan, W. Y., To, Y. M., Paton, V. O., & Lai, S.-L. (2009). Measuring self-regulation in online and blended learning environments. *The Internet and Higher Education*, *12*(1), 1–6. http://doi.org/10.1016/j.iheduc.2008.10.005

Barron, B. (2003). When Smart Groups Fail. *Journal of the Learning Sciences*, *12*(3), 307–359. http://doi.org/10.1207/S15327809JLS1203_1

Biemans, H., Mariën, H., Fleur, E., Tobi, H., Nieuwenhuis, L., & Runhaar, P. (2016). Students' Learning Performance and Transitions in Different Learning Pathways to Higher Vocational Education. *Vocations and Learning*, *9*(3), 315–332. http://doi.org/10.1007/s12186-016-9155-6

Boddy, C. R. (2016). Sample size for qualitative research. *Qualitative Market Research: An International Journal*, *19*(4), 426–432. http://doi.org/10.1108/QMR-06-2016-0053

Boelens, R., & De Wever, B. (2016). Scripting inquiry-based learning in a teacher education program for adult learners: how learners use and experience structuring interventions. In *EARLI SIG 20 Conference: inquiry in and outside classrooms*.

Boelens, R., & De Wever, B. (2017). Collaborating on a shared document: Vocational and technical students' approaches and experiences. In *Proceedings of the European Conference on e-Learning, ECEL* (pp. 56–64). Retrieved from https://search.proquest.com/docview/1968935738?accountid=11077%0A

Boelens, R., & De Wever, B. (under review). How can teachers support student interaction during computer-supported collaborative learning? An exploratory case study in a higher education setting for hands-on learners. *British Journal of Educational Technology*.

Boelens, R., Voet, M., & De Wever, B. (2018). The design of blended learning in response to student diversity in higher education: Instructors' views and use of differentiated instruction in blended learning. *Computers & Education*, *120*, 197–212. http://doi.org/10.1016/j.compedu.2018.02.009

Brand-Gruwel, S., & Gerjets, P. (2008). Instructional support for enhancing students' information problem solving ability. *Computers in Human Behavior*, *24*(3), 615–622. http://doi.org/10.1016/j.chb.2007.01.020

Brand-Gruwel, S., Wopereis, I., & Vermetten, Y. (2005). Information problem solving by experts and novices: analysis of a complex cognitive skill. *Computers in Human Behavior*, *21*(3), 487–508. http://doi.org/10.1016/j.chb.2004.10.005

Brand-Gruwel, S., Wopereis, I., & Walraven, A. (2009). A descriptive model of information problem solving while using internet. *Computers & Education*, *53*(4), 1207–1217. http://doi.org/10.1016/j.compedu.2009.06.004

Bråten, I., & Samuelstuen, M. S. (2007). Measuring strategic processing: comparing task-specific self-reports to traces. *Metacognition and Learning*, *2*(1), 1–20. http://doi.org/10.1007/s11409-007-9004-y

Brophy, J. (2006). Observational research on generic aspects of classroom teaching. In P. A. Alexander & P. H. Winne (Eds.), *Handbook of Educational Psychology* (pp. 755–780). Mahwah, New Jersey, London: Lawrence Erlbaum Associates.

Cohen, L., Manion, L., & Morrison, K. (2007). *Research Methods in Education* (Sixth edit). London and New York: Routledge. http://doi.org/10.1111/j.1467-8527.2007.00388_4.x

Damşa, C. I., & Nerland, M. (2016). Student Learning Through Participation in Inquiry Activities: Two Case Studies in Teacher and Computer Engineering Education. *Vocations and Learning*, *9*(3), 275–294. http://doi.org/10.1007/s12186-016-9152-9

De Bruin, A. B. H., Thiede, K. W., Camp, G., & Redford, J. (2011). Generating keywords improves metacomprehension and self-regulation in elementary and middle school children. *Journal of Experimental Child Psychology*, *109*(3), 294–310. http://doi.org/10.1016/j.jecp.2011.02.005

Deci, E. L., & Ryan, R. M. (2000). The "what" and "why" of goal pursuits: Human needs and the self-determination of behavior. *Psychological Inquiry*, *11*(4), 227–268. http://doi.org/10.1207/S15327965PLI1104_01

