Collaborative Work Practices for Management Education: Using Collaboration Engineering to Design a Reusable and Scalable Collaborative Learning Instructional Design

Sarah Oeste-Reiß
University of Kassel, Germany
oeste-reiss@uni-kassel.de

Matthias Söllner
University of Kassel, Germany
soellner@uni-kassel.de

Jan Marco Leimeister
University of St. Gallen, Switzerland
& University of Kassel, Germany
janmarco.leimeister@unisg.ch

Abstract

Pandemics like COVID-19 highlight the needs and pitfalls of inclusive and equitable education in a digital society. IT-based instructional designs are needed to increase learners' expertise, and to develop higher-order thinking skills. Instructional designs for collaborative learning (CL) seem to be a promising solution. However, they are mostly suitable for face-to-face and not for distance teaching. The core problem that impedes their reusability and scalability is a 'collaboration problem' for which collaboration engineering (CE) provides guidance. Therefore, we deploy a design science research study and contribute to CL and CE literature. We develop requirements and provide the design of an IT-based collaborative work practice fostering CL. We provide empirical evidence with an online experiment in a large-scale lecture with undergraduate business information students. This reveals that groups of learners who followed our CL experience achieve higher levels of expertise than those who followed a traditional ad hoc CL experience.

Keywords: collaborative learning, instructional design, collaboration engineering, collaborative work practice, design science research

1. Introduction

The United Nations define quality education as sustainable development goal and call for inclusive and equitable education opportunities for all (OECD, 2021). A key toward this is inherent in increasing learners' expertise and training job-related higher-order thinking skills like problem-solving, communication and cooperation (OECD, 2016, 2020). One might argue, that predominantly developing countries face such demands. But the COVID-19 pandemic has taught us otherwise. The pandemic showed what happens when traditional learning approaches in the classroom are not possible - i.e., demand for digital learning revealed infrastructural and pedagogical pitfalls and pent-up demands for our supposedly digital society. Massive Open Online Courses (MOOCs) are, despite of their weaknesses (e.g., dropout-rates, frontal teaching, less interaction), on the rise and change the educational landscape (Manca & Meluzzi, 2020). In this light, reusable and scalable teaching learning approaches that foster higher-level-learning (HLL) experiences are needed more than ever. HLL refers to the upper levels of Bloom's taxonomy of educational objectives (e.g. evaluate information, construct, critique, and defend positions, or reason beyond available information to create novel knowledge) (Krathwohl, 2002). However, due to few interaction opportunities and structures, large class sizes or digital settings often do not foster HLL in a useful manner. Thus, when developing contemporary digital learning approaches, interaction opportunities should be considered. Thus, novel instructional designs should emphasis on expanding learners' expertise by training higher-order thinking skills in an interactive, reusable and scalable manner. Not taking this into account would bear the risk to cut people from education in times of an digital and pandemic prone world (OECD, 2021).

Collaborative learning (CL) provides a solution toward this challenge. CL is an instructional design that grounds on constructivist learning theory (Arbaugh, 2010; Damon, 1984; Jones, 2014; Moll, 2013; Topping, 2005). Characteristics are ad hoc collaboration among learners and facilitation guidance from a lecturer. From a collaboration point of view it is mentionable, that group effectiveness tends to decline as group size increases beyond six participants (Ingham et al., 1974). Therefore, CL is still considered as being less reusable and scalable. The core problem of CL that impedes reusability and scalability can be summarized as a *'recurring collaboration problem'*. For example, field experiences show that most individuals do not have an intuitive grasp of how to collaborate effectively. Left to themselves, most groups tend to evolve inefficient and ineffective work practices (Briggs et al., 2013). A solution toward this problem could be, that lecturers develop structured collaborative work practice and take

URI: https://hdl.handle.net/10125/102631
978-0-9981331-6-4

the facilitator role. However, this would require sophisticated facilitation expertise by lecturers and scalability won't be achieved. Another solution could be, that lecturers in their facilitator role make themselves superfluous. This would call for instructional designs that use digital technologies, package facilitation-expertise in the design, foster collaboration among learners and empower learners to execute the instructional design without the ongoing support from lecturers (OECD, 2021). The body of Collaboration Engineering (CE) provides *"an approach to designing collaborative work practices for high-value recurring tasks, and deploying those designs for practitioners to execute for themselves without ongoing support from professional facilitators"* (Briggs et al., 2006). Since the core of the illustrated pedagogical problem is a "collaboration problem", CE provides promising design guidance to develop reusable and scalable instructional designs that have the potential to increase learners' expertise and foster HLL. Instructional designs in the form of collaborative work practices (CWPs) with packaged facilitation expertise constitute a promising solution. To achieve empirical evidence of the effects CE can unfold in the educational domain, we address the following *research question*: How can principles of CE be used to develop CL instructional designs that increase learners' expertise and train higher-order thinking skills in large classes without the ongoing support from a lecturer? On this basis we derive two design goals (DG): *DG 1:* Create a reusable and scalable CWP for an instructional design that supports learners to collaboratively increase expertise and higher-order thinking skills in large scale lectures. *DG 2:* Package collaboration expertise in the instructional design so that practitioners (i.e., learners) can deploy and execute it without training in tools or techniques.

2. Design Science Research Approach

We use a design science research (DSR) approach and use the three cycle view to structure our paper (Hevner, 2007) (Fig. 1). We started a relevance cycle by outlining the real-world and research problem (sec. 1). Then, we started a rigor cycle to position the outlined problem in the bodies of CE and CL literature (sec. 3).

