Learning and Instruction 75 (2021) 101487

Contents lists available at ScienceDirect

Learning and Instruction

journal homepage: www.elsevier.com/locate/learninstruc

Learning to diagnose collaboratively – Effects of adaptive collaboration scripts in agent-based medical simulations

Anika Radkowitsch[a,b,*], Michael Sailer[a,b], Ralf Schmidmaier[a,c], Martin R. Fischer[a,d], Frank Fischer[a,b]

[a] Munich Center of the Learning Sciences (MCLS), Ludwig-Maximilians-Universität München, Munich, Germany
[b] Department Psychologie, Ludwig-Maximilians-Universität München, Munich, Germany
[c] Medizinische Klinik und Poliklinik IV, LMU Klinikum, Ludwig-Maximilians-Universität München, Munich, Germany
[d] Institut für Didaktik und Ausbildungsforschung in der Medizin, LMU Klinikum, Ludwig-Maximilians-Universität München, Munich, Germany

ARTICLE INFO

Keywords:
Adaptive collaboration scripts
Basic psychological needs
Collaborative diagnostic
Reasoning
Medical education

ABSTRACT

We investigated how medical students' collaborative diagnostic reasoning, particularly evidence elicitation and sharing, can be facilitated effectively using agent-based simulations. Providing adaptive collaboration scripts has been suggested to increase effectiveness, but existing evidence is diverse and could be affected by unsystematic group constellations. Collaboration scripts have been criticized for undermining learners' agency. We investigate the effect of adaptive and static scripts on collaborative diagnostic reasoning and basic psychological needs. We randomly allocated 160 medical students to one of three groups: adaptive, static, or no collaboration script. We found that learning with adaptive collaboration scripts enhanced evidence sharing performance and transfer performance. Scripting did not affect learners' perceived autonomy and social relatedness. Yet, compared to static scripts, adaptive scripts had positive effects on perceived competence. We conclude that for complex skills complementing agent-based simulations with adaptive scripts seems beneficial to help learners internalize collaboration scripts without negatively affecting basic psychological needs.

Diagnosing collaboratively is part of many physicians' daily routines. For instance, physicians discuss their patients' symptoms with other physicians from different medical subspecialities. Although collaborative diagnostic reasoning is a crucial competence in routine medical care, empirical research is largely lacking (Kiesewetter, Fischer, & Fischer, 2017). The few available studies suggest that collaborative diagnosing is difficult, and physicians often fail to pool their knowledge appropriately, which can lead to wrong diagnoses (Tschan et al., 2009). Hence, understanding and facilitating collaborative diagnostic reasoning, and particularly sharing and elicitation of information, seems necessary. Simulations have been found to be effective for enhancing the learning of complex skills (Gegenfurtner, Quesada-Pallarès, & Knogler, 2014). Simulations are models of real-world scenarios in which learners can act as if they were in that situation, thereby practicing complex skills (Gegenfurtner et al., 2014). However, particularly during early stages of the development of complex skills, it seems beneficial to provide learners with additional scaffolding beyond providing realistic problem solving opportunities (Chernikova et al., 2020). Collaboration scripts are scaffolds that structure collaboration (Fischer, Kollar, Stegmann, & Wecker, 2013) and were found to be effective for facilitating collaborative learning (Radkowitsch, Vogel, & Fischer, 2020). Yet, collaboration scripts have been criticized for being too coercive, by that reducing learners' self-determination and thus reducing motivation and impairing learning (Dillenbourg, 2002). A promising solution for the respective criticism seems to provide adaptive support that adjusts to the learners' needs. However, the evidence for the effectiveness of adaptive scaffolding to increase motivation and learning is ambiguous (Stegmann, Mu, Gehlen-Baum, & Fischer, 2011). Hence, the present study addresses the questions to what extent adaptive collaboration scripts (1) enhance the learning of complex skills such as collaborative diagnostic reasoning in medicine and (2) ensure self-determination when learning with simulations.

* Corresponding author. Department Psychologie, Ludwig-Maximilians-Universität München, Leopoldstr. 13, 80802 München, Germany.
E-mail addresses: Anika.Radkowitsch@psy.lmu.de (A. Radkowitsch), michael.sailer@psy.lmu.de (M. Sailer), ralf.schmidmaier@med.uni-muenchen.de (R. Schmidmaier), martin.fischer@med.uni-muenchen.de (M.R. Fischer), frank.fischer@psy.lmu.de (F. Fischer).

https://doi.org/10.1016/j.learninstruc.2021.101487
Received 20 May 2020; Received in revised form 22 February 2021; Accepted 21 March 2021
Available online 19 April 2021
0959-4752/

1. Background

1.1. Collaborative diagnostic reasoning

As in other collaborative problem-solving contexts (e.g., OECD, 2017), physicians need to be able to diagnose individually and to engage in collaborative activities when diagnosing collaboratively. For a proper understanding of collaborative diagnostic reasoning, examining both, individual and collaborative cognitive processes as well as their interaction is thus necessary. *Individual* diagnostic reasoning is an epistemic process (Heitzmann et al., 2019) with the goal to identify an accurate diagnosis (Simmons, 2010) and to reduce uncertainty to the degree that enables taking appropriate action (Charlin et al., 2012). Generally, epistemic reasoning processes have been described as coordination between hypotheses and evidence by generating hypotheses and deriving predictions from and testing of hypotheses in the light of evidence (Klahr & Dunbar, 1988). In the context of medical diagnostic reasoning, physicians suggest differential diagnoses (i.e., hypotheses) based on findings and symptoms (i.e., evidence) which in turn allow deriving predictions about further findings and symptoms. To test these predictions, the generation of further evidence is often crucial. For example, physicians perform specific physical examinations or conduct laboratory tests to get more information about the patient's health status based on prior hypotheses (Charlin et al., 2012). These diagnostic reasoning processes are based on different types of knowledge such as conceptual biomedical or strategic knowledge. These types of knowledge and their efficient organization are essential for efficient diagnostic reasoning processes (Feltovich & Barrows, 1984; Klein, Otto, Fischer, & Stark, 2019; Stark, Kopp, & Fischer, 2011). According to the illness script theory, medical knowledge is stored in so-called *illness scripts* which organize medical knowledge based on disease entities, their underlying pathophysiological processes, the resulting signs and symptoms, and enabling conditions. An efficient organization develops with increasing medical experience which allows to relate a patient's signs, symptoms, and enabling conditions to the respective disease. Thus, illness-scripts allow physicians to diagnose accurately and fast based on pattern recognition (Charlin et al., 2007; Feltovich & Barrows, 1984).

Yet, complex cases often require that different medical experts combine their efforts to diagnose the patient *collaboratively*. For example, attending physicians do not generate additional evidence themselves but consult more specialized diagnosing physicians such as radiologists or pathologists to generate and evaluate evidence. In such situations, collaboration has the function to pool knowledge and skills distributed among collaborators in order to reach the solution of a problem (OECD, 2017). Collaborators do so by engaging in collaborative activities of which sharing, elicitation, negotiation, and regulation were particularly in the focus of recent models of collaboration (e.g., Hesse, Care, Buder, Sassenberg, & Griffin, 2015; Liu, Hao, von Davier, Kyllonen, & Zapata-Rivera, 2015; OECD, 2017; Sun et al., 2020). By sharing and eliciting information, collaborators contribute to both the processing of information on a group level (Hinsz, Tindale, & Vollrath, 1997), as well as the formation of a shared mental representation of the problem and its possible solutions (Meier, Spada, & Rummel, 2007; Roschelle & Teasley, 1995) and requires collaborators to take the audience's background into account. Negotiating is particularly important in case of conflicts and can prevent groups from premature closure or from ignoring dissenting evidence (Nickerson, 1988; Patel, Kaufman, & Arocha, 2002). Regulation refers to coordinating goals and strategies to reach these goals and requires collaborators to reflect on their own collaborative activities (Järvelä & Hadwin, 2013). In short, when diagnosing collaboratively, physicians share, negotiate, or elicit information and the results of the individual diagnostic activities (e.g., evidence or hypotheses) and coordinate their collaborative diagnostic reasoning. By that physicians pool their knowledge and effort in order to reach the common goal which is, for instance, to identify the patients' disease or a suitable treatment.

In which collaborative activities physicians engage in depends on the situational needs. However, the pooling of information (i.e., sharing and elicitation) has received much attention in psychological and medical research as it has been found to be difficult but, at the same time, highly relevant for the success of collaboration. A number of studies in the medical context showed that physicians often fail to share or elicit crucial patient information among each other which negatively affected individual diagnostic reasoning processes such as the generation of hypotheses (Larson, Christensen, Franz, & Abbot, 1998; Tschan et al., 2009), or the generation and evaluation of evidence (Davies et al., 2018). Brady, Laoide, McCarthy, & McDermott (2012) showed that in radiology diagnostic errors often occur due to missing medical information. For instance, radiologists are not informed about prior surgeries or secondary diagnoses which may lead to misinterpretations of radiologic evidence and, thus, mislead their diagnostic reasoning process. Therefore, although general collaborative problem-solving processes often also require negotiating and regulation, in this paper we focus on the sharing and elicitation of evidence when collaboratively generating evidence since these collaborative diagnostic processes seem crucial but often deficiently functioning.