DiCicco-Bloom, B., & Crabtree, B. F. (2006). The qualitative research interview. *Medical Education*, *40*, 314–321. http://doi.org/10.1111/j.1365-2929.2006.02418.x

Dillenbourg, P. (2002). Over-scripting CSCL: The risks of blending collaborative learning with instructional design. In P. A. Kirschner (Ed.), *Three worlds of CSCL: Can we support CSCL?* (pp. 61–91). Heerlen: Open Universiteit Nederland. http://doi.org/10.1007/s11165-004-8795-y

Drexler, W. (2010). The networked student model for construction of personal learning environments: Balancing teacher control and student autonomy. *Australasian Journal of Educational Technology*, *26*(3), 369–385. http://doi.org/10.14742/ajet.1081

Dubeau, A., Plante, I., & Frenay, M. (2017). Achievement Profiles of Students in High School Vocational Training Programs. *Vocations and Learning*, *10*(1), 101–120. http://doi.org/10.1007/s12186-016-9163-6

Edelson, D. C. (2006). What we learn when we engage in design: Implications for assessing design research. In J. van den Akker, K. Gravemeijer, S. McKenney, & N. Nieveen (Eds.), *Educational Design Research* (pp. 156–165). Routledge.

Edelson, D. C., Gordin, D. N., & Pea, R. D. (1999). Addressing the Challenges of Inquiry-Based Learning Through Technology and Curriculum Design. *Journal of the Learning Sciences*, *8*(3–4), 391–450. http://doi.org/10.1080/10508406.1999.9672075

Garrison, D. R. (2007). Online community of inquiry review: social, cognitive, and teaching presence issues. *Journal of Asynchronous Learning Networks*, *11*(1), 61–72. http://doi.org/10.1128/JB.05513-11

Graham, C. R., Henrie, C. R., & Gibbons, A. S. (2014). Developing models and theory for blended learning research. In A. G. Picciano, C. D. Dziuban, & C. R. Graham (Eds.), *Blended learning: Research perspectives* (Volume 2, pp. 13–33). New York: Routledge.

Green, A., Preston, J., & Sabates, R. (2003). Education, Equality and Social Cohesion: A distributional approach. *Compare: A Journal of Comparative and International Education*, *33*(4), 453–470. http://doi.org/10.1080/0305792032000127757

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research*, *17*(4), 397–431. http://doi.org/10.2190/7MQV-X9UJ-C7Q3-NRAG

Haerens, L., Vansteenkiste, M., Aelterman, N., & Van den Berghe, L. (2016). Toward a Systematic Study of the Dark Side of Student Motivation: Antecedents and Consequences of Teachers' Controlling Behaviors. In W. C. Liu, J. C. K. Wang, & R. M. Ryan (Eds.), *Building Autonomous Learners: Perspectives from research and practice* (pp. 59–81). Singapore: Springer. http://doi.org/10.1007/978-981-287-630-0_4

Hämäläinen, R., De Wever, B., Malin, A., & Cincinnato, S. (2015). Education and working life: VET adults' problem-solving skills in technology-rich environments. *Computers and Education*, *88*, 38–47. http://doi.org/10.1016/j.compedu.2015.04.013

Hämäläinen, R., & Oksanen, K. (2012). Challenge of supporting vocational learning: Empowering collaboration in a scripted 3D game - How does teachers' real-time orchestration make a difference? *Computers and Education*, *59*(2), 281–293. http://doi.org/10.1016/j.compedu.2012.01.002

Hammersley, M., & Atkinson, P. (2007). *Ethnography: Principles in Practice.* (Third edit, Vol. 15). Routledge. http://doi.org/10.2307/2070079

Hannafin, M. J., & Hill, J. R. (2007). Resource-based Learning. In *Handbook of research on educational communications and technology* (3rd ed., pp. 525–536). http://doi.org/10.1007/978-1-4419-1428-6_449

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of Educational Research*, *77*, 81–112. http://doi.org/10.3102/003465430298487

Hmelo-Silver, C. E., Duncan, R. G., & Chinn, C. A. (2007). Scaffolding and Achievement in Problem-Based and Inquiry Learning: A Response to Kirschner, Sweller, and Clark (2006). *Educational Psychologist*, *42*(2), 99–107. http://doi.org/10.1080/00461520701263368