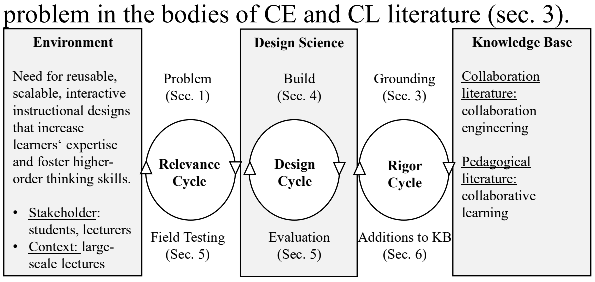

Figure 1: Design science cycles

With these insights, we conducted a design cycle and derived generalizable requirements (sec. 4.1), developed an instructional design in the form of a conceptual design of a CWP and an exemplar instance in the form of a process support application (PSA) (sec. 4.2). We took the PSA back to the field and developed an experimental design to evaluate the PSA in a large-scale lecture with undergraduate business information students (sec. 5). To complete the design and relevance cycle, we focus on empirical evidence toward our design goals, refer to the experimental procedure (sec. 5.1), constructs (sec. 5.2) and testable propositions (sec. 5.3) used and report and discuss the achieved results (sec. 5.4). To complete the rigor cycle, we further discuss contributions, limitations, and future research implications and make additions to the bodies of CE and CL literature (sec. 6).

3. Theoretical Background

3.1 Collaboration Engineering

CE is an approach to design CWPs for high value-recurring tasks. High-value tasks are those that create substantial value or reduce risk of loss of substantial value (G.-J. de Vreede & Briggs, 2019). Such engineered CWPs package facilitation expertise in such a way that they can be deployed to practitioners (i.e., non-collaboration experts like learners) to execute for themselves without or less training in tools and techniques (G.-J. de Vreede & Briggs, 2019). A CWP is a series of reusable collaborative activities performed by multiple teammates to achieve a group goal (Winkler et al., 2019). Groups that execute such engineered work practices can outperform groups left to their own designed procedures (d. G.-J. Vreede et al., 2009). The CE approach is tailored to two unique roles – Collaboration Engineers and Practitioners. *Collaboration Engineers* are familiar with the CE approach and have sophisticated collaboration skills. They have capabilities to design a CWP on a recurring basis. *Practitioners* are people that are skilled in their domain and familiar with the team's task. They have no collaboration expertise. Practitioners can facilitate or participate in an engineered work practice (G.-J. de Vreede & Briggs, 2019). The so-called iterative *Collaboration Process Design Approach (CoPDA)* supports Collaboration Engineers with design guidance. It describes design activities and design choices that a Collaboration Engineer needs to consider when developing a CWP. The CoPDA consists of five steps – (1) task diagnosis; (2) task decomposition; (3) thinkLet choice; (4) agenda building; (5) validation (Kolfschoten & Vreede, 2009). To transfer engineered CWP designs to practitioners, Collaboration Engineers create so-called Facilitation Process Models (FPM). A FPM is a visual abstract overview of the whole process flow (Briggs et al., 2013). In this light, instructional designs that increase learners' expertise and foster higher-order thinking skills, are a suitable application context of CE.

3.2. Collaborative Learning

CL builds on constructivist learning theory. Based on that, learners acquire knowledge and extent their knowledge base through interactions with their environments (e.g. learning content, learners, lecturers) (Moll, 2013; Topping, 2005). This reveals three types of interactions (Moore, 1989; Schrum & Berge, 1997) (Thurmond & Wambach, 2004): *Learner-lecturer* interactions can stimulate cognitive learning mechanisms (e.g., learners test and expand their knowledge through clarification requests). This can trigger the juxtaposition of related knowledge frames in working memory. *Learner-content* interactions occur by e.g., reading text, listening to audio, or watching video. However, there are fewer opportunities for feedback compared to learner-lecturer interaction. *Learner-learner* interactions occur among learners explaining concepts or debating positions to one another. Learners usually have different levels of domain knowledge. The less advanced learners ask clarification questions and benefit from the answers of the advanced learners. In turn, the advanced learners required to provide easy understandable answers and explicate their tacit knowledge which in turn challenges them to reason beyond their available knowledge (Smith et al., 2009; van Dijk et al., 2020). Moreover, such social involvement can also increase learner motivation. A further value of CL is inherent in improving skills such as communication, cooperation, and problem-solving. Thus, we conclude that integrating learner-learner and learner-content interactions in a CL instructional design that is reusable and scalable, has a promising potential. It can create HLL experiences that increase learners' expertise and foster higher-order thinking skills. Problem-solving, communication and cooperation skills can be classified as higher-order thinking skills. Considering the revised Bloom's Taxonomy of learning objectives, higher-order thinking skills like problem-solving, communication and cooperation address the upper levels of the taxonomy (i.e., learning objectives: apply, analyze, evaluate and create) (Krathwohl, 2002).

