

Perceived authenticity across three forms of educational simulations—the role of interactant representation, task alignment, and continuity of simulation

Caroline Corves^{1,2} · Matthias Stadler¹ · Martin R. Fischer¹

Received: 28 August 2023 / Revised: 28 February 2024 / Accepted: 29 February 2024 / Published online: 16 April 2024 © The Author(s) 2024

Abstract

Authenticity in simulation-based learning is linked to cognitive processes implicated in learning. However, evidence on authenticity across formats is insufficient. We compared three case-based settings and investigated the effect of discontinuity in simulation on perceived authenticity. In a quasi-experiment, we compared formats of simulation in the context of medical education. All formats simulated anamnestic interviews with varying interactant and task representations using highly comparable designs. Interactants (patients) were simulated by (a) live actors (standardized patients), (b) live fellow students (roleplays), or (c) question menus and videoclips (virtual patients). The continuity of simulations varied. We measured perceived authenticity with three subscales: Realness, Involvement, and Spatial Presence. We employed confirmatory factor analysis (CFA) to assess measurement invariance across settings and analysis of variance on authenticity ratings to compare the effects of setting and discontinuous simulation. CFA supported the assumption of invariance. Settings differed in Realness and Spatial Presence but not Involvement. Discontinuous simulations yielded significantly lower ratings of authenticity than continuous simulations. The compared simulation modalities offer different advantages with respect to their perceived authenticity profiles. Lower levels of interactivity and reduced subtask representation do not necessarily lead to lower ratings of perceived authenticity. Spatial Presence can be as high for media-based simulation as for roleplays. Discontinuation of simulations by offering scaffolding impairs perceived authenticity. Scaffolds may be designed to avoid discontinuation of simulation to uphold perceived authenticity.

Keywords Simulation-based learning \cdot Authenticity \cdot Standardized patients \cdot Virtual patients \cdot Roleplay \cdot Scaffolding

Simulation-based learning has been shown to facilitate complex skill development in higher education (Chernikova et al., 2020) and is an established practice in the fields of science, technology, engineering, and mathematics (STEM) (d'Angelo et al., 2014) and medicine (Cook et al., 2013; Hegland et al., 2017). Simulations represent aspects of the real world in an interactive manner that allows learners to be immersed in the learning

Extended author information available on the last page of the article

environment (Gaba, 2004; Jeffries, 2012). Simulations represent relevant aspects of reallife problems and professional practice (Grossman, 2021), offer opportunities for decisionand action-based learning (Kaufman & Ireland, 2016), and become indispensable when stakes are high and training becomes unethical in real-life situations (Friedrich, 2002; Salas et al., 2009). Balancing the authenticity of simulations with other processes implicated in learning poses a challenge for their design (Codreanu et al., 2020; Stadler et al., 2021). In this study, we aim to compare three formats of simulation in medical education associated with varying design characteristics. We derive predictions for perceived authenticity in three forms of educational simulations in the context of medical education and compare these forms across different facets of perceived authenticity. Moreover, we investigate the effect of discontinuous simulation through intermittent scaffolding on perceived authenticity.

Representational reduction in simulation-based learning

Simulation is an active educational method that consists of replacing real experiences with guided experiences that replicate aspects of the real world in an interactive manner (Gaba, 2004). As simulations *model* real-life situations, they are reductive in nature, thus offering potential benefits of lowering complexity on the basis of pragmatism (Stachowiak, 1973).

In this latter context, correspondence between the *model* (simulation) and *target system* (real-life situation) involves which aspects are represented (and how), along with the resulting detail or granularity of the model (Maier et al., 2017). Detail is not a value per se, as adding more detail to a model does not necessarily result in a higher quality of that model: "As a model grows more realistic, it also becomes just as difficult to understand as the real-world processes it represents" (Dutton & Starbuck, 1971). By leaving out some aspects, the ones represented become more prominent, thus facilitating their processing or handling. This provides opportunities to target learners' experience by simulation design to facilitate knowledge or skill acquisition.

Cognitive architecture poses limits and implicates effective design choices

The availability of cognitive resources is limited (Peterson & Peterson, 1959; Miller, 1956) and more so in novel environments, as representation in working memory depends on prior knowledge (Ericsson & Kintsch, 1995; Kirschner et al., 2006). This point has been used to argue for direct instructional guidance and against minimal guidance approaches like discovery-based learning (Kirschner et al., 2006).