Jossberger, H., Brand-Gruwel, S., van de Wiel, M. W. J., & Boshuizen, H. (2018). Learning in workplace simulations in vocational education: A student perspective. *Vocations and Learning*, *11*(2), 179–204. http://doi.org/10.1007/s12186-017-9186-7

Jossberger, H., Brand-Gruwel, S., van de Wiel, M. W. J., & Boshuizen, H. P. A. (2015). Teachers' Perceptions of Teaching in Workplace Simulations in Vocational Education. *Vocations and Learning*, *8*(3), 287–318. http://doi.org/10.1007/s12186-015-9137-0

Karoğlu, A. K., Kiraz, E., & Ozden, M. Y. (2014). Good Practice Principles in an Undergraduate Blended Course Design. *Education and Science*, *39*(173), 249–263.

King, A. (1992). Comparison of Self-Questioning, Summarizing, and Notetaking-Review as Strategies for Learning From Lectures. *American Educational Research Journal*, *29*(2), 303–323. http://doi.org/10.3102/00028312029002303

Könings, K. D., Brand-Gruwel, S., & van Merriënboer, J. J. G. (2005). Towards more powerful learning environments through combining the perspectives of designers, teachers, and students. *The British Journal of Educational Psychology*, *75*, 645–660. http://doi.org/10.1348/000709905X43616

Liu, X., Bonk, C. J., Magjuka, R. J., Lee, S., & Su, B. (2005). Exploring four dimensions of online instructor roles: A program level case study. *Journal of Asynchronous Learning Networks*, *9*, 29–48. Retrieved from http://sloanconsortium.org/jaln/v9n4/exploring-four-dimensions-online-instructor-roles-program-level-case-study

McDonald, P. L. (2014). Variation in adult learners' experiences of blended learning in higher education. In A. G. Picciano, C. D. Dziuban, & C. R. Graham (Eds.), *Blended learning: Research perspectives* (Volume 2, pp. 215–234). New York: Routledge.

McKenney, S., & Reeves, T. C. (2012). *Conducting educational design research*. London: Routledge.

McKenney, S., & Reeves, T. C. (2018). *Conducting educational design research* (2nd ed.). London: Routledge.

McLaren, B. M., Van Gog, T., Ganoe, C., Karabinos, M., & Yaron, D. (2016). The efficiency of worked examples compared to erroneous examples, tutored problem solving, and problem solving in computer-based learning environments. *Computers in Human Behavior*, *55*, 87–99. http://doi.org/10.1016/j.chb.2015.08.038

Merrill, M. D. (2007). First principles of instruction: a synthesis. In R. A. Reiser & J. V. Dempsey (Eds.), *Trends and issues in instructional design and technology, 2nd edition* (Vol. 2, pp. 62–71). Upper Saddle River, NJ: Merrill/Prentice-Hall. http://doi.org/10.1007/BF02504986

Miles, M. B., & Huberman, M. A. (1994). *Qualitative data analysis: An expanded sourcebook (2nd ed.)*. London: Sage.

Nadolski, R. J., Kirschner, P. A., & Merriënboer, J. J. G. (2005). Optimizing the number of steps in learning tasks for complex skills. *British Journal of Educational Psychology*, *75*(2), 223–237. http://doi.org/10.1348/000709904X22403

Nadolski, R. J., Kirschner, P. a., Merriënboer, J. J. G., & Hummel, H. G. K. (2001). A model for optimizing step size of learning tasks in competency-based multimedia practicals. *Educational Technology Research and Development*, *49*(3), 87–101. http://doi.org/10.1007/BF02504917

Nadolski, R. J., Kirschner, P. A., & van Merriënboer, J. J. G. (2006). Process support in learning tasks for acquiring complex cognitive skills in the domain of law. *Learning and Instruction*, *16*(3), 266–278. http://doi.org/10.1016/j.learninstruc.2006.03.004

Nerland, M. (2012). Changes in Knowledge Cultures and Research on Student Learning. Berkeley. Retrieved from https://cshe.berkeley.edu/publications/changes-knowledge-cultures-and-research-student-learning