4. Collaborative Work Practice for a Collaborative Learning Instructional Design

4.1. Generalizable Requirements

The generalizable requirements (GR) for our solution of a CWP for a CL instructional design were derived from the class of unsolved problem (i.e., instructional designs that increase learners' expertise and foster higher-order thinking skills) (sec. 1, 2): Universities and lecturers have limited resources to procure novel digital learning infrastructure. Tools for specialized teaching purposes and courses are hardly existent and thus, bespoke, or custom solutions are expensive. Therefore, <u>GR1. Leverage Available Resources:</u> *It should be possible to create an instance of an instructional design for large classes using infrastructure that are already available at universities.* As group size increases beyond six people, the difficulty of group work tends to rise and the effectiveness of groups tends to decrease (Ingham et al., 1974). It is necessary to prevent students from such distraction and to free their cognitive resources to cope with CL tasks. Therefore, <u>GR 2. Influence Group Behavior by Group Size:</u> *It should be possible to subdivide a large class into parallel breakout groups of 6 or fewer learners to minimize the emergence of dysfunctional group behaviors that could interfere with learning.* Lecturers in large classes experience high demands on their attention (Allais, 2014). They may decline to adopt novel digital instructional designs that require from them a steep learning curve for the implementation and deployment in the classroom. Therefore, <u>GR3. Ensure Transferability:</u> *It should be possible that a lecturer is able to implement the solution in the classroom and deploy it to learners with little or no training.* Most people do not have an intuitive grasp on how to design and conduct effective collaboration (Briggs et al., 2013; G.-J. de Vreede & Briggs, 2019). Therefore: <u>GR4. Provide Prescribed Procedures:</u> *The solution should prescribe procedures that a learner can follow without training in tools or techniques.* Learners increase their expertise (i.e. domain knowledge) when they chunk multiple knowledge schemas into one bigger schema (Sweller et al., 2011). For that purpose, they evaluate the quality of information, reason from first principles to a new position, judge the merits of proposed solutions (Krathwohl, 2002). Therefore: <u>GR5. Foster Schema Building:</u> *The solution should challenge learners with tasks that require them to synthesize multiple knowledge schemas into one more-complex knowledge schema.* Newly chunked knowledge schemas often incorporate incomplete understandings (Sweller et al., 2011). Learners can verify the validity of a new schema by using it to attempt a task. Working memory, however, fades within seconds unless it is refreshed, and new schema in long-term memory fade if they are not reinforced. Thus, it is useful for learners to use acquired new higher-level knowledge as quickly as possible and get an assessment of its quality. Therefore: <u>GR6. Provide Rapid Feedback:</u> *The solution should provide opportunities that learners quickly receive feedback on* *their newly acquired domain knowledge.* Learners usually arrive in a class with differing levels of domain knowledge. Collaboration research shows that groups comprising people with different levels of domain knowledge can achieve greater gains in productivity than homogeneous groups (Ries et al., 2013). Learners can benefit from asking and answering clarification questions; formulating, advancing, and critiquing positions of others and defending their own positions (Smith et al., 2009; van Dijk et al., 2020). To create knowledge on the upper levels of Bloom's taxonomy, learners must have mastered basic knowledge. Therefore, *GR7. Generate Shared Knowledge Base: The solution should foster the give-and-take to create shared understanding. It should give opportunities for learners to compare, challenge and reinforce one another's understandings. GR8. Heterogeneous Learning Groups: The solution should assure that each breakout group has a mix of less advanced and advanced learners.* Achieving knowledge on the upper levels of Bloom's taxonomy places a high intrinsic cognitive load on less advanced learners. Such learners perceive the learning task as more difficult than advanced learners as they need to acquire basic knowledge and build required schema in long term memory (Sweller et al., 2011). Inventing ad hoc collaboration adds an additional extraneous cognitive load (e.g., shape the collaborative procedure; maintain goal congruence). Therefore, *GR9. Minimize Extraneous Cognitive Load: The solution should minimize cognitive load for actions not directly related to learning, to maximize cognitive resources (i.e., germane cognitive load) available for learning.* Learning experiences targeting to advanced learners may be too difficult for novice learners, while those for novice learners may give advanced learners no opportunity to learn. Therefore, *GR9. Ensure Reciprocity: The solution should not be too difficult for less-advanced learners to understand yet should not be too easy that it wastes the time of advanced learners.*

4.2. Collaborative Work Practice: Design and Characteristics of the PSA

A CWP describes a collaboration process in terms of a group goal, work product, and a series of collaborative activities. Adhering to the CoPDA (Kolfschoten & Vreede, 2009), we started with '*step 1: task diagnosis*' and defined a group goal. It refers to a desired state or outcome and motivates individuals for action (Briggs et al., 2008). The *'group goal'* of our use case is *"Learners increase expertise and achieve higher-order thinking skills (abilities for problem-solving, communication, and cooperation) by collaboratively creating a novel solution for a real-world problem from a case study in two hours".* Instrumental for attaining a group goal are the type and design of a *'collaboration task'*. The task points out the expectations of the work product that the group should create. For our use case, we designed a collaboration task in the form of a case study with subtasks (Fig. 2). Taking the 'collaboration task' into account, the *'work product'* that the learners collaboratively create per subtask is *"a correct and well-structured solution for the real-world problem of the case. On five slides learners document the solution and visualize domain knowledge concepts from multiple sources that form new, integrated, and meaningful domain knowledge".*

Case description: "You are a trainee in a small trading company, which sells its goods via stationary trading and additionally wants to set up an online shop. The company´s innovation management wants to address the increasing digitalization and gives you the following tasks."

- Subtask 1: In your team gather examples how the digitalization influences the retail trade, its data models and business processes. Illustrate the model-based task solving on the example of a payment system introduction. Explain the process from the as-is till the target state as well the process from the target till the as-is state.
- Subtask 2: Develop in a team a reference model for the online payment procedure in a small trading company. Follow sector-specific purchase procedures of the well-known online shops. Explain with the help of your model different construction techniques, which are used to design reference models.
- Subtask 3: Explain the possible applications of a customer relationship management (CRM) system in a company. Concentrate on how a CRM system supports the user-, benefit- and usage-orientation. Refer to the online as well the stationary trading and give examples, where CRM systems connect online trading with the stationary trading. Explain on this example the relation between the CRM and ERP system.
- Subtask 4: Explain the ERP implementation in a trading company and describe the benefit of ERP systems within the SCM. Explain the benefit of ERP systems for the operative, middle, and top management.

Figure 2: Collaboration task

We further followed the CoPDA and completed the *'step 2: task decomposition'*, *'step 3: thinkLet choice'*, *'step 4: agenda building'* to develop a conceptual design of our CWP. Since the design choices are not the focus of the study presented in the paper, we subsequently describe the final *'collaborative procedure'*. We refer to characteristics of our developed group decomposition and our developed collaborative process flow (Fig. 3) that illustrates the collaborative activities. Since our design goal 1 is to create a reusable and scalable CWP in large scale lectures, the CWP uses a hybrid approach and group decomposition (i.e., a plenary group with distributed IT-supported subgroups and breakout groups). *Plenary group:* This is the total number of participants (i.e., all learners who participate in the large class setting). *Breakout groups (max. 6 per group):* Learners from the plenary group are divided into breakout groups. Each group has an equal number of less advanced and advanced learners. *Subgroups:* Breakout groups are assigned to larger subgroups. The number of breakout groups per subgroup depends on the number of subtasks. The *'Facilitation Process Model (FPM)'* represents the skeleton of five core collaborative activities that characterize the CWP.