Independent of instructional guidance, representational reduction may be seen as another way to respond to limited resources: Learners may benefit from *representational reduction* by allocating attentional resources to the remaining aspects, thus enhancing cognitive processing via perceptual organization and incorporating information into mental representations as a basis for learning to complement direct instructional support. This may include reductions in sensory stimuli to facilitate perceptual organization and object recognition, in represented objects competing for attentional allocation, in professionally relevant or non-relevant subtasks, and in a complex decision space. Task complexity has been shown to complicate memory retrieval (Anderson et al., 1996). Reductive design may thus systematically increase the potential for mastery by making the simulated environment manageable, keeping learners in their zone of proximal development (Vygotsky, 1978), providing meaning to success and failure, and accounting for motivational factors. The benefits of reduction may depend on learners' characteristics, including their knowledge and prior experience with the simulated context. Simulation can thus provide inexperienced learners with opportunities to practice for challenging professional situations (Groot et al., 2020). Positive effects of reducing complexity have been demonstrated in simulation-based learning (Chernikova et al., 2020).

Simulation is based on an active and ongoing process of mental construction in the learner

A participant's experience relies on an active and ongoing process of constructing and maintaining mental representations on the basis of provided stimuli. These stimuli may be consistent with or may conflict with participants' mental representations of the simulated situation, either facilitating or complicating consistent mental representations. *We conceptualize simulations as experienced-based processes that are based on the settings' facilitation of learners' ongoing mental representations of target situations*. Educational simulation involves modeling a relevant complex target situation via selective representation in a way designed to facilitate perceptual processing, attentional allocation, action selection, and transferable knowledge formation. In accordance with previous conceptions, simulations can be technically supported or unsupported (Chernikova et al., 2020; Davidsson & Verhagen, 2017; Heitzmann et al., 2019).

Thus, representational reduction can be seen as an elementary potential benefit of simulation, to provide learning opportunities in response to constraints of processing inherent to our cognitive architecture, but may compromise the alignment of the model and target situation.

Fidelity and authenticity in simulation-based learning

The quality of simulations or simulators in representing the real world is usually referred to as fidelity. Earlier, fidelity was usually assessed on the basis of a simulation's technology or physical resemblance to the modeled situation, but this approach was criticized for its overemphasis on engineering attributes and based on a lack of implications for learning (Bland et al., 2014). Hamstra et al. (2014) suggested prioritizing concepts of functional task alignment (i.e., replication of the demands of a real clinical task) over physical resemblance and avoiding the term fidelity. Instead, the term authenticity was suggested to describe the fidelity experienced by the learner (Lavoie et al., 2020). Authenticity is often considered conceptually distinct from fidelity and is framed on the basis of experience in a simulation: Authenticity has been referred to as learners' interpretation of the veracity of a situation in which they interact with a context, other learners, and a simulator (Bland et al., 2014). Learners' judgments of what is credible vary and depend at least partially on prior experience. Increased fidelity might thus not necessarily result in increased authenticity, and low-fidelity simulations may feel authentic (Bland et al., 2014; Grossman et al., 2015). Chernikova et al.'s (2020) meta-analysis on simulation-based learning found stronger effects on learning for increasing degrees of authenticity when assessing the focus and scope of correspondence between simulations and real-life situations.

We take perceived authenticity to indicate experience that involves representational reductions in alignment with learners' expectations of the target situation's relevant aspects and the aptness of the setting to evoke and facilitate mental representations of the simulated environment, such that the learner feels present and engaged in this environment.

The implication is that perceived authenticity is potentially compromised by conditions that compromise perceived alignment and compromise allocation of attention to the simulated environment or stimuli that are inconsistent with mental representations of the simulated environment.