Under which conditions are collaborators successful? Different strands of research (e.g., transactive memory theory, shared mental models) highlight the relevance of *meta-knowledge*, that is knowledge about the other team members' roles, their responsibilities and their tasks (Cannon-Bowers, Salas, & Converse, 1993; Engelmann & Hesse, 2011; Wegner, 1987). Meta-knowledge indicates how knowledge is distributed among the team and allows team members to anticipate the team members' activities and to adapt the own activities accordingly. Thus, meta-knowledge is likely to affect which information is shared with or elicited from collaboration partners (Fiore et al., 2010).

In sum, collaborative diagnostic reasoning requires combining individual diagnostic skills such as the generation of evidence and collaborative skills such as sharing and elicitation of diagnostically relevant information. To successfully diagnose collaboratively, physicians need to apply medical knowledge as well as knowledge about the team members' roles and responsibilities to a patient case.

1.2. Facilitating collaborative diagnostic reasoning with agent-based simulations

Offering learners opportunities to apply their knowledge to realistic cases is considered crucial for the development of complex skills (Kolodner, 1992) such as collaborative diagnostic reasoning. By applying their knowledge to cases, physicians reorganize and encapsulate their knowledge to build efficient illness scripts (Feltovich & Barrows, 1984). Simulations provide learners with the opportunity to apply knowledge to specific cases in standardized settings (e.g., rare diseases, situations with high stakes) while simultaneously allowing to reduce the complexity of a real practice situation (Siebeck et al., 2011). Typically, these simulations focus on a smaller range of sub-skills, and allow for repetition, error, time-outs, and systematic debriefing. Several systematic reviews have been conducted, concluding that learning with simulations is beneficial for the development of complex skills in a broad range of conditions (Cook, 2014; Gegenfurtner et al., 2014). Thus, simulations offer beneficial learning opportunities by engaging learners with important aspects of a task, enabling them to apply knowledge and practice certain skills. However, prior research has repeatedly shown that unsupported problem solving is likely to overwhelm learners, particularly during early phases of skill development (Belland, Walker, Kim, & Lefler, 2017; Kirschner, Sweller, & Clark, 2006). Designing simulation-based learning environments with additional instructional support such as scaffolding has been found to be a promising way to enhance further the effectiveness of simulations (Chernikova et al., 2020).

When simulating collaborative tasks, a recent approach is the collaboration with simulated computer agents (Graesser et al., 2018). In agent-based collaboration, one or more learners collaborate with one or more computer agents which are computer programs designed to act similar to humans to solve a problem or task (Rosen, 2015). Such agents can fulfill different roles, for example as pedagogical agent (e.g., AutoTutor, Nye, Graesser, & Hu, 2014), or as substitute for a collaboration partner (e.g., Stadler, Herborn, Mustafić, & Greiff, 2019). Simulating collaboration partners during collaborative tasks allows controlling the effect of possibly influencing variables such as motivation or expertise of the collaborators, thus offering highly standardized research settings (Graesser et al., 2018). The targeted aspects of collaboration are often selected based on empirical analyses showing a particular need for advancing these skills.

1.3. Scaffolding collaborative diagnostic reasoning with collaboration scripts

To learn complex skills, instructional support beyond providing a problem scenario seems beneficial (Wood, Bruner, & Ross, 1976). External collaboration scripts are a scaffold used for facilitating interaction during collaborative learning by prompting specific collaborative activities (Fischer et al., 2013). For example, learners receive prompts to share particular information (Noroozi, Biemans, Weinberger, Mulder, & Chizari, 2013). The idea of external collaboration scripts is based on the *script theory of guidance* (Fischer et al., 2013), which assumes that internal collaboration scripts guide any behavior and cognition concerning collaboration. Such internal collaboration scripts are assumed to structure knowledge about specific collaborative practices. During collaborative practices, this knowledge is flexibly activated depending on the situational characteristics and the collaborators' goals. External collaboration scripts complement less functional cognitive script components. By engaging learners in beneficial collaborative activities, external collaboration scripts can facilitate the internalization of functional script components (Fischer et al., 2013). Beyond these theoretical considerations, there is empirical evidence for the effectiveness of collaboration scripts for advancing the development of collaboration skills. For example, Rummel and Spada (2005) found that dyads of medical students and psychological students who were supported with a collaboration script showed less deviations from an exemplary collaborative process compared to dyads not supported by an external collaboration script. In a meta-analysis, Radkowitsch et al. (2020) found that collaboration scripts effectively facilitate the learning of collaboration skills in the context of computer-supported collaborative learning. Yet, most of the prior studies targeted advancing domain learning, and none of the studies employed simulations systematically to advance collaboration skills. Thus, it is not clear whether collaboration scripts are effective when implemented in simulations and used to advance profession-specific skills such as sharing or elicitation of evidence.

1.3.1. Adaptive collaboration scripts

As described above, external collaboration scripts aim to complement less functional internal collaboration scripts. That means that to be effective, collaboration scripts should be adapted to the learners' actual proficiency level (Fischer et al., 2013). When providing too detailed guidance – for which collaboration scripts have been criticized (Dillenbourg, 2002) – the learning of more advanced learners is hindered because it restricts the learners' natural collaboration processes (i.e., their internal collaboration scripts). This phenomenon became known as *over-scripting* (Dillenbourg, 2002). Yet, most studies that aim to avoid over-scripting and to take learners' zone of proximal development (Vygotsky, 1978) into account, use static techniques such as fading. For example, Stegmann et al. (2011) compared static and fading collaboration scripts supporting argumentation to unstructured collaborative learning. They found that learning with static and fading collaboration scripts enhanced the quality of argumentation with descriptively larger effects for learning with the static collaboration script. By fading after a predefined sequence, learners' internal collaboration script could have been misaddressed, posing unnecessary extraneous cognitive load on learners. Consequently, learners could have failed to internalize the collaboration scripts (Wecker & Fischer, 2011).

Thus, to exploit the full potential of adaptive support modeling learners' collaborative activities and compare them to an ideal collaborative process seems beneficial. In the case of previously defined deviations, learners receive just-in-time support in the form of so-called adaptive collaboration scripts (Karakostas & Demetriadis, 2011; Tchounikine, Rummel, & McLaren, 2010). There are different ways to implement such adaptive support with varying degrees of complexity. For example, in a small-scale study on conducting chemical experiments, Tsovaltzi et al. (2010) used a wizard of Oz approach in which a human adept intervened in the collaborative process in predefined situations. Other studies used the amount of contribution of each collaborator (Constantino-Gonzalez, Suthers, & de los Santos, 2003), or learners' automatically assessed problem-solving strategies (Diziol, Walker, Rummel, & Koedinger, 2010) as indices for quality of collaboration. Yet, studies on adaptive collaboration scripts are scarce and provide no clear picture of their effectiveness for advancing collaboration. A reason for this could be that research on collaboration often takes place in rather unstandardized human-to-human collaborative settings. That means that the composition of learners affects collaboration processes (Kirschner, Sweller, Kirschner, & Zambrano, 2018) and consequently could also influence the effects of collaboration scripts. Thus, the reported effects of collaboration scripts could be affected by their particular implementation and by the high noise due to several real collaboration partners. Therefore, an agent-based realization of collaboration provides a standardized setting for investigating the effect of adaptive collaboration scripts.

1.3.2. Collaboration scripts and psychological need satisfaction

Collaboration scripts were criticized for restricting learners and thereby negatively affecting learners' self-determination and, thus, intrinsic motivation (Dillenbourg, 2002). This criticism relates to the self-determination theory (Deci & Ryan, 1985) that describes that the feelings of autonomy, competence, and social relatedness are basic psychological needs and thus crucial determinants of intrinsic motivation and, consequently, human behavior. The feeling of autonomy refers to whether the learners perceive their actions as congruent and volitional. The feeling of competence relates to the feeling of efficacy of own actions. The perception of social relatedness refers to a feeling of being connected with a specific community (Deci & Ryan, 1985). Whereas critics argue that collaboration scripts reduce the perceived autonomy (Wise & Schwarz, 2017), collaboration scripts might also enhance the feeling of competence and social relatedness by enabling them to experience successful collaboration (Radkowitsch et al., 2020). Yet, the empirical evidence for a negative effect of collaboration scripts on self-determination or intrinsic motivation is scarce. In their meta-analysis, Radkowitsch et al. (2020) identified only six studies investigating the effect of collaboration scripts on motivation, which yielded a combined null effect on motivation.

If collaboration scripts negatively affected self-determination due to their limitation of choices, then collaboration scripts that are less structured should have a less negative impact on learners' self-determination. Stegmann et al. (2011) found initial evidence for that as learners supported either with a high or low structured collaboration script were descriptively less intrinsically motivated compared to learners who collaborated freely. However, this effect was not significant, and the study does not allow for differentiated conclusions on how learners' basic psychological needs were affected by learning with collaboration scripts. Besides, participants of the study could be novices and they could have required a higher degree of structure for successful collaboration. Then, the collaboration script should not automatically interfere with the learners' motivation. As the learning environment – and hence also the collaboration script – is considered crucial for fostering psychological need satisfaction (Deci & Vansteenkiste, 2004), we argue that collaboration scripts could have a negative impact on learners' perceived autonomy if the provided structure interferes with the learners' needs for support. Thus, adaptive collaboration scripts should affect learners' perceived autonomy less than collaboration scripts that are not adapted to the learners' specific needs for support. Concerning the feeling of competence, we assume that adaptive collaboration scripts are better tailored to the learners' needs, which could enable learners to adapt their collaborative activities better. Concerning the feeling of social relatedness, we assume that collaboration scripts could have positive effects. As learning with collaboration scripts could lead to equal participation and increase team functioning, we assume that both adaptive and static collaboration scripts increase the perception of social relatedness with higher effects of an adaptive collaboration script.