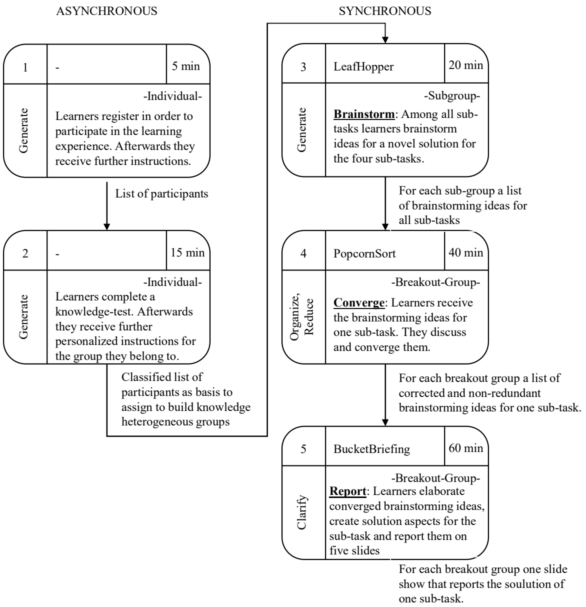

Figure 3: Facilitation process model

The FPM elements show in the upper left, the number of the current activity; in the upper right, the duration of the current activity; in the upper middle, the chosen thinkLet (i.e., a named, scripted collaborative activity that gives rise to a known pattern of collaboration among people working toward a goal (G.-J. de Vreede & Briggs, 2019)); on the left, the pattern of collaboration that will be achieved with this activity; in the center a brief activity summary. The work product of each activity is shown below each FPM element. Activities 1 and 2 represent asynchronous individual learning activities. These pre-collaboration steps are mandatory to communicate the expectations of the 'CL experience' to the learners (e.g., learning units to be acquired; date of the distributed synchronous collaboration session; technical requirements for the technical devices to join the session; walkthrough video on how to use the collaboration space). In addition, they create the conditions to assign learners into knowledge heterogeneous breakout groups. With the start of the synchronous collaborative activities of a distributed synchronous collaboration session (activities 3-5), learners receive step-by-step instructions and the learning case. To complete these activities, learners work in sub-groups and breakout groups to create a solution for the collaboration task. When the learners come together in the lecturer hall, the lecturer randomly chooses these solutions and starts a plenary discussion.

Having this conceptual design of the CWP in mind, an expository instance is required to deploy it in a large class lecture. This is the physical representation of our CWP. For that purpose, we developed a so-called *'Process Support Application (PSA)'*. We used and configured following IT-tools: To communicate expectations and instructions to learners and run a knowledge-test, we used an existing course in the learning management system (LMS) "Moodle" with LMS capabilities for registration, quiz, and separated groups. For the distributed synchronous collaboration session, we configured a collaborative working space with shared writing pages. For that purpose, we prepared shared writing pages with GoogleDocs and GoogleSlides for the various subgroups and breakout groups. To give learners access to the shared writing pages, we copied the URLs into the before configured separated groups in the LMS. This allowed us to assign each learner to a separated subgroup and breakout group and provide individualized instructions and URLs to execute the activities. The core technical requirement for learners to join the collaboration session was to bring their own notebook and ensure a stable internet access.

5. Evaluation

This section provides empirical evidence whether our PSA can foster learners' expertise compared to a traditional CL instructional design by which learners need to invent ad hoc collaboration *(Design Goal 1)*. It also provides empirical evidence whether our PSA can be executed by practitioners (i.e., learners) without an intensive training in tools or techniques *(Design Goal 2)*. Thereto, we conducted an online experiment in a large-scale lecture with business information students.

5.1. Characteristics of the Experimental Design

Experimental Procedure: Our online experiment is characterized by a five-step procedure (Fig. 4): (1) The lecturer announced in the lecturer hall and in Moodle the opportunity to join in a *'CL experience'*. Learners can achieve up to four bonus points for the final exam. (2) Learners used Moodle to register for the 'CL experience', received access to a knowledge-test and then were guided to a pre-survey on Limesurvey. (3) Registered learners received a walkthrough video to become familiar with the expectations of the 'CL experience' and capabilities of the collaboration space. (4) Based on the knowledge-test performance, we randomly assigned learners to knowledge-heterogeneous breakout groups with an assigned subtask. We assigned these breakout groups randomly to one of two samples – i.e., experimental sample vs. control sample.