As a form of experience-based learning, simulation provides opportunities for the formation of episodic memory. The theory on episodic memory formation and retrieval states that memories can be retrieved through mental search or by cue-prompted automatic activation in recall and that internal and external retrieval cues present at the time of recall determine the accessibility of memory engrams (Frankland et al., 2019; Tulving & Pearlstone, 1966). Different authors have proposed its directive function on future action which includes tasks like problem-solving and decision-making (Bar, 2009; Pillemer, 2003). When the cues of the recall environment resemble those in the learning environment, this facilitates memory activation and recall (Godden & Baddeley, 1975; Tulving & Thomson, 1973). We consider this relevant in the context of simulation-based learning and authenticity, because the simulation's alignment with the future target situation affects the encoding and retrieval of long-term memory. In addition to providing opportunities to practice relevant skills, design may target optimization for later retrieval and guidance of action. Perceived authenticity might thus to a certain extent reflect a factor important for the formation of transferable memory.

Perceived authenticity as a multidimensional construct

A factor-analytic study (Schubert et al., 2001) originally designed to distinguish the constituents of the experience of presence suggested three elements of perceived authenticity: Realness, Involvement, and Spatial Presence. Involvement and Spatial Presence were identified as two separate factors of the experience of presence. Additionally, the authors found a third subjective factor: judgments of Realness. Lessiter et al. (2001) found three corresponding factors as separate constituents of participants' experiences in interactive or non-interactive simulations.

Realness describes the degree to which a person believes that a situation and its characteristics resemble a real-life context (Schubert et al., 2001). Involvement, on the other hand, is defined as a feeling of cognitive immersion and the judgment that a situation has personal relevance (Hofer, 2016). In contrast to cognitive immersion, Spatial Presence denotes the feeling of *physical* immersion in a situation. Separating these constructs has implications for the assessment and design of simulation-based learning. The authors of the study inferred the question of which characteristics of a given simulation environment affect which factors of participants' experiences and judgments (Schubert et al., 2001). Schubert et al. (2000) reported that providing *both real and illusory ways to interact* with a virtual environment through bodily movement enhances Spatial Presence but only marginally enhances Involvement and Realness.

Perceived opportunities for action increase the sense of being physically immersed in a simulated situation (Schubert et al., 2000). Thus, the connections that learners' experience in simulations reflect the simulations' aptness for suggesting action opportunities and may offer a way to measure learners' mental representations of simulated environments in terms

of action orientation. As such, these connections provide benefits for transferable memory formation to provide orientation in the target situation.

Formats of simulation-based learning

Simulation-based learning comprises formats with varying characteristics as a common part of medical curricula, where simulation-based learning is well established (Cook et al., 2013): standardized patients (SPs), roleplays (RPs), and virtual patients (VPs). SPs are actors or volunteers trained to act as if they are displaying symptoms of a disease (Vu & Barrows, 1994). Such displays may include non-symptom-based human behavior to teach students how to avoid common communication fallacies. RPs are used in small or larger groups where students usually prepare for their part with a script (Joyner & Young, 2006) and play the role of doctor, patient, or observer. Although the distinction between an *active* (*doctor, patient*) and a *passive (observer) role* has been made, observers are engaged and benefit from actively assessing the situation, especially by means of instructional support, such as an observational script or protocol (Stegmann et al., 2012). Hence, we suggest the terms *interactive* (doctor, patient) and *observational*.

A study on second-year medical students reported that students found SP feedback to be more instructive, whereas the interaction itself was judged to be more engaging due to its heightened emotional intensity (Taylor et al., 2019).

On the other hand, RPs require less organizational or financial effort (Taylor et al., 2019; Weller et al., 2012), and only RPs promote the change in perspective and empathy (Bosse et al., 2012).

The term VP has been used to denote different forms of simulations, including human SPs, high-fidelity manikins, and virtual SPs, with the term virtual being used to describe the simulative aspect (Kononowicz et al., 2015). In our context, we define VPs as computer-based simulations in which the representing model can be manipulated by the participant (de Jong, 2010; Berman et al., 2016) with varying means of patient representation (avatars, videos of SPs) and degrees of interaction (Kononowicz et al., 2015; Villaume et al., 2006). Production costs may be high (Heitzmann et al., 2019), but the method's scalability and permanence may offset the expenditure (Cook & Triola, 2009; Jennebach et al., 2022). Moreover, VPs offer the advantage of a possible maximally standardized case presentation (Berman et al., 2016). A recent study compared the effect of RPs and web-based task simulations on knowledge gain and found a benefit of RPs but not of web-based task simulations (McAlpin et al., 2023) and concluded that RPs may be a better environment for practicing the integration of complex skills.