1.4. The present study

We seek to identify conditions under which facilitating collaborative diagnostic reasoning of medical students using agent-based simulation and information sharing scripts (ISS) is effective. In this study, ISS are collaboration scripts that focus on facilitating the sharing and elicitation of diagnostic information during collaboration. More particularly, we aim at facilitating the elicitation and sharing of evidence, which are considered essential subskills of collaborative diagnostic reasoning. The goals of the present study are twofold: We examine the effects of adaptive and static ISS on (1) collaborative diagnostic reasoning, and (2) the satisfaction of basic psychological needs. We pose the following research questions:

(1) What are the effects of an adaptive and a static ISS on collaborative diagnostic reasoning and more specifically on a) evidence elicitation and b) evidence sharing in an agent-based simulation? We hypothesize that both static and adaptive collaboration scripts have positive effects on a) evidence elicitation and b) evidence sharing compared to an unsupported control group, with adaptive ISS resulting in larger effects than static ISS.
(2) What are the effects of an adaptive and a static collaboration script on the basic psychological need satisfaction in the context of diagnosing collaboratively in an agent-based simulation? We hypothesize that both ISS have negative effects on the perceived autonomy but positive effects on the perceived competence and perceived social relatedness compared to an unsupported control group. Learning with adaptive ISS should result in higher perceived autonomy, higher perceived competence, and higher perceived social relatedness than static ISS.

2. Method

2.1. Sample and design

We conducted an experiment with a one-factorial design with the three *levels adaptive ISS, static ISS*, and no further instructional support (*control group*). Before recruiting participants, we conducted a power analysis based on medium effect sizes reported in prior studies (Radkowitsch et al., 2020). To detect a medium effect size of $f = 0.25$ in an ANCOVA design, a minimum sample of 159 participants (53 per group) was needed (presuming $\alpha = .05$, $1 - \beta = 0.80$). Medical students in their 3rd clinical year and higher were recruited to participate in the study voluntarily. These medical students have usually already completed medical clerkships in which collaboration with other physicians is necessary. However, systematic training of professional collaboration is not a formal part of medical education and their prior experience in medical collaboration can, thus, be described as low. Further, they had no experience with the agent-based simulations that we used in the experiment. All participants were randomly distributed to one of three groups. The final sample consisted of 160 participants ($N_{female} = 110$) of whom 54 were in the adaptive ISS condition and 53 in the static ISS and control condition each. Participants were, on average, 25 years old (SD = 3.1) and in their 5.24 years of medical school (SD = 1.07) of a 6-year program.

2.2. The agent-based medical simulation

To foster and measure collaborative diagnostic reasoning, we developed a simulation as well as cases of suitable fictitious patients suffering from fever of unknown origin. We chose the collaborative generation of evidence between internists and radiologists as simulated scenario since this is a situation that is common in emergency departments but often deficiently functioning (e.g., Brady et al. (2012)) and which offers a high potential for future standardization. In these situations, internists consult radiologists to generate evidence in order to reduce the uncertainty with respect to a specific diagnosis. Radiologists are better able to reliably contribute to the diagnostic process if internists precisely specify and justify the kind of evidence needed to reduce uncertainty (i.e., elicitation of evidence) and if internists report any relevant signs, symptoms, and prior conditions of the patient that could influence the radiologists' diagnostic process (i.e., sharing of evidence). As such, this situation does not represent a mere distribution of tasks since the quality of the activity of one person depends on the quality of collaborative activities of another person. For the internist to optimally collaborate with the radiologist, meta-knowledge of radiologists' tasks, role, and responsibilities is beneficial (e.g., Brady et al., 2012).

The simulation was developed in collaboration with software developers, medical educators, physicians, and psychologists and validated in expert workshops, a pilot study, as well as a comprehensive validation study comparing cognitive processes of medical students and experienced internists when working in the simulation (see Radkowitsch, Fischer, Schmidmaier, & Fischer, 2020). We implemented the simulation in the learning platform CASUS (http://www.casus.net) with which most participants already had prior experience in their curriculum.

Participants acted in the role of an internist working in an emergency department and were required to collaborate with a simulated radiologist to generate further evidence in each patient case. Overall, the interaction between the participant and the simulated radiologist in each case consists of filling in the request form and receiving e-mail-like text messages from the simulated radiologist containing a short introduction, the decline or acceptance of the request, and the report of radiological findings. The participants first received the electronic patient file containing information about the patient's admission, medical history, physical examination, and laboratory tests. In the next step, the participants requested a radiological test to generate further evidence about the cause of the presented symptoms in order to reduce uncertainty respective a potential diagnosis. For that, participants first contacted the radiologist by pressing the button "request radiological test". A simulated radiologist then sent an e-mail-like prompt shortly introducing herself as the radiologist on duty and asking to fill in a request form. Participants then requested the radiologic test by choosing a specific radiologic method (e.g., computer tomography) and a body part (e.g., abdomen) using the form. The request form further required participants to share patient information and suspected diagnoses that justified the test and helped the radiologist to interpret the radiologic findings. For that learners could tick off patient information from a thematically clustered list containing all information from the electronic patient file and type in diagnoses in a free text field. The test chosen by the participant specified the required information. For example, tests based on x-rays required, inter alia, information about a potential pregnancy. The minimum amount of information necessary to justify the request was determined in advance by medical experts. Only when participants shared sufficient patient information, the simulated radiologist actually conducted the respective test and interpreted the findings. Participants then received a detailed report on the generated evidence in form of a text message. Otherwise, the simulated radiologist rejected the request stating that the shared information was not sufficient to justify the specific radiologic test and asked the participant to revise and resubmit the request. Finally, the medical students concluded the patient case by suggesting and justifying a final diagnosis.

2.3. Treatment

We supported participants in the treatment groups either with an adaptive or a static ISS that were both provided in text messages by the simulated radiologist during the learning phase. Both ISS consisted of three types of prompts containing meta-knowledge (i.e., information about the radiologists' role, task, and responsibilities). Firstly, the ISS included general, case-independent details on the radiologists' task and information that is helpful for them to complete the task (first type of ISS prompt). For instance, the radiologist explains that the request should, inter alia, provide information on the patient's main symptoms and their course as this helps the radiologist to judge what and where to look for. This first type of ISS prompt addressed the learners' evidence sharing skill since it should help learners to precisely identify information relevant for the collaborating radiologist to complete their task. Secondly, the ISS contained case-specific meta-knowledge about how radiologists generate evidence for specific suspected diagnoses (second type of ISS prompt). This information was included for 45 differential diagnoses that were most relevant for the patient cases. For instance, the radiologist explains that if the patient is suspected to suffer from a pneumonia, the radiologist typically tries to differentiate air-filled parts in the lung from liquid-filled parenchyma, and that x-rays often are not sufficient for a differentiated evaluation due to overlays. This type of ISS prompt should help learners to specify and justify the type of evidence needed from the collaborator, the radiologist, and thus addresses the learners' evidence elicitation skill. Thirdly, the ISS provided meta-knowledge about specific radiologic examinations and about how such imaging procedures could potentially harm patients (third type of ISS prompt). Here, the radiologist explained which information helps radiologists to judge the risk of a specific radiologic test. For example, the radiologist explains that radiocontrast can have negative effects on the patients' kidneys which is why radiologists require the kidney status in order to weighing up the benefits and risks of using radiocontrast. This ISS prompt particularly addressed evidence sharing since learners are supported in their decision about whether a specific information is relevant for the collaborator (i.e., the radiologist) or not. As the simulation, the collaboration script prompts were developed in collaboration with experts from medical education, medicine (internists and radiologists), and psychology. The ISS prompts suggested to engage in specific collaborative processes and provided information about why these collaborative processes were meaningful. But the ISS prompts themselves did not force the participants to engage in a specific step at a specific point in time and, thus, can be described as low coercive.

In the *adaptive ISS* condition, participants received the first type of ISS prompt at the beginning of the interaction with the simulated radiologist during the introduction of the radiologist. The second and third prompts were provided by the simulated radiologist whenever the participants submitted a request that was not adequately justified according to the criteria described above. The simulated radiologist checked whether the presented diagnoses and symptoms were compatible with the requested test (second type of ISS prompt) and whether all necessary information for the respective radiologic tests was given (third type of ISS prompt) and answered with the respective prompts. For instance, when participants requested a test with radiocontrast but failed to share information on the kidney function, the request was rejected, and participants received a prompt providing meta-knowledge about the importance of kidney status for the radiologist and asking the learner to share respective information.