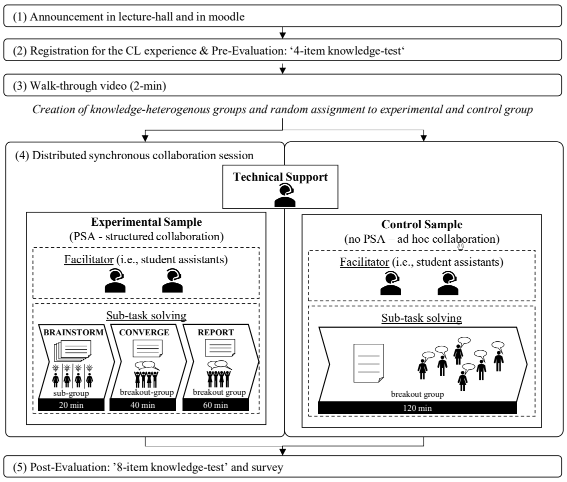

Figure 4: Experimental procedure

We used a spreadsheet to randomly assign three top learners and three bottom learners to each of 16 breakout groups (treatment: N= 48; control: N=56). The control breakout groups had a seventh member because the number of participants was not evenly divisible by 16. This had the potential to skew performance measures in favor of the control breakout groups. Research shows that, collaboration technology that permits simultaneous input by all participants, can increase group productivity in groups up to 30 (Gallupe et al., 1992). Since all were in the control sample, this will make it harder for us to show the value of our treatment sample. To verify that the stratification process produced subject pools with approximately equal levels of ability, we compared the distributions of pre-test scores of both samples with a Mann Whitney U test. Results show no statistically significant difference in the distribution of pre-test scores by experimental sample assignment (U=1076.5, p=0.171). This indicates that both samples started with no bias with respect to lower-level basic knowledge. To control for potential differences in task difficulty in both samples, each subtask was per sample assigned to two breakout groups. A post-hoc Bonferroni test revealed no statistically significant differences in task difficulty for subtasks 1, 2, and 4. Subtask 3 was more difficult than subtask 1. That difference, however, was balanced across treatments. At the date of the distributed synchronous collaboration session, learners of the experimental sample followed our developed PSA with packaged facilitation expertise (Fig. 3). In contrast, learners of the control sample received a subtask of the collaboration task and had the opportunity to invent ad hoc collaboration to develop a solution. Both samples were observed by one person that provided technical support and each sample by two facilitators. Facilitators were student assistants that used our internal agenda with brief facilitation instructions (e.g., reminder for the remaining time; stop editing rights in the documents). Each facilitator observed 4 breakout groups. The use of facilitators served more as backup for the case that the 'CL-experience' could collapse in its first deployment. (5) After the collaboration session, learners were guided to a knowledge-test and post-survey on Limesurvey.

Differences of the experimental and control sample: Learners of the *treatment sample* followed a 'CL experience' in the form of our PSA. This incorporated our engineered CWP with packaged facilitation expertise. The step-by-step instructions on the shared writing pages of each collaborative activity focused on how to collaborate with each other as well as outlined the expectations for a good solution. Learners of the *control group* followed a digital, but traditional 'CL experience'. Full instructions and expectations for a good solution were provided at the beginning. The instructional design allowed learners to develop ad hoc collaboration to create a solution for a subtask. The learners of the control sample received a similar, but less structured collaboration space – i.e., shared writing pages in one googleDocs and one GoogleSlides. The writing pages contained the learning task (i.e., case description and one subtask) but no further instructions on how to organize their collaboration.

Background and context of the learners: Research object was a large-scale lecture at a German university with 150 undergraduate business information students. We conducted the 'CL experience' for one time in semester in the course of a selected learning unit (i.e., domain knowledge). 104 voluntary students registered for our online experiment. Students who completed the whole experimental procedure (Fig. 4) received up to 4 bonus points. In total, 101 students (31 males, 70 females, aged 19-39 years [Mean =23 years; SD = 3,2]) completed the whole experiment.

5.2. Measures

During a pre- / post-evaluation, we collected data from two audiences (Table 1): *(1) Data from learners*: The collected learner data provided insights toward changes in learners' expertise regarding their basic factual knowledge (i.e., lower-level-learning (LLL)) and satisfaction. To measure knowledge changes, we used a pre/post knowledge-test with true/false questions. The pre knowledge-test (i.e., LLL_KT_1357) comprised four questions each with a reference to one of the four subtasks from the case. The post knowledge-test (i.e., LLL_KT_ 12345678) comprised eight questions. It contained the same four questions from the pre-test, and four new questions each with a reference to one of the four subtasks from the case. To measure learners' satisfaction, we used established constructs (i.e., tool difficulty [TOOLDIF]; satisfaction with process (SP); satisfaction with outcome (SO) (Briggs et al., 2013); efficiency; productivity (Kolfschoten & Santanen, 2007), perceived team performance (Benalian, 201X)) on a 7-point Likert scale (1- disagree / 7- agree).

Table 1: Overview of measures

	pre-evaluation	post-evaluation
performance measures (individual - knowledge test by learner)	4-item knowledge test • pre-and-post-questions (pre_post_LLL_1357)	8-item knowledge test • 8-LLL-questions (LLL_12345678) • post-only-questions (post_LLL_2468)
performance measures (group- lecturer assessment)	-	• LLL_level_of_correctness • HLL_level-of_sophistication
satisfaction measures (survey by learner)	-	TOOLDIF, SP, SO, Effic., Effect., Prod., TP

(2) *Data from independent lectures (domain experts):* The collected data provided insights toward the LLL and higher-level learning (HHL) performance between the two samples. We asked five treatment-blind, independent raters (i.e., lecturers) to assess the collaborative work products (i.e., 5-slide solutions of the breakout groups). All raters had teaching experience in the knowledge area of information systems. Thus, they were familiar with the topic of the use case. Therefore, we can assume that they provide a high-quality assessment. For that assessment, we developed an evaluation scale (Table 2). The interrater reliability measured by Cronbach's Alpha was 0.85.

Table 2: Group performance measures

Vari-able	Items
LLL performance: **' Level of correctness '**	Raters evaluated the 'level of correctness' of the domain knowledge represented by the amount of a correct solution aspects in the students' work products (i.e., 5-slides). To evaluate the work product, raters used a 7-point Liktert scale: *1) = The group did not submit a solution;* *2) = number of correct aspects is 0%;* *3) = 20%; 4) = 40%; 5) = 60%; 6) = 80%; 7) = 100%.*
HHL Performance **' Level of sophistication'**	Raters evaluated the 'level of sophistication' of the domain knowledge represented by the students' work products (i.e., 5-slides). This refers to connections among differentiated characteristics of domain knowledge. Raters used a 7-point Likert scale: *1) The group did not submit a solution; 2) No visual representation of relationships among concepts/ copied or long unfocused textbook phrases; (...); 7) Complete visualization of relationships among concepts/ clear and concise phrases.*

5.3. Hypothesis and Testable Propositions

To test, whether our PSA achieves our design goals, we formulated a hypothesis (H). To analyze and prove the effects beyond the hypothesized HLL effect, we derived exploratory propositions (P).