SPs, RPs, and VPs entail different characteristics potentially implicated in perceived authenticity

SP-based simulations include an increased opportunity for action as well as detailed and highly aligned representations of the interactant (the patient): trained impersonation offering coherent behavior, including situation-specific behavior, emotional displays, and high responsiveness to participants' actions. Additionally, SPs can represent the professionally relevant subtasks needed to obtain relevant information from the patient: free recall, choice, effective phrasing of questions, and flexible communication. Moreover, they offer continuous simulation of the target situation. Correspondingly, previous research has reported high authenticity for SPs (Barrows, 1993).

Similar to SPs, RPs offer similar opportunities for action, model professional and situation-specific task requirements, and a continuous experience while arguably offering a less believable and coherent object representation (patient) on average for the learner taking part as a doctor, potentially reducing perceived authenticity. Concerning the representation of humans in particular, observing the actions of others may include action simulation and prediction (Brass et al., 2000; Burke et al., 2010; Graf et al., 2007; Wilson & Knoblich, 2005), error monitoring, and the prediction of higher cognition (Schuch & Tipper, 2007; van Schie et al., 2004) as well as covert mimicking of emotion (Hatfield et al., 1993). As other cognitive and affective states are part of people's representation (Apperly et al., 2009; Premack & Woodruff, 1978; Schlaffke et al., 2015), inconsistent behavioral display including emotion may interfere with participants' prediction and provide conflicting stimuli.

VPs usually offer less opportunity for action or interactivity, including reduced subtask representation, such as preselection of anamnestic questions, no direct relationship between the learner and patient in space, and discontinuous interactant representation, but in contrast to RPs, may offer a more believable representation of this interactant during its overt presentation (patient videoclips). In contrast to RPs and SPs, VPs also limit the focus on a selection of the space containing the patient, which might offer an advantage if the space of the live simulation setting does not closely resemble the environment of the target situation.

In terms of situation-, task-, and interactant-specific representation and immersive characteristics, constraints and opportunities differ across these three forms of simulations. A selective summary of these differences can be found in Table 1. Given these differences in opportunities to simulate relevant aspects of clinical tasks, their differences in interactivity, and their representation of the object of interaction (patient), these settings are expected to vary in their ability to induce the experience of authenticity. Although the role of learners' experience in the simulation has been shown to be relevant for learning processes, comparisons of these forms of simulation are lacking. Potential variations in VP design are practically limitless, yet few studies have rigorously explored design issues (Cook & Triola, 2009). Direct comparisons of perceived authenticity between SPs, RPs, and VPs as three different forms of simulation are lacking. However, such comparisons provide insights for medical curricula designers who need to orchestrate available possibilities to provide learning opportunities that approximate professional practice (Grossman et al., 2009), thus preparing learners for situations they are likely to encounter in their professional practices (Berman et al., 2016).

Mental representation, conflicting stimuli, and effects of scaffolding

Representational reduction in simulations can be complemented by additionally enhancing salience or more explicit instructional guidance (e.g., prompting). Incorporating a way to facilitate learning into a simulation may aid learners' knowledge or competence acquisition (Heitzmann et al., 2019) but may reduce the correspondence between a simulation and the modeled target situation and affect judgment of Realness.

According to Schubert et al. (2001), inconsistent stimuli warrant suppression for the maintenance of coherent mental representation of the situational experience. Temporary structured support (scaffolding, Wood et al., 1976) may entail stimuli conflicting with learners' representation.

	Catting A: stondordized notiants (CDs)	Satting B. rolanlow (DDc)	Satting C. vintual notiante (V/De)
	ocume A. Standardized patients (of s)	ocuming D. HOUCPHAYS (NES)	Schning C. Virtual patients (VIS)
Patient embodiment	Actors trained for their roles in a separate session, representing situation-specific behavior and emotional display	Fellow students received prior information directly before the simulation. They were told that the focus was on providing correct infor- mation, whereas acting was not required	Trained actors delivered scripted answers (vide- oclips), including representations of situation- specific behavior and emotional displays
Interaction	Free dialogue	Free dialogue between doctor and patient	Interactive learning platform by means of question menus and connected videoclips
Situational setting	One-on-one live setting	Live in mostly triadic groups (doctor, patient, observer) or including a second observer	Online learning platform accessed from university facilities or from home



Scaffolding is often associated with discontinuous cuing of simulated environments. When it is associated with pausing the simulation of an ongoing process—while not being perceived as a property of the simulated target environment itself—the resulting situation can be interpreted as being in conflict with the mental representation of a target situation, hindering the experience of actually being in that situation. Interruptions by stimuli or subtasks that are inconsistent with the simulation of the target situation may thus require mental reconstruction and stimulus suppression and might reduce learners' perceived authenticity. However, research investigating this connection in our context is sparse.