Participants in the *static ISS* condition received the ISS in form of a letter and a booklet from the radiologist at the beginning of the learning phase. In the letter, the radiologist first explained the general procedure of a radiologist (first type of ISS prompt) and which specific information was needed for potentially harmful radiologic tests (third type of ISS prompt). A booklet further provided meta-knowledge about the evaluation of specific diagnoses (second type of ISS prompt). The learners could access the letter and booklet any time and as often as they wanted to. When participants in the static ISS condition failed to adequately justify their request, their requests were also rejected by the radiologist. Participants were then required to find the relevant information in the letter and booklet for themselves. Thus, the main difference between the static and the adaptive ISS is that learners in the adaptive ISS condition receive only the ISS prompt the system identified as their current need for support whereas learners in the static ISS condition received all ISS prompts at once. Therefore, adaptive ISS should be less likely to interfere with the learners' need for autonomy compared to the static ISS. Moreover, since the adaptive ISS should help learners to implement the script prompts and consequently enhance the learners' perceived competence. Finally, in the adaptive ISS condition, the simulated radiologist reacts more directly to the learners' action. Therefore, learners supported with an adaptive ISS should perceive a higher social relatedness compared to learners in the static ISS condition. After the learning phase, learners were asked to return the letter and the booklet.

The requests of participants in the *control condition* were also rejected when they failed to share and elicit the necessary information. However, participants in the control condition did not receive any meta-knowledge prompts from the radiologist.

2.4. Dependent variables

2.4.1. Collaborative diagnostic reasoning performance

We assessed the participants' performance of *evidence sharing* and *evidence elicitation* as two subskills of collaborative diagnostic reasoning. For that, we used log files produced during the interaction with the simulation. The log files consist of all clicks and text entries of the users within the simulation, such as the patient information or diagnoses chosen to share with the radiologist. We used R Studio Version 4.0.2 (R Core Team, 2020) to automatically evaluate the quality of evidence elicitation and evidence sharing based on the expert solutions produced from medical experts. More specifically, the patient information and diagnoses shared, and radiologic tests requested from the learners were automatically matched to relevant patient information, diagnoses and radiologic tests as defined by the expert solutions. Thus, no manual coding was necessary to evaluate the measures evidence elicitation and evidence sharing. The expert solution and scoring procedure are described in more detail below for each measure.

(Transfer) performance of evidence elicitation. As indicator for evidence elicitation, we used the medical relevance of the radiologic tests elicited from the simulated radiologist as defined by the expert solution. This indicator assesses whether learners are able to identify how radiologists would generate the needed evidence, and thus justify their request accordingly in order to convince the radiologist to conduct the test. For each requested test, learners received 1 point if the test was appropriate with respect to the indicated diagnosis and 0 points if learners chose an inappropriate radiologic test. The mean points for all requested radiologic tests were calculated for each patient case scenario. Hence, for each patient case scenario, a maximum of 1 point was possible. To analyze the *performance of evidence elicitation*, we calculated the mean evidence elicitation across all learning cases solved during the intervention. To analyze the *transfer performance of evidence elicitation*, we used the evidence elicitation score of an unsupported posttest case. The internal consistency across all cases, as indicated by Cronbach's alpha was .60.

(Transfer) performance of evidence sharing. We used the relevance of the evidence shared with the radiologist as defined by the expert solution as indicator for evidence sharing. This indicator assesses whether learners are able to identify which information a radiologist would need to optimally conduct the radiologic test and interpret its results. We evaluated the evidence shared during the first request depending on which test was chosen and calculated the proportion of shared relevant evidence to all relevant evidence. Hence, values range between 0 and 1 point with 1 point indicating that all relevant information was shared with the radiologist and 0 points indicating that no relevant information was shared. To analyze the *performance of evidence sharing*, we calculated the mean quality of evidence sharing across all learning cases solved during the intervention. To analyze the *transfer performance of evidence sharing*, we used the evidence sharing score of the unsupported posttest case. The internal consistency across all cases, as indicated by Cronbach's alpha was .76.

2.4.2. Psychological need satisfaction

We assessed psychological need satisfaction directly after the intervention using a scale adapted from Sailer, Hense, Mayr, and Mandl (2017). The scale consisted of three subscales assessing perceived competence, perceived autonomy, and perceived social relatedness. Participants answered all items on a seven-point Likert scale ranging from 1 (I do not agree) to 7 (I totally agree). Perceived competence was measured with four items (Cronbach's alpha = .92). A sample item is "I felt competent during the activity.". The subscale perceived autonomy consisted of three items (Cronbach's alpha = .85). A sample item is "I was able to decide for myself what I would do during the activity.". Social relatedness was measured using three items (Cronbach's alpha = .85). A sample item is "I felt like a part of a team.".

2.4.4. Treatment check

We assessed how many requested radiologic tests were rejected from the simulated radiologist during the intervention to determine whether learners in the adaptive ISS condition did receive any support. For that, we calculated the average absolute number of rejections for the first two test requests and additionally calculated the absolute number of participants whose requests were rejected at least once during the intervention. We included only the first two turns since for solving most patient cases two radiologic examinations are meaningful. The treatment check is successful if learners were rejected at least once per case. The results of the treatment check are reported below.

2.5. Procedure

All participants first answered demographic questions (age, sex, semester, 3 min), and then solved an unsupported pretest case within the agent-based simulation (15 min). During the intervention phase, all participants solved four patient cases (20 min each) in variations of the simulation corresponding to their experimental condition (see above). All patient cases were presented in the same order. Directly after the intervention, we assessed the participants' psychological need satisfaction (5 min). Finally, all participants solved an unsupported posttest case (15 min). All patient cases covered diseases related to fever with unknown origin.

2.6. Statistical analyses

We conducted all statistical analyses with R Studio using the R Version 4.0.2 (R Core Team, 2020) and report inferential results based on a 5% alpha level. We first examined correlations between pretest variables and outcome variables. We found small to moderate correlations between prior evidence elicitation and evidence elicitation performance ($r = 0.26$, $p < .01$) and between prior evidence sharing and evidence sharing performance ($r = 0.36$, $p < .01$) as expected and thus included the respective pretest measures as a covariate.

To analyze whether adaptive and static ISS enhance the performance of evidence elicitation during the intervention and transfer performance of evidence elicitation in an unsupported posttest (Research Question 1a), we conducted two ANCOVAs with the prior evidence elicitation performance as a covariate. To address research question 1b, we conducted two ANCOVAs with evidence sharing performance and transfer performance of evidence sharing as the dependent variable, respectively, and prior evidence sharing performance as a covariate. We addressed the second research question by conducting three ANOVAs with the three basic psychological needs measures perception of competence, autonomy, and social relatedness as dependent variables. Further, examining Q-Q plots and histograms yielded that measures for evidence elicitation, transfer performance of evidence sharing, and autonomy suffered from a non-normal distribution. Therefore, we additionally conducted non-parametric Kruskal-Wallis-Tests for these variables. For all analyses, we tested hypotheses with planned contrasts and analyzed further differences between groups using Tukey corrected post-hoc tests.

3. Results

3.1. Treatment check

As a treatment check, we descriptively analyzed the sum of rejections during the intervention. The control group was rejected most often ($M = 5.528$, $SD = 4.304$), followed by the adaptive ISS condition ($M = 3.759$, $SD = 3.291$) and the static ISS condition ($M = 3.528$, $SD = 3.129$). In the adaptive ISS, there were 5 participants (9.3%) who received no rejections on their requests at all during the intervention. In the static condition, the requests of 6 participants (11.3%) were not rejected by the radiologist. In the control condition, the requests of 4 participants (7.5%) were not rejected. For more detailed number of rejections per group see Table 1. Since these numbers are comparable between the treatment groups, we decided not to exclude these participants to not reduce the power of analyses. These findings show that the majority of participants in the adaptive ISS condition did receive instructional support during the intervention and numbers of participants whose requests were not rejected by the radiologist were comparable between groups. Thus, the treatment check was successful.

3.2. Effects of adaptive and static information sharing scripts on evidence elicitation

Concerning the *evidence elicitation performance*, the descriptive results show that learners supported with the static ISS scored highest, followed by learners in the adaptive ISS condition. Learners in the control group showed the lowest performance of evidence elicitation (see Table 2). The ANCOVA indicates a large effect of the intervention with significant differences between conditions ($F(2,156) = 13.362$, $p <$

Table 1
Absolut number of participants receiving a number of rejections during the intervention.

Number of Rejections	Adaptive ISS	Static ISS	Control condition
0	5	6	4
1	12	11	3
2	6	8	5
3	5	4	5
4	8	10	6
5	6	4	5
6	5	2	8
7	0	0	7
8	3	4	4
9	1	2	1
10	1	0	1
More than 10	2	2	4

Note: ISS = information sharing script.

Table 2
Means and standard deviations for collaborative diagnostic reasoning performance per condition.

	Adaptive ISS	Static ISS	Control condition
	M (SD)	M (SD)	M (SD)
Pretest			
Prior EE perf.	0.728 (0.335)	0.723 (0.312)	0.741 (0.330)
Prior ES perf.	0.644 (0.207)	0.641 (0.239)	0.660 (0.223)
Intervention phase			
EE perf.	0.748 (0.170)	0.858 (0.153)	0.681 (0.232)
ES perf.	0.854 (0.117)	0.815 (0.111)	0.756 (0.143)
Posttest			
EE transfer perf.	0.898 (0.240)	0.896 (0.207)	0.896 (0.241)
ES transfer perf.	0.831 (0.162)	0.767 (0.181)	0.762 (0.188)

Note: ISS = information sharing script, EE = evidence elicitation, ES = evidence sharing. For all variables, the theoretical minimum is 0 and the theoretical maximum is 1.