- ***H1***: *Breakout-groups that execute the PSA will score higher on HLL performance ('level of sophistication') than breakout-groups that invent ad-hoc collaboration in a CL experience.*

Overall effects on LLL: To prove whether CL leads to an increase in learners' LLL performance, we expect LLL performance increases in both samples (i.e., experimental / control). For that purpose, we analyze differences in the LLL measured by a comparison of pre and post knowledge-test sores.

- ***P1***: *Learners score better on 'pre-and-post_ LLL 1357' questions in the post-test than in the pre-test.*

Differences in LLL by treatment: Even though the main objective is to measure HLL performance effects, we assume that the PSA increases learners' LLL performance. Compared to the control sample, learners in the experimental sample do not need to invent ad hoc collaboration and have the chance to focus their attention on the learning content.

- ***P2***: *Learners in the experimental sample score better on the '8-LLL-12345678' questions in the post-test than learners in the control sample.*
- ***P3***: *Learners in the experimental sample score better on the 'pre_post_1357' questions in the post-test than learners in the control sample.*
- ***P4***: *Learners in the experimental sample score better on the 'post_2468' questions in the post-test than learners in the control sample.*
- ***P5***: *Breakout-groups in the experimental sample achieve better 'LLL_level_of_correctness' scores than breakout-groups in the control sample.*

Differences in satisfaction by treatment: Next to designing a PSA that increases expertise (design goal 1), it was important that the PSA can be executed without training in tools or techniques (design goal 2). We assume, that comparable satisfaction scores in both samples are a suitable indicator. Traditional CL experiences call for learning experiences that are not restricted and allow for ad hoc collaboration.

- ***P6***: *Learners in the experimental sample are more satisfied (SP, SO, TOOLDIF, Effic., Effec., Prod., TP) with the CL experience than those in the control sample.*

5.4. Results and Discussion

To ensure that learners in both samples started with no bias in LLL knowledge, we compared the means of the pre-test questions ('pre_post_LLL_1357'). A t-test revealed that there is no significant difference by treatment (t=-1.166 I p=0.171).

Design Goal 1: Table 3 reports the means of HLL performance (i.e., lecturer assessment of breakout-groups' work product: 5 slides) per sample, t-statistics, and p-value. Regarding ***H1***, the means in the experimental sample (M = 5.425) are significantly better than those the control sample (M = 3.825). The large effect size accounts for 52.4% of the variance in structural assessment scores. We attribute this result to the execution of our PSA. It seems that following a well-structured collaboration process supports learners with better HLL conditions. An explanation may be inherent in Cognitive Load Theory (Sweller et al., 2011) and the use of Collaboration Engineering (G.-J. de Vreede & Briggs, 2019). This also supports previous literature that considers CWPs for CL experiences and the transfer and documentation of knowledge in small classes (Oeste-Reiß et al., 2016; Oeste-Reiß et al., 2017). Learners in the experimental sample received step-by-step instructions and thus, had less extraneous cognitive load. They were able to rather focus on the learning content than to invent ad hoc collaboration. Those learners gained more exposure to the learning content, and thus achieved faster discussions and contributions for representing a high-level of sophistication among domain knowledge. To produce a higher-level knowledge solution, they must go beyond lower-level knowledge (i.e., evaluate information, critique, and defend positions, and to reason beyond information). During the 'brainstorming' and 'convergence' activities (Fig. 3), learners created a common understanding of domain knowledge. They elaborated this knowledge in the 'reporting' activity. Contributions from the advanced learners may helped the less advanced learners to assemble domain knowledge. Therefore, the results suggest that we achieved our design goal 1.

Table 3: HLL performance by sample

	Experimental Sample		Control Sample		t(df)	p-value (1-t.)
	N	Mean	N	Mean		
HLL_ Sophistication	8	5.425	8	3.825	t(14)= 3.933	p = 0.001**
Statistical significance *p<0.05, **p<0.01 / 7-point Likert scale						

Table 4 reports the means of learners' LLL performance regardless of the sample during the pre- and during the post-test. Regarding ***P1***, learners performed significantly better (p=0.036*) on the 'pre_post_LLL_1357' questions in the post-test (M=0.5644) than in the pre-test (M=0.5099). It is to mention that advanced learners already had high LLL performance scores in the pre-test. Therefore, we attribute performance increases to the less advanced learners. The results may also confirm the positive value of CL experiences in general. Interestingly, this may indicate that LLL can be achieved in large classes with a traditional, but IT-based 'CL experience' (i.e., invent ad hoc collaboration; receive no process restrictions).

Table 4: Overall effects on LLL

	Pre-test LLL_1357		Post-test LLL_1357		t(df)	p-value (1-t.)
	N	Mean	N	Mean		
P1	100	0.5099	100	0.5644	t(100) =1.817	p = 0.036*
Statistical significance *p<0.05, **p<0.01						

Table 5 reports the differences in LLL performance by experimental and control sample, t-statistics, and p-value. Regarding ***P2***, learners in the experimental sample (M=0.6622) performed significantly better on all post-test 'LLL_12345678' questions than those in the control sample (M=0.5841). This further confirms the results of our H1 and the value that our PSA can unfold. The structured collaboration may help learners to juxtapose knowledge concepts and relationships in working memory. In contrast, students in the control sample may have been distracted as they had to invent ad hoc collaboration. Regarding ***P3***, in the post-test, there was no significant difference on the four 'pre_post_LLL_1357' questions between samples. This might be attributed to a priming as learners had seen the questions already in the pre-test. Regarding ***P4***, learners in the experimental sample (M=0.7181) performed significantly better on the 'post_LLL_2468' questions than those in the control sample (M=0.6346). An explanation for this effect could be similar to that of P2. Regarding **P5**, the lecturer assessment revealed no significant difference of the breakout-groups' work product in terms of its 'LLL_level_of-correctness'. An explanation could be attributed to the positive CL effects in terms of LLL (similar to P1).