Study aims, research questions, and hypotheses

We aimed to compare perceived authenticity across (a) three forms of simulation with varying characteristics in higher education, SPs, RPs, and VPs, and (b) conditions of continuous and discontinuous simulations for its implications for simulation design and scaffolding.

Representational variation across simulation settings in medical education perceived authenticity profiles across SPs, RPs, and VPs

The three common settings of SPs, RPs, and VPs vary in characteristics that potentially contribute to perceived authenticity as described above, suggesting differences in authenticity ratings on the subscale level. As SPs may offer high alignment (concerning task and interactant), action opportunities, interactivity, and continuous simulation, we expected perceived authenticity to be high on all subscales and Realness to be highest for SPs. We expected a difference in Realness between RPs and VPs. The direction of this difference may depend on whether Realness judgments depend more on simulated patients' features (e.g., emotional displays, situation-specific behavior) that are represented in the VP videoclips or on subtask representation, which is higher for RPs than for VPs. We further expected spatial presence to be higher in RPs than VPs, as the live setting of RPs involves more choices for taking action and a direct physical relationship with the simulated patient and movement in space. On the other hand, VPs may be better able to facilitate the construction of a mental representation that entails a professional setting. Here, mental representations are based on cues provided through virtual stimuli that are distinct from the actual setting and thus may be associated with less conflicting stimuli about the actual location of a participant.

Thus, we hypothesized:

H1.0: Perceived authenticity (i.e., Realness, Involvement, and Spatial Presence) will differ across simulation settings (i.e., SPs, RPs, VPs).

H1.1: Realness, Involvement, and Spatial Presence will differ unequally across SPs, RPs, and VPs.

H1.2: Realness will be highest for SPs.

H1.3: Realness will differ between VPs and RPs.

H1.4: Spatial Presence will be higher for SPs than for VPs.

H1.5: Spatial Presence will differ between VPs and RPs.

Continuity of simulation—the effect of discontinuous simulation through scaffolding

We suggest that scaffolding could affect the experience of being present in the modeled situation by presenting stimuli that conflict with learners' representations, thus warranting suppression so that learners can maintain coherent mental representations of their situational experience.

We propose that intermittent scaffolding through the induction of reflection should effectively result in a discontinuous experience of the simulated interview and the need for reallocation of attentional resources and reconstruction of the simulated environment, thus affecting spatial presence. Moreover, the prompting of reflection might not be considered to align with situational properties (i.e., participants might not conceive of the task as being part of the task in the target situation), thus affecting ratings of Realness. We hypothesized:

H2.0: Perceived authenticity will be lower in conditions of discontinuous simulation than in conditions of continuous simulation.

H2.1: Specifically, there will be a decrease in authenticity ratings for discontinuous simulation in terms of Spatial Presence.

Methods

Participants

The sample consisted of 568 participants who took part in one of three settings simulating an anamnestic interview situation, 499 of whom had complete information, including role and scaffolding condition: SP simulation (n=81), RP simulation (n=298; doctor=96, patient=101, observer=101), or the VP simulation (n=120). All participants were German medical students at LMU Munich in years 3–6 of a 6-year program. Good command of German was a prerequisite for participation. Participation in the study was voluntary, and students received course credit or monetary compensation for participating. All participants were over 18 years of age, and all gave informed consent.

Procedure and materials

Data were collected in the context of three original experimental studies, each examining diagnostic accuracy and potentially supportive interventions through scaffolding in one of three simulation-based learning scenarios (case-based learning) from 2018 to 2022. Hence, participants' distribution across studies—and thus setting—was quasiexperimental, while their allocation to conditions within each study was randomized.

Setting A (SPs) Patients were portrayed by trained actors, selected in accordance with the patient-case characteristics, and coached by a physician in separate sessions. They were instructed to abstain from introducing medical