.001, partial $\eta^2 = 0.146$). The robust Kruskal-Wallis-Test also yielded significant differences ($\chi^2(2) = 22.431$, $p < .001$). The planned contrasts reveal that the scripted groups significantly outperform unscripted groups, whereby learning with the static ISS led to significantly higher evidence elicitation performance than learning with the adaptive ISS (see Table 3). One-sided Tukey-corrected post-hoc comparisons further yielded, that the static ISS significantly enhanced evidence elicitation performance compared to the control group ($M_{Difference} = 0.179$, $SE = 0.035$, $p < .001$), but the adaptive ISS did not ($M_{Difference} = 0.069$, $SE = 0.035$, $p = .066$).

Concerning the *transfer performance of evidence elicitation*, the descriptive results show similar scores in all three groups with the highest score for learners supported with an adaptive ISS during the intervention (see Table 2). The ANCOVA yielded a null effect and no significant differences between conditions ($F(2,156) = 0.003$, $p = .997$, partial $\eta^2 = 0.000$). The robust Kruskal-Wallis-Test also yielded no significant differences ($\chi^2(2) = 0.109$, $p = .947$).

Overall, the results suggest that, as hypothesized, learning with static ISS enhanced evidence elicitation performance, but in contrast to our expectation, learning with an adaptive ISS did not. Further, the effects did not transfer to an unsupported posttest.

3.3. Effects of adaptive and static information sharing scripts on evidence sharing

Concerning the *evidence sharing performance*, learners supported with an adaptive ISS yielded the highest score followed by learners supported with a static ISS. Learners who did not receive any additional support yielded the lowest score for the performance of evidence sharing (see Table 2). The ANCOVA revealed significant and medium-sized effects of the intervention ($F(2,156) = 10.633$, $p < .001$, partial $\eta^2 = 0.120$). Planned contrasts showed that learning with ISS significantly enhanced evidence sharing performance compared to learning without ISS. Yet, there were no significant differences between both ISS conditions (see Table 3). One-sided Tukey-corrected post-hoc tests further show that the adaptive ISS ($M_{Difference} = 0.101$, $SE = 0.022$, $p < .001$) and the static ISS ($M_{Difference} = 0.063$, $SE = 0.022$, $p = .007$) conditions significantly differ from the control condition.

Concerning the *transfer performance of evidence sharing*, the descriptive results yielded the highest score for learners supported with an adaptive ISS, followed by learners supported with a static ISS. Learners who received no additional support during the intervention showed the lowest score in the skill of evidence sharing (see Table 2). The ANCOVA yielded significant small differences between conditions ($F(2,156) = 3.145$, $p = .046$, partial $\eta^2 = 0.039$). The robust Kruskall-Wallis-Test yielded no significant differences ($\chi^2(2) = 5.674$, $p = .059$). The planned contrasts revealed no significant differences between both ISS conditions and the control condition, but significant differences between the adaptive and the static ISS (see Table 3). One-sided Tukey-corrected post-hoc comparisons further show that only learning with adaptive ISS enhances the transfer performance of evidence sharing in an unsupported posttest compared to unstructured learning ($M_{Difference} = 0.074$, $SE = 0.032$, $p = .030$), but learning with a static ISS did not ($M_{Difference} = 0.010$, $SE = 0.032$, $p = .696$).

Hence, we find support for the hypothesis that adaptive and static ISS enhance evidence sharing performance. Yet, adaptive ISS had no larger effects than static ISS. Further, we find support for the hypothesis that adaptive ISS affects the transfer performance of evidence sharing. Yet, static ISS did not affect the transfer performance of evidence sharing.

3.4. Effects of adaptive and static information sharing scripts on basic psychological need satisfaction

Descriptively, learners supported with an adaptive ISS reported the highest *perceived competence*. The lowest feeling of competence was reported by learners supported by the static ISS (see Table 4). The ANOVA yielded a significant small effect of scripting on perceived competence ($F(2,157) = 3.568$, $p = .031$, $\eta^2 = 0.043$). The planned contrasts reveal that both ISS conditions did not differ significantly from the control condition concerning the perceived competence. Yet, the adaptive ISS differed significantly from the static ISS (see Table 3). One-sided Tukey-corrected post-hoc comparisons further show that neither learners supported with an adaptive ISS perceived significantly higher competence compared to the control group ($M_{Difference} = -0.383$, $SE = 0.250$, $p = .157$), nor did learners supported with the static ISS ($M_{Difference} = 0.283$, $SE = 0.251$, $p = .998$).

Table 4
Means and standard deviations for basic psychological need satisfaction per condition.

	Adaptive ISS	Static ISS	Control condition
	M (SD)	M (SD)	M (SD)
Perceived competence	4.963 (1.241)	4.297 (1.303)	4.580 (1.338)
Perceived autonomy	5.735 (1.100)	5.478 (1.395)	5.597 (1.185)
Perceived social relatedness	3.704 (1.458)	3.233 (1.334)	3.704 (1.660)

Note: ISS = information sharing script. For all variables the theoretical minimum is 1 and the theoretical maximum is 7.

Table 3
Planned contrasts for significant group comparisons.

DV	Contrast	*t*	*p*	*r*
EE performance	Scripted groups vs. control group	4.099	<.001	0.312
	Adaptive vs. static collaboration script	−3.173	.002	0.246
ES performance	Scripted groups vs. control group	4.271	<.001	0.324
	Adaptive vs. static collaboration script	1.715	.088	0.136
ES Transfer performance	Scripted groups vs. control group	1.506	.134	0.120
	Adaptive vs. static collaboration script	1.998	.048	0.158
Perceived competence	Scripted groups vs. control group	0.229	.819	0.018
	Adaptive vs. static collaboration script	2.660	.009	0.208

Note: EE = evidence elicitation, ES = evidence sharing.

Learners reported a relatively high *perceived autonomy*, with learners in the adaptive ISS condition reporting the slightly stronger feelings of autonomy compared to the other conditions. Learners supported with static ISS reported the lowest feeling of autonomy (see Table 4). The ANOVA showed that scripting had a non-significant effect on the perceived autonomy ($F(2,157) = 0.581$, $p = .560$, $\eta^2 = 0.007$). The robust Kruskal-Wallis-Test also yielded no significant differences ($\chi^2(2) = 0.657$, $p = .720$).

Concerning the *perceived social relatedness*, learners supported with adaptive ISS and unsupported learners reported the same level of feeling of social relatedness. Learners supported with static ISS reported the lowest feeling of social relatedness (see Table 4). The ANOVA revealed a non-significant effect of scripting on the perceived social relatedness ($F(2,157) = 1.774$, $p = .173$, $\eta^2 = 0.022$).

Overall, the results are in support of the hypothesis that adaptive scripting enhances the feeling of competence. Yet, the findings do not support the hypotheses that learning with adaptive or static ISS affects the perceived autonomy and perceived social relatedness.

4. Discussion

4.1. Effects of adaptive and static ISS on collaborative diagnostic reasoning

The presented study suggests that collaboration scripts can support specific collaborative diagnostic reasoning processes when learning with simulations. More specifically, the study shows that providing learners with knowledge about the collaboration partners' responsibilities and tasks helped them to successfully collaborate (Fiore et al., 2010) in a professional medical situation between internists and radiologists. In contrast to our hypotheses, static and adaptive collaboration scripts differed in their effectiveness for specific subskills during the intervention. We hypothesized that the adaptive collaboration script would better address the learners' zone of proximal development (Vygotsky, 1978) and, therefore, better help learners to apply the script to the collaborative tasks. Yet, the results support those assumptions only for the subskill evidence sharing, but not for evidence elicitation. The performance of evidence elicitation was only facilitated by the static ISS. These findings challenge the assumption that scaffolds adapted to the learners' needs consistently outperform static support. An explanation why evidence elicitation was only facilitated by the static ISS could lie in the implementation of the ISS. The static ISS was implemented as a letter and a booklet that were constantly present during the intervention phase. In contrast to the adaptive ISS which was only presented to the learners when they showed deviations from an optimal collaboration, the learners could apply the static ISS from the beginning and avoid errors. Besides, learners in the static ISS condition could have focused their attention on the booklet which visually dominated the letter due to its length and contained evidence elicitation support. Thus, implementing static or adaptive scaffolds could impact the learners' focus on the learning material.

An interesting finding is that only the effects of the adaptive ISS on evidence sharing transferred to an unsupported posttest, whereas effects of the static ISS on evidence sharing did not. This means that learners supported with a static collaboration script relied on the availability of the information and thus failed to internalize the meta-knowledge necessary to share patient information. Pea (2004) calls such static scaffolds "distributed intelligence" (p. 431) since they support the learners' momentary activity, but the required skills are not internalized and transferred to similar learning situations. At this point it is important to note that we did not directly measure script internalization. However, internalization of the collaboration script is a plausible explanation for the learners' processes in a posttest in which learners were not supported by a collaboration script.

A further explanation might be that the static collaboration script posed a higher extraneous cognitive load on learners than the adaptive collaboration script since learners were required to search through learning materials to identify relevant support. Thus, the high demands of the complex learning environment combined with the static collaboration script could have exceeded learners' cognitive capacity and impeded learning (Sweller, van Merriënboer, & Paas, 2019). It seems that – when combining complex learning environments such as simulations with other types of instructional support – adaptation of scaffolds is necessary in order to not overwhelm learners. Prior findings support this line of argumentation (e.g., Radkowitsch et al., 2020; Schwaighofer et al., 2017). For example, Schwaighofer et al. (2017) found that when combining different scaffolds, sequencing the scaffolds seems particularly relevant for learners with low working memory capacity. Moreover, compared to prior findings showing moderate to large positive effects of collaboration scripts on learning to collaborate in other learning settings (Radkowitsch et al., 2020), the effects of the collaboration scripts on learning to collaborate within this study (i.e. when learning with simulations) were rather small.