Table 5: LLL performance by sample

	Experimental Sample		Control Sample		t(df)	p-value (1-t.)
	N	Mean	N	Mean		
P2 (LLL_ 12345678)	47	0.6622	52	0.5841	t(97) = 2.113	p = 0.0185*
P3 (LLL_ 1357)	47	0.6064	52	0.5337	t(97) = 1.539	p = 0.0635
P4 (LLL_ 2468)	47	0.7181	52	0.6346	t(97) = 1.803	p = 0.0375*
P5 (LLL_ Correc.)	8	5.625	8	5.525	t(14) = -0.362	p = 0.3615
Statistical significance *p<0.05, **p<0.01						

Design Goal 2: Table 6 reports the satisfaction means by experimental and control sample. Regarding ***P7***, the analysis reveals that there is no significant difference in learners' satisfaction ('SP, SO, TOOLDIF, Effic., Effec., Prod.') by sample. The results show that learners in both samples were motivated and had the chance to take desired actions for creating the group deliverable. This indicates that learners experience the 'CL experience' equally satisfied. For several reasons, we interpret the non-significant results as positive. First, they may indicate that learners in the experimental sample feel, despite of process restrictions of the PSA (i.e., step-by-step instructions, pre-structured writing pages), comparably free in their CL experience like learners in the control sample. Second, in traditional CL instructional designs a training in tools and techniques is typically not necessary. Thus, we can assume that our PSA packages facilitation expertise and can be executed without a training in tools and techniques. Third, all learners were able to execute the 'CL experience' (i.e., distributed synchronous collaboration session) without problems. Each of the 16 breakout groups submitted a joint work product (i.e., 5 slides). In addition, learners in the experimental sample (M=5.2319) perceived the 'team performance' (TP) better than those in the control sample (M=4.7630). This indicates that learners not only performed better objectively, but also experienced a high team performance. Thus, the results indicate that we achieved our design goal 2.

Table 6: Satisfaction by sample

	Experimental Sample		Control Sample			
	N	Mean	N	Mean	t(df)	p-value (2-t.)
P6 (SP)	47	4.5390	52	4.7308	t(97) = 0.668	p = 0.506
P6 (SO)	47	4.8628	49	5.1316	t(94) = 1.031	p = 0.305
P6 (TOOLDIF)	47	3.4734	49	3.4745	t(94) = 0.008	p = 0.994
P6 (Effic.)	47	4.9901	51	5.1739	t(96) = 0.852	p = 0.396
P6 (Effec.)	47	4.8613	48	5.1128	t(93) = 1.158	p = 0.250
P6 (Prod.)	46	4.9967	50	5.1420	t(94) = 0.706	p = 0.482
P6 (TP)	46	5.2319	45	4.7630	t(89) = 2.225	p = 0.029*
Statistical significance *p<0.05, **p<0.01 / 7-point Likert scale						

6. Conclusion, Contribution, Limitations and Future Research

In this DSR study we addressed the research question of how CE principles can be used to develop CL instructional designs that increase learners' expertise and foster higher-order thinking skills in large-scale lectures. For that purpose, we derived two design goals and report the core findings from our DSR study. We developed a CL instructional design inherent in a CWP (Fig. 3), build an expository instance in the form of a PSA and evaluated the PSA in an online experiment in a large-scale lecture with undergraduate business information students. Our results provide evidence that we were able to achieve both design goals. The results indicate that learners in the experimental sample significantly outperformed learners in the control sample in terms of HLL performance. Our results prove that engineered CWPs for management education benefit from structured collaboration with packaged facilitation expertise. Interestingly, the results are contrary to traditional CL literature which argues that learner interactions should be rather ad hoc and less restricted by processes (Dillenbourg, 2002). In contrast, CE literature argues that process restrictions can increase the number, quality, and creativity of ideas under certain conditions (Briggs et al., 2013). In our study, we provide evidence that process restrictions of learners' behavior do not impede group performance. The execution of our PSA helps learners to significantly perform better and avoid distractions than the execution of a rather traditional, but IT-based CL experience. Thus, our results contribute to the body of CE and CL literature. For the CE body of knowledge, we a) provide generalizable requirements for developing CWPs for CL; b) provide the design for a novel CWP in the educational domain; c) provide implications on how to build a PSA and deploy it in the field with practitioners; d) provide empirical evidence by testing the PSA with undergraduate business information students. This provides further evidence that education is a promising application domain of CE. From an educational perspective, we developed a novel instructional design for IT-supported, reusable, and scalable CL experiences for large scale lectures. We provide evidence that, engineered CWPs are suitable a) to enrich CL experiences and support learners in increasing expertise and achieve higher-order thinking skills; b) to make IT-supported CL instructional designs executable by practitioners (i.e., learners); c) to make CL experiences reusable and scalable for large-scale lectures; d) to reinvent traditional CL instructional designs by enriching them with IT, packaged facilitation expertise and process restrictions.

Nevertheless, our study is not without limitations that provide space for future research. First, there might be a self-selection bias as participating in our 'CL experience' was voluntary. Learners had the chance to receive up to four bonus points for their final exam. It is possible that only highly motivated learners participated in our experiment. However, in our case, most students of the information systems lecture participated in the 'CL experience'. Second, our PSA is applicable to one case in one knowledge domain in one large-scale lecture. To demonstrate the generalizability of the PSA, future research could adapt the PSA for a different lecture with a different case. Third, the collaboration task of the PSA focused on a learning case with four sub-tasks to increase learners' expertise and higher-order thinking skills. Such learning cases are suitable and important learning tasks to foster HLL, but not the only opportunity. Therefore, future research could focus on other learning tasks and group deliverables and create further CWPs for management education.

Acknowledgements

The research was partially funded by the German Federal Ministry of Education and Research in course of a junior research group funding for the project HyMeKI, (FKZ 01IS20057B).

References

Arbaugh, J. B. (2010). *Online and Blended Business Education for the 21st Century: Current Research and Future Directions*. Chandos Pub.