Nicol, D. J., & Macfarlane-Dick, D. (2006). Formative assessment and self-regulated learning: a model and seven principles of good feedback practice. *Studies in Higher Education*, *31*(2), 199–218. http://doi.org/10.1080/03075070600572090

Onrubia, J., & Engel, A. (2009). Strategies for collaborative writing and phases of knowledge construction in CSCL environments. *Computers & Education*, *53*(4), 1256–1265. http://doi.org/10.1016/j.compedu.2009.06.008

Onrubia, J., & Engel, A. (2012). The role of teacher assistance on the effects of a macro-script in collaborative writing tasks. *International Journal of Computer-Supported Collaborative Learning*, *7*(1), 161–186. http://doi.org/10.1007/s11412-011-9125-9

Pintrich, P. R. (2000). The Role of Goal Orientation in Self-Regulated Learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of Self-Regulation* (pp. 451–502). San Diego, CA: Academic Press. http://doi.org/10.1016/B978-012109890-2/50043-3

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, *16*(4), 385–407. http://doi.org/10.1007/s10648-004-0006-x

Reeves, T. (2011). Can Educational Research Be Both Rigorous and Relevant ? The Pendulum Swings : Rigor and Relevance. *Educational Designer*, *1*(4), 1–24. Retrieved from http://www.educationaldesigner.org/ed/volume1/issue4/article13/pdf/ed_1_4_reeves_11.pdf%5Cnhttp://www.educationaldesigner.org/ed/volume1/issue4/article13/

Roediger, H. L., & Butler, A. C. (2011). The critical role of retrieval practice in long-term retention. *Trends in Cognitive Sciences*, *15*(1), 20–27. http://doi.org/10.1016/j.tics.2010.09.003

Roschelle, J., & Penuel, W. (2006). Co-design of innovations with teachers: Definition and dynamics. In *The 7Th International Conference on the Learning Sciences* (pp. 606–612). Retrieved from http://dl.acm.org/citation.cfm?id=1150122

Rovai, A. P. (2003). In search of higher persistence rates in distance education online programs. *Internet and Higher Education*, *6*(1), 1–16. http://doi.org/10.1016/S1096-7516(02)00158-6

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and Extrinsic Motivations: Classic Definitions and New Directions. *Contemporary Educational Psychology*, *25*(1), 54–67. http://doi.org/10.1006/ceps.1999.1020

Salomon, G. (1996). Studying novel learning environments as patterns of change. In S. Vosniadou, E. De Corte, R. Glaser, & H. Mandl (Eds.), *International perspectives on the design of technology supported learning environments* (pp. 363–377). Mahwah, NJ: Erlbaum.

Sandoval, W. A. (2014). Conjecture Mapping: An Approach to Systematic Educational Design Research. *Journal of the Learning Sciences*, *23*(1), 18–36. http://doi.org/10.1080/10508406.2013.778204

Smith, P. J. (2000). Flexible delivery and apprentice training: Preferences, problems and challenges. *Journal of Vocational Education and Training*, *52*(3), 483–503. http://doi.org/10.1080/13636820000200132

Smith, P. J. (2001). Learners and their workplaces: towards a strategic model of flexible delivery of training in the workplace. *Journal of Vocational Education & Training*, *53*(4), 609–628. http://doi.org/10.1080/13636820100200180

Tomlinson, C. A., & Imbeau, M. B. (2013). Differentiated instruction. In B. J. Irby, G. Brown, R. Lara-Alecia, & S. Jackson (Eds.), *The Handbook of Educational Theories* (pp. 1097–1117). Charlotte, North Carolina: Information Age Publishing, Inc.

Tsai, C.-W., & Shen, P.-D. (2009). Applying web-enabled self-regulated learning and problem-based learning with initiation to involve low-achieving students in learning. *Computers in Human Behavior*, *25*(6), 1189–1194. http://doi.org/10.1016/j.chb.2009.05.013

van Gog, T., Paas, F., & van Merriënboer, J. J. G. (2008). Effects of studying sequences of process-oriented and product-oriented worked examples on troubleshooting transfer efficiency. *Learning and Instruction*, *18*(3), 211–222. http://doi.org/10.1016/j.learninstruc.2007.03.003

van Merriënboer, J. J. G. (2013). Perspectives on problem solving and instruction. *Computers & Education*, *64*, 153–160. http://doi.org/10.1016/j.compedu.2012.11.025

van Merriënboer, J. J. G., Clark, R. E., & de Croock, M. B. M. (2002). Blueprints for Complex Learning : The 4C/ID-Model. *Educational Technology Research and Development*, *50*(2), 39–64.

van Merriënboer, J. J. G., & Kirschner, P. A. (2013). *Ten steps to complex learning: a systematic approach to four-component instructional design* (Second edi). New York: Routledge.