Concerning evidence elicitation, learners in all groups scored high in the transfer test. That indicates that learning of evidence elicitation was easier compared to learning of evidence sharing and that using the agent-based simulations for trial-and-error strategies was successful for internalizing this subskill. Yet, the rather high scores on the pretest and posttest case make a differentiated analyses difficult and further analyses with more differentiated measures are necessary. Nevertheless, for the learning of more complex skills such as evidence sharing, complementing agent-based simulations with adaptive collaboration scripts seems beneficial for internalizing the scripts.

4.2. Effects of adaptive and static ISS on basic psychological need satisfaction

Self-determination theory stresses the importance of the environment for the intrinsic motivation of learners (Deci & Vansteenkiste, 2004). Offering a learning environment enriched with collaboration scripts – which guide learners through collaborative activities and therefore have been criticized for diminishing learners' agency (Wise & Schwarz, 2017) – could thus reduce learners' self-determination and intrinsic motivation. Yet, our findings do not support this line of argumentation since we found no significant effects of adaptive and static ISS on perceived autonomy. Though, it is important to stress that both ISS were designed in a way that they relatively little affected learners' choices when compared to collaboration scripts that, for instance, structure learners' communication with sentence openers. Nevertheless, our findings show that it is generally possible to design static and adaptive collaboration scripts that have little negative impact on perceived autonomy.

Our results suggest that an adaptive design of collaboration scripts can have positive effects on perceived competence since we found that adaptive collaboration scripts significantly increased the perceived competence compared to learning with static ISS. We assumed that just-in-time prompts challenged learners at the right level (Vygotsky, 1978), enabling them to easily adjust their collaborative diagnostic activities, which could have led to an increased feeling of competence (Deci & Ryan, 1985). Thus, our findings suggest that instead of having the proposed detrimental effect on motivation, adaptive collaboration scripts even have the potential to increase perceived competence and, thus, intrinsic motivation.

Further, perceived social relatedness was not significantly affected by scripting. Descriptively, learners in the static ISS condition rated their social relatedness the lowest which is surprising, since, in contrast to the control condition, these learners did receive additional information from the simulated radiologist and, hence, interacted more with the agent than did the control group. Overall, the perceived social relatedness was rather low which could indicate that learners did not immerse in the collaboration as expected. It is, however, unclear, whether this finding is specific to the simulation and the text-based realization of the agent or due to a rather low social relatedness in the simulated situation itself.

4.3. Limitations

The presented study is not without limitations that need consideration when interpreting the results. A first limitation concerns the measure of evidence elicitation that suffered from a rather low internal consistency indicating that learners showed varying degrees of skills between cases. This is a rather common problem in knowledge-based domains such as medical education (Wimmers, Splinter, Hancock, & Schmidt, 2007). The simulation was developed by researchers with many years of experience in medicine, medical education, and educational psychology, and was positively evaluated in expert workshops, a pilot study, as well as a comprehensive validation study. Therefore, we are confident that the simulation has high external validity. Further, most learners scored rather high on the pretest and very high on the posttest, indicating that the respective tests may have been too easy to differentiate well between different levels of competence. The relatively low reliability and variance could have contributed to the non-significant effect of evidence elicitation in the posttest.

Further limitations concern the implementation of the static and adaptive ISS. So far, collaboration scripts were mostly used to scaffold collaborative problem-solving or collaborative co-construction of knowledge (e.g., Rummel & Spada, 2005; Stegmann, Weinberger, & Fischer, 2007). These collaboration scripts differ from the scripts used in this study as they scaffolded rather unidirectional elicitation and sharing processes (i.e., a subset of collaboration skills) which is mainly due to the agent-based realization of the study. Although this limits the comparability of our findings to other collaboration script studies, we are convinced that through careful development of the simulation and the collaboration script, we scaffold important collaboration skills and achieve a high standardization within the learning environment.

Beyond that, it is important to consider that the implementation of the collaboration scripts affects the learners' choices minimally which could explain the lack of influence on the perceived autonomy. Thus, these results may not generalize to more coercive collaboration scripts. Nevertheless, our findings demonstrate that criticizing collaboration scripts per se as coercive (Wise & Schwarz, 2017) does not reflect the range of possibilities for implementing collaboration scripts.

Moreover, it is important to mention that we used a text-based implementation of the computer agent. This means that the collaboration could have been perceived as low immersive since learners did not see or hear their collaboration partner. Yet, an agent-based collaboration (text-based or video-based) is per se more artificial than real human-to-human collaboration since spontaneous reactions typical for human-to-human interactions are very limited in human-to-agent interactions. The ultimate question that arises from the use of a text-based agent is, thus, whether the results generalize to the real professional collaborative situation. This question ultimately requires empirical testing. This is particularly the case since we investigated only some collaborative subskills of collaborative diagnostic reasoning (i.e., sharing and elicitation), but not others (e.g., negotiating). However, because of the careful development and empirical evaluation of the agent-based simulation as well as the collaboration scripts, we consider that the simulation has a certain level of external validity (see Radkowitsch et al., 2020). Beyond that we are convinced that the agent-based simulation offers a high degree of standardization which we consider important for the thorough empirical examination in the context of basic instructional research (Graesser et al., 2018). Particularly for the learning of very specific (sub-)skills, agents can provide meaningful learning tools. Form and text-based interaction is very close to clinical reality for the skills we have identified to require training. This even makes it possible to dispense with too much resource intense face-to-face communication and provide a high degree of standardization. So far, educational and psychological research has neglected the standardization of collaborative situations to large extents.

A final limitation concerns the statistical power of the analyses. Based on prior studies that reported moderate to large effects of collaboration scripts on learning to collaborate, we conducted the a priori power analyses based on a moderate effect. The effects found in the transfer tests were rather small and the posterior power for the effect of collaboration scripts on transfer performance of evidence sharing was 61%. Probably due to the low power, the robust analyses were not significant. In contrast to the ANCOVA, the robust analyses did not include the variation of the pretest variable which is why we rely and interpret findings of the ANCOVA. Yet, future research on the combined effect of different instructional means should assume small effects for the calculation of power and replicate the findings using larger samples.

4.4. Implications and further research

In this study, we advanced collaborative diagnostic reasoning, for which individual diagnostic activities (e.g., generation of evidence or hypotheses) and collaborative activities (e.g., sharing, negotiating) are necessary (e.g., OECD, 2017), by using agent-based simulations and collaboration scripts. We focused on sharing of evidence adjusted to a partner with different knowledge background and elicitation of evidence from such a partner in order to reduce the uncertainty within the own diagnostic reasoning processes. These skills are considered two important subskills of collaborative diagnostic reasoning (e.g., Tschan et al., 2009). Our results suggest that collaboration scripts have positive effects beyond learning with simulations, and although the adaptive collaboration script was not generally better than the static collaboration script, the adaptive script helped learners to internalize the collaboration script. Besides, the adaptive collaboration script had positive effects on the perceived competence as compared to static collaboration scripts. Taking these findings together suggests that although adaptive support requires more effort during development, its effect is relevant. Still, to get a clear picture when and how adaptive collaboration scripts are effective, systematization of research is necessary (Plass & Pawar, 2020; Rummel, Walker, & Aleven, 2016). Future research should systematically vary the mechanism of adaptivity (e.g., adaptive or fading out), the bases for decision (e.g., prior knowledge or performance in the process), and the skill targeted by the scaffold (e.g., elicitation or sharing). Beyond that it seems important to consider the extent to which learners are exposed to the treatment when using adaptive support. We considered this by conducting a treatment check. An alternative approach for future research could be to examine the effect of treatment exposition on learning. Agent-based collaboration could be a promising means to provide the necessary standardization for such analyses.

We used meta-knowledge prompts to explicitly guide learners' evidence sharing and elicitation processes in the context of collaborative diagnostic reasoning. Research and theory on collaboration scripts have so far focused on how engaging in and prompting specific collaborative activities affects learning (Radkowitsch et al., 2020). The role of meta-knowledge for engaging in collaborative processes was discussed instead in group awareness research (Engelmann & Hesse, 2011). Our findings show that collaboration scripts including meta-knowledge prompts indeed affect the learning of collaboration skills. Hence, a theory about collaboration scripts (Fischer et al., 2013) should address the role of knowledge about the collaboration partners explicitly.

Furthermore, we provided further counterevidence against the criticism that collaboration scripts were prone to undermine learners' agency (Wise & Schwarz, 2017). These results are in line with the findings of a meta-analysis (Radkowitsch et al., 2020). Our results provide more detailed insights as we analyzed differentiated effects on basic psychological needs. The results of our study suggest that if the negative effects of collaboration scripts on autonomy existed at all, the effects must be minimal. In contrast, adaptive scripts enhanced the perceived competence of learners. For generalizing our results, future research should focus on replicating these effects in different contexts with different types of collaboration scripts and investigate long-term effects on basic psychological needs, intrinsic motivation, and learning. Given the evidence provided by this and previous studies (Radkowitsch et al., 2020; Stegmann et al., 2011) and the moderate to large positive effects on learning to collaborate (Radkowitsch et al., 2020), we can recommend the use of collaboration scripts for learning to collaborate.