Benalian, A. (201X). Are We Aligned…Enough? The Effects of Perceptual Congruence Between Service Teams and Their Leaders on Team. *Jouranl of Service Research (JSR)*, *XX*(X), 1–17.

Briggs, R. O., Kolfschoten, G., Vreede, G.-J. de, & Douglas, D. (2006). Defining Key Concepts for Collaboration Engineering. In *Americas Conference on Information Systems,* Acapulco, Mexico.

Briggs, R. O., Kolfschoten, G. L., Vreede, G.-J. de, Lukosch, S., & C, A. C. (2013). Facilitator-in-a-Box: Process Support Applications to Help Practitioners Realize the Potential of Collaboration Technology. *JMIS*, *29*(4), 159–193.

Briggs, R. O., Reinig, B. A., & Vreede, G.-J. de (2008). The Yield Shift Theory of Satisfaction and Its Application to the IS/IT Domain. *JAIS*, *9*(5), 267–293.

Damon, W. (1984). Peer Education: The Untapped Potential. *Journal of Applied Developmental Psychology*, *5*, 331–343.

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?*. Open Universiteit Nederland.

Gallupe, R. B., Dennis, A. R., Cooper, W. H., Valacich, J. S., Bastianutti, L. M., & Nunamaker, J. F. (1992). Electronic brainstorming and group size. *Academy of Management Journal*, *35*(2), 350–369.

Hevner, A. R. (2007). A Three Cycle View of Design Science Research. *Scandinavian Journal of Information Systems*, *19*(2), 87–92.

Ingham, A. G., Levinger, G., Graves, J., & Peckham, V. (1974). The Ringelmann effect: Studies of group size and group performance. *Journal of Experimental Social Psychology*, *10*(4), 371–384.

Jones, J. M. (2014). Discussion Group Effectiveness is Related to Critical Thinking through Interest and Engagement. *Psychology Learning and Teaching*, *13*(1), 12–24.

Kolfschoten, G. L., & Santanen, E. L. (2007). Reconceptualizing Generate thinkLets: the Role of the Modifier. In *40th Hawaii International Conference on System Sciences*.

Kolfschoten, G. L., & Vreede, G.-J. de (2009). A Design Approach for Collaboration Processes: A Multimethod Design Science Study in Collaboration Engineering. *JMIS*, *26*(1), 225–256.

Krathwohl, D. R. (2002). A Revision of Bloom's Taxonomy: An Ovierview. *Theory into Practice*, *41*(4), 212–218.

Manca, F., & Meluzzi, F. (2020). *Strengthening online learning when schools are closed: The role of families and teachers in supporting students during the COVID-19 crisis*. https://read.oecd-ilibrary.org/view/?ref=136_136615-o13x4bkowa&title=Strengthening-online-learning-when-schools-are-closed

Moll, L. C. (2013). *L.S. Vygotsky and Education*. Taylor & Francis.

Moore, M. G. (1989). Editorial: Three Types of Interaction. *American Jn. of Distance Education*, *3*(2), 1–7.

OECD. (2016). *Skills for a Digital World: OECD Digital Economy Papers.*

OECD. (2020). *Going Digital integrated policy framework: OECD Digital Economy Papers* (No. 292).

OECD. (2021). *Principles for an Effective and Equitable Educational Recovery*. Paris.

Oeste-Reiß, S., Bittner, E. A. C., & Söllner, M. (2017). Yes You Can – Empowering Lecturers to Simulate Collaboration among Learners in the Disciplines of Problem-Solving and Critical Thinking Regardless of Class Size. In *Internat. Tagung Wirtschaftsinformatik,* St.Gallen, Schweiz.

Oeste-Reiß, S., Söllner, M., & Leimeister, J. M. (2016). Development of a Peer-Creation-Process to Leverage the Power of Collaborative Knowledge Transfer. In *Hawaii International Conference on System Sciences,* Kauai, Hawaii, USA.

Ries, B. C., Diestel, S., Shemla, M., Christina, L. S., Jungmann, F., Wegge, J., & Schmidt, K.-H. (2013). Age Diversity and Team Effectiveness. In C. M. Schlick, E. Frieling, & J. Wegge (Eds.), *Age-Differentiated Work Systems.* Springer.

Schrum, L., & Berge, Z. L. (1997). Creating Student Interaction within the Educational Experience: A Challenge for Online Teachers. *Canadian Jn. of Educational Communication*, *26*(3), 133–144.

Smith, M. K., Wood, W. B., Adams, W. K., Wieman, C., Knight, J. K., & Guild, T. T. S. (2009). Why Peer Discus. Improves Student Performance IN-Class Concept Questions. *SCIENCE*, *323*, 122–124.

Sweller, J., Ayres, P., & Kalyuga, S. (2011). *Cognitive Load Theory*. Springer.

Thurmond, V., & Wambach, K. (2004). Understanding interactions in distance education: A review of the literature. *International Jn. of Instructional Technology and Distance Learning*, *1*(1), 9–26.

Topping, K. J. (2005). Trends in peer learning. *Educational Psychology*, *25*(6), 631–645.

van Dijk, A. M., Eysink, T. H. S., & de Jong, T. (2020). Supporting Cooperative Dialogue in Heterogeneous Groups in Elementary Education. *Small Group Research*, *51*(4), 464–491.

Vreede, d. G.-J., Briggs, R. O., & Massey, A. P. (2009). Collaboration Engineering: Foundations and Opportunities. *JAIS*, *10*, 121–137.

Vreede, G.-J. de, & Briggs, R. O. (2019). A Program of Collaboration Engineering Research and Practice: Contributions, Insights, and Future Directions. *JMIS.*, *36*(1), 74–119.

Winkler, R., Briggs, R. O., Vreede, G.-J. de, Leimeister, J. M., Oeste-Reiß, S., & Söllner, M. (2019). *Towards a Technique for Modeling New Forms of Collaborative Work Practices – The Facilitation Process Model 2.0.*