Van Oers, B., & Hännikäinen, M. (2001). Some thoughts about togetherness: an introduction. *International Journal of Early Years Education*, *9*(2), 101–108. http://doi.org/10.1080/713670686

Vandewaetere, M., Manhaeve, D., Aertgeerts, B., Clarebout, G., Van Merriënboer, J. J. G., & Roex, A. (2014). 4C/ID in medical education: How to design an educational program based on whole-task learning: AMEE Guide No. 93. *Medical Teacher*, (93), 1–17. http://doi.org/10.3109/0142159X.2014.928407

Weinstein, C. E., Husman, J., & Dierking, D. R. (2000). Self-regulation interventions with a focus on learning strategies. In *Handbook of self-regulation* (pp. 727–747). http://doi.org/10.1016/B978-012109890-2/50051-2

Whetten, D. A. (1989). What Constitutes a Theoretical Contribution? *Academy of Management Review*, *14*(4), 490–495. http://doi.org/10.2307/258554

Workman, J. J., & Stenard, R. A. (1996). Student support services for distance learners. *DEOSNEWS*, *6*(3), 1–11. Retrieved from http://www.ed.psu.edu/acsde/deos/deosnews/deosnews6_3.asp

Yin, R. K. (2009). *Case Study Research: Design and Methods* (Fourth edi). London and Singapore: Sage.

Appendix A. Students' background information

Name	Sex	Age	Diploma VS=Vocational secondary TS=Technical secondary	Group	Pre-course perceptions related to open-ended tasks
Victor	M	33	VS-painter and decoration	1	He has no experience with open-ended task, and feels therefore somewhat uncomfortable and insecure. He prefers to have weekly face-to-face meetings in which the teacher explains the content, because he thinks it would be difficult to combine online learning activities with his family responsibilities
Lucas	M	39	No senior secondary education degree, completed junior VS (to age 16)-welding metals	1	He indicates that starting, and also completing, every learning activity that requires a certain amount of autonomy is a big step for him
Esra	F	23	VS-hairdressing	1	She has positive perceptions about open-ended tasks in which students have to take an active role in their learning process, as this can be done at her own pace, and she can choose how much time to spend on the task
Emma	F	24	VS-beauty care	2	She reports that she can easily deal with tasks that require a certain amount of responsibility and autonomy
Sophie	F	21	VS-hairdressing	2	She prefers to get everything explained by the teacher
Eva	F	23	TS+adult education-beauty care and pedicure	3	She indicates that she has no problems with processing certain content independently
Charlotte	F	21	TS-beauty care	3	She indicates that the risk of a high amount of autonomy is that she might procrastinate her school work. She needs someone who motivates and encourages her to perform the tasks in time

Anna	F	26	TS-beauty care	4	She reports positive experiences with open-ended tasks: when you have to search for information by yourself, you understand the content better
Mia	F	34	TS-beauty care	4	She reports rather negative experiences: “if I have to work independently, I easily put these things aside. It is hard to combine with my job”
Olivia	F	27	VS-hairdressing	4	She has no experience with open-ended task. Although she is afraid to fail because she has a lot of work and tasks in the program, she thinks she can handle it

Note. F = female; M = male

For further information on these studies please contact the corresponding authors:

The Effect of Cues for Reflection on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

The Effect of Cues for Calibration on Learners' Self-Regulated Learning through Changes in Learners' Learning Behaviour and Outcomes

Stijn Van Laer, KU Leuven – Centre for Instructional Psychology and Technology

stijn.vanlaer@ppw.kuleuven.be

How can teachers support student interaction during computer-supported collaborative learning? An exploratory case study in a higher education setting for hands-on learners

Conjecture mapping to support hands-on adult learners in open-ended tasks

Ruth Boelens, Ghent University – Department of Educational Studies

ruth.boelens@ugent.be