5. Conclusion

We investigated facilitative conditions for collaborative diagnostic reasoning in a standardized agent-based simulation in undergraduate medical education by using adaptive and static collaboration scripts. To date, agents have mainly been used to assess competences (e.g., OECD, 2017) or as pedagogical agents to support specific competences of individuals or groups of learners (e.g., intelligent tutoring systems, Steenbergen-Hu & Cooper, 2014). We showed that using an agent to simulate a collaboration partner is a suitable means to effectively facilitate the learning of collaborative competences without the confounding influence of variables related to group composition (Fransen, Weinberger, & Kirschner, 2013). For complex competences such as collaborative diagnostic reasoning, such agent-based simulation can provide an economical alternative to face-to-face team training and further allow to focus on essential but specific subskills. This seems important since results of this study suggest that knowledge about the collaboration partners and their roles, tasks, and responsibilities substantially affects collaboration. Furthermore, this study shows that combining simulations with adaptive instructional support helps learners to internalize complex skills without negatively affecting learners' basic psychological needs. Yet, adaptive support is no panacea, and systematizing research on adaptive support is necessary to better understand under which conditions adaptive support enhances the learning of collaboration.

We conclude that by complementing agent-based simulations with adaptive collaboration scripts, we identified conditions to effectively help medical students learn important aspects of collaborative diagnostic reasoning.

Author note

Ethical clearance was declared by the ethics committee of LMU Klinikum, Ludwig-Maximilians-Universität, prior to data collection.

Funding information

The research presented in this contribution was funded by a grant of the Deutsche Forschungsgemeinschaft (DFG, FOR 2385) to Frank Fischer, Martin R. Fischer, and Ralf Schmidmaier (Project 6 CoSiMed, FI 792/11-1).

CRediT authorship contribution statement

Anika Radkowitsch: Writing – original draft, Conceptualization, Resources, Methodology, Formal analysis, Project administration. **Michael Sailer:** Conceptualization, Supervision, Writing – review & editing. **Ralf Schmidmaier:** Funding acquisition, Resources, Conceptualization, Writing – review & editing. **Martin R. Fischer:** Funding acquisition, Resources, Conceptualization, Writing – review & editing. **Frank Fischer:** Conceptualization, Funding acquisition, Supervision, Resources, Writing – review & editing.

Declaration of competing interest

The authors declare that they have no conflict of interest.

Acknowledgments

The authors would like to express their gratitude to Felix Behrmann, Dr. Hannah Gerstenkorn, Dr. Victoria Reitmeier, Dr. Alexander Sachs, and Dr. Hanna Zimmermann for contributing their medical expertise during the developement of the experimental material and to Felix Behrmann, Laura Brandl, and Andreas Wildner for their support before and during the collection of data.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.learninstruc.2021.101487.

References

Radkowitsch, A., Fischer, M. R., Schmidmaier, R., & Fischer, F. (2020). Learning to diagnose collaboratively: Validating a simulation for medical students. *GMS Journal of Medical Education, 37*(5). https://doi.org/10.3205/zma001344. Doc51.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A meta-analysis. *Review of Educational Research, 87*(2), 309–344. https://doi.org/10.3102/0034654316670999

Brady, A., Laoide, R.Ó., McCarthy, P., & McDermott, R. (2012). Discrepancy and error in radiology: Concepts, causes and consequences. *Ulster Medical Journal, 81*(1), 3. PMID: 23536732; PMCID: PMC3609674 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3609674/.

Cannon-Bowers, J. A., Salas, E., & Converse, S. (1993). Shared mental models in expert team decision making. In N. J. Catellan (Ed.), *Individual and group decision making: Current issues* (pp. 221–246). Hillsdale, NJ: Erlbaum.

Charlin, B., Boshuizen, H. P. A., Custers, E. J., & Feltovich, P. J. (2007). *Scripts and clinical reasoning. Medical Education* (Vol. 41). https://doi.org/10.1111/j.1365-2923.2007.02924.x

Charlin, B., Lubarsky, S., Millette, B., Crevier, F., Audétat, M. C., Charbonneau, A., et al. (2012). Clinical reasoning processes: Unravelling complexity through graphical representation. *Medical Education, 46*(5), 454–463. https://doi.org/10.1111/j.1365-2923.2012.04242.x

Chernikova, O., Heitzmann, N., Fink, M. C., Timothy, V., Seidel, T., & Fischer, F. (2020). Facilitating diagnostic competences in higher education - a meta-analysis in medical and teacher education. *Educational Psychology Review, 32*, 157–196. https://doi.org/10.1007/s10648-019-09492-2

Constantino-Gonzalez, M. A., Suthers, D. D., & de los Santos, J. G. E. (2003). Coaching web-based collaborative learning based on problem solution differences and participation. *International Journal of Artificial Intelligence in Education, 13*(2–4), 263–299.

Cook, D. A. (2014). How much evidence does it take? A cumulative meta-analysis of outcomes of simulation-based education. *Medical Education, 48*(8), 750–760. https://doi.org/10.1111/medu.12473

Davies, S., George, A., Macallister, A., Barton, H., Youssef, A., Boyle, L., et al. (2018). "It's all in the history": A service evaluation of the quality of radiological requests in acute imaging. *Radiography, 24*(3), 252–256. https://doi.org/10.1016/j.radi.2018.03.005

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York, NY: Plenum.

Deci, E. L., & Vansteenkiste, M. (2004). Self-determination theory and basic need satisfaction: Understanding human development in positive psychology. *Ricerche di Psicologia, 27*(1), 23–40.

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). Netherlands: Open Universiteit Nederland. Heerlen.

Diziol, D., Walker, E., Rummel, N., & Koedinger, K. R. (2010). Using intelligent tutor technology to implement adaptive support for student collaboration. *Educational Psychology Review, 22*(1), 89–102. https://doi.org/10.1007/s10648-009-9116-9

Engelmann, T., & Hesse, F. W. (2011). Fostering sharing of unshared knowledge by having access to the collaborators' meta-knowledge structures. *Computers in Human Behavior, 27*, 2078–2087. https://doi.org/10.1016/j.chb.2011.06.002

Feltovich, P. J., & Barrows, H. S. (1984). Issues of generality in medical problem solving. *Tutorials in problem-based learning: A new direction in teaching the health professions* (pp. 128–142). Assen, NL: Van Gorcum.

Fiore, S. M., Rosen, M. A., Smith-Jentsch, K. A., Salas, E., Letsky, M., & Warner, N. (2010). Toward an understanding of macrocognition in teams: Predicting processes in complex collaborative contexts. *Human Factors, 52*, 203–224. https://doi.org/10.1177/0018720810369807

Fischer, F., Kollar, I., Stegmann, K., & Wecker, C. (2013). Toward a script theory of guidance in computer-supported collaborative learning. *Educational Psychologist, 48*, 56–66. https://doi.org/10.1080/00461520.2012.748005

Fransen, J., Weinberger, A., & Kirschner, P. A. (2013). Team effectiveness and team development in CSCL. *Educational Psychologist, 48*(1), 9–24. https://doi.org/10.1080/00461520.2012.747947

Gegenfurtner, A., Quesada-Pallarès, C., & Knogler, M. (2014). Digital simulation-based training: A meta-analysis. *British Journal of Educational Technology, 45*(6), 1097–1114. https://doi.org/10.1111/bjet.12188

Graesser, A. C., Fiore, S. M., Greiff, S., Andrews-Todd, J., Foltz, P. W., & Hesse, F. (2018). Advancing the science of collaborative problem solving. *Psychological Science in the Public Interest, 19*(2), 59–92. https://doi.org/10.1177/1529100618808244

Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M., et al. (2019). Facilitating diagnostic competences in simulations: A conceptual framework and a

research agenda for medical and teacher education. *Frontline Learning Research, 7*(4), 1–24. https://doi.org/10.14786/flr.v7i4.384
Hesse, F., Care, E., Buder, J., Sassenberg, K., & Griffin, P. (2015). A framework for teachable collaborative problem solving skills. In P. Griffin, & E. Care (Eds.), *Assessment and teaching of 21st century skills: Methods and approach* (pp. 37–56). Springer.
Hinsz, V. B., Tindale, R. S., & Vollrath, D. A. (1997). The emerging conceptualization of groups as information processors. *Psychological Bulletin, 121*(1), 43–64. https://doi.org/10.1037/0033-2909.121.1.43
Järvelä, S., & Hadwin, A. F. (2013). New frontiers: Regulating learning in CSCL. *Educational Psychologist, 48*(1), 25–39. https://doi.org/10.1080/00461520.2012.748006
Karakostas, A., & Demetriadis, S. (2011). Enhancing collaborative learning through dynamic forms of support: The impact of an adaptive domain-specific support strategy. *Journal of Computer Assisted Learning, 27*(3), 243–258. https://doi.org/10.1111/j.1365-2729.2010.00388.x
Kiesewetter, J., Fischer, F., & Fischer, M. R. (2017). Collaborative clinical reasoning – a systematic review of empirical studies. *Journal of Continuing Education in the Health Professions, 37*(2), 123–128. https://doi.org/10.1097/ceh.0000000000000158
Kirschner, P. A., Sweller, J., & Clark, R. E. (2006). Why minimal guidance during instruction does not work: An analysis of the failure of constructivist, discovery, problem-based, experiential, and inquiry-based teaching. *Educational Psychologist, 41*(2), 75–86. https://doi.org/10.1207/s15326985ep4102_1
Kirschner, P. A., Sweller, J., Kirschner, F., & Zambrano, J. (2018). From cognitive load theory to collaborative cognitive load theoy. *International Journal of Computer-Supported Collaborative Learning, 13*, 213–233. https://doi.org/10.1007/s11412-018-9277-y
Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science, 12*, 1–48. https://doi.org/10.1207/s15516709cog1201_1
Klein, M., Otto, B., Fischer, M. R., & Stark, R. (2019). Fostering medical students' clinical reasoning by learning from errors in clinical case vignettes: Effects and conditions of additional prompting procedures to foster self-explanations. *Advances in Health Sciences Education: Theory and Practice, 24*(2), 331–351. https://doi.org/10.1007/s10459-018-09870-5
Kolodner, J. L. (1992). An introduction to case-based reasoning. *Artificial Intelligence Review, 6*(1), 3–34.
Larson, J. R., Christensen, C., Franz, T. M., & Abbot, A. S. (1998). Diagnosing groups: The pooling, management, and impact of shared and unshared case information in team-based medical decision making. *Journal of Personality and Social Psychology, 75*(1), 93–108. https://doi.org/10.1037/0022-3514.75.1.93
Liu, L., Hao, J., von Davier, A. A., Kyllonen, P., & Zapata-Rivera, D. (2015). A tough nut to crack: Measuring collaborative problem solving. In Y. Rosen, S. Ferrara, & M. Mosharraf (Eds.), *Handbook of research on computational tools for real-world skill development* (pp. 344–359). Hershey, PA: IGI Global.
Meier, A., Spada, H., & Rummel, N. (2007). A rating scheme for assessing the quality of computer-supported collaboration processes. *International Journal of Computer-Supported Collaborative Learning, 2*(1), 63–86.
Nickerson, R.-S. (1988). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology, 2*(2), 175–220. https://doi.org/10.1037/1089-2680.2.2.175
Noroozi, O., Biemans, H. J. A., Weinberger, A., Mulder, M., & Chizari, M. (2013). Scripting for construction of a transactive memory system in multidisciplinary CSCL environments. *Learning and Instruction, 25*, 1–12. https://doi.org/10.1016/j.learninstruc.2012.10.002
Nye, B. D., Graesser, A. C., & Hu, X. (2014). AutoTutor and family: A review of 17 years of natural language tutoring. *International Journal of Artificial Intelligence in Education, 24*(4), 427–469. https://doi.org/10.1007/s40593-014-0029-5
OECD. (2017). *PISA 2015 Assessment and analytical framework: Science, reading, mathematic, financial literacy and collaborative problem solving, revised edition*. Paris: France. OECD Publishing.
Patel, V. L., Kaufman, D. R., & Arocha, J. F. (2002). Emerging paradigms of cognition in medical decision-making. *Journal of Biomedical Informatics, 35*, 52–75.
Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences, 13*(3), 423–451. https://doi.org/10.1207/s15327809jls1303_6
Plass, J. L., & Pawar, S. (2020). Toward a taxonomy of adaptivity for learning. *Journal of Research on Technology in Education, 52*(3), 275–300. https://doi.org/10.1080/15391523.2020.1719943
R Core Team. (2020). R: A language and environment for statistical computing. In *R foundation for statistical computing*Vienna: Austria. https://www.R-project.org/.
Radkowitsch, A., Vogel, F., & Fischer, F. (2020). Good for learning, bad for motivation? A meta-analysis on the effects of computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 15*(1), 5–47. https://doi.org/10.1007/s11412-020-09316-4
Roschelle, J., & Teasley, S. D. (1995). The construction of shared knowledge in collaborative problem solving. In C. Oḿalley (Ed.), *Computer-supported collaborative learning* (pp. 69–97). Berlin, Germany: Springer.
Rosen, Y. (2015). Computer-based assessment of collaborative problem solving: Exploring the feasibility of human-to-agent approach. *International Journal of Artificial Intelligence in Education, 25*(3), 380–406. https://doi.org/10.1007/s40593-015-0042-3
Rummel, N., & Spada, H. (2005). Learning to collaborate: An instructional approach to promoting collaborative problem solving in computer-mediated settings. *The Journal of the Learning Sciences, 14*, 201–241. https://doi.org/10.1207/s15327809jls1402_2
Rummel, N., Walker, E., & Aleven, V. (2016). Different futures of adaptive collaborative learning support. *International Journal of Artificial Intelligence in Education, 26*(2), 784–795. https://doi.org/10.1007/s40593-016-0102-3
Sailer, M., Hense, J. U., Mayr, S. K., & Mandl, H. (2017). How gamification motivates: An experimental study of the effects of specific game design elements on psychological need satisfaction. *Computers in Human Behavior, 69*, 371–380.
Schwaighofer, M., Vogel, F., Kollar, I., Ufer, S., Strohmaier, A., Terwedow, I., et al. (2017). How to combine collaboration scripts and heuristic worked examples to foster mathematical argumentation – when working memory matters. *International Journal of Computer-Supported Collaborative Learning, 12*(3), 281–305. https://doi.org/10.1007/s11412-017-9260-z
Siebeck, M., Schwald, B., Frey, C., Röding, S., Stegmann, K., & Fischer, F. (2011). Teaching the rectal examination with simulations: Effects on knowledge acquisition and inhibition. *Medical Education, 45*, 1025–1031. https://doi.org/10.1111/j.1365-2923.2011.04005.x
Simmons, B. (2010). Clinical reasoning: Concept analysis. *Journal of Advanced Nursing, 66*(5), 1151–1158. https://doi.org/10.1111/j.1365-2648.2010.05262.x
Stadler, M., Herborn, K., Mustafić, M., & Greiff, S. (2019). Computer-based collaborative problem solving in PISA 2015 and the role of personality. *Journal of Intelligence, 7*(3), 15. https://doi.org/10.3390/jintelligence7030015
Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: Two experimental studies in undergraduate medical education. *Learning and Instruction, 21*(1), 22–33. https://doi.org/10.1016/j.learninstruc.2009.10.001
Steenbergen-Hu, S., & Cooper, H. (2014). A meta-analysis of the effectiveness of intelligent tutoring systems on college students' academic learning. *Journal of Educational Psychology, 106*(2), 331. https://doi.org/10.1037/a0034752
Stegmann, K., Mu, J., Gehlen-Baum, V., & Fischer, F. (2011). The myth of over-scripting: Can novices be supported too much?. In H. Spada, G. Stahl, N. Miyake, & N. Law (Eds.), *Connecting computer-supported collaborative learning to policy and practice: CSCL2011 conference proceedings. Volume I. Hong-kong, China: International society of the learning sciences*.
Stegmann, K., Weinberger, A., & Fischer, F. (2007). Facilitating argumentative knowledge construction with computer-supported collaboration scripts. *International Journal of Computer-Supported Collaborative Learning, 2*, 421–447. https://doi.org/10.1007/s11412-007-9028-y
Sun, C., Shute, V. J., Stewart, A., Yonehiro, J., Duran, N., & D'Mello, S. (2020). Towards a generalized competency model of collaborative problem solving. *Computers & Education, 143*. https://doi.org/10.1016/j.compedu.2019.103672
Sweller, J., van Merriënboer, J. J., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review, 31*, 261–292. https://doi.org/10.1007/s10648-019-09465-5
Tchounikine, P., Rummel, N., & McLaren, B. M. (2010). Computer supported collaborative learning and intelligent tutoring systems. In R. Nkambou, J. Bourdeau, & R. Mizoguchi (Eds.), *Advances in intelligent tutoring systems* (pp. 447–463). Berlin, Heidelberg: Springer. Berlin Heidelberg.
Tschan, F., Semmer, N. K., Gurtner, A., Bizzari, L., Spychiger, M., Breuer, M., et al. (2009). Explicit reasoning, confirmation bias, and illusory transactive memory: A simulation study of group medical decision making. *Small Group Research, 40*(3), 271–300. https://doi.org/10.1177/1046496409332928
Tsovaltzi, D., Rummel, N., McLaren, B. M., Pinkwart, N., Scheuer, O., Harrer, A., et al. (2010). Extending a virtual chemistry laboratory with a collaboration script to promote conceptual learning. *International Journal of Technology Enhanced Learning, 2*(1–2), 91–110. https://doi.org/10.1504/IJTEL
Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
Wecker, C., & Fischer, F. (2011). From guided to self-regulated performance of domain-general skills: The role of peer monitoring during the fading of instructional scripts. *Learning and Instruction, 21*(6), 746–756. https://doi.org/10.1016/j.learninstruc.2011.05.001
Wegner, D. M. (1987). Transactive memory: A contemporary analysis of the group mind. In B. Mullen, & G. R. Goethals (Eds.), *Theories of group behavior* (pp. 185–208). New York, NY: Springer.
Wimmers, P. F., Splinter, T. A., Hancock, G. R., & Schmidt, H. G. (2007). Clinical competence: General ability or case-specific? *Advances in Health Sciences Education, 12*(3), 299–314. https://doi.org/10.1007/s10459-006-9002-x
Wise, A. F., & Schwarz, B. B. (2017). Visions of CSCL: Eight provocations for the future of the field. *International Journal of Computer-Supported Collaborative Learning, 12*(4), 423–467. https://doi.org/10.1007/s11412-017-9267-5
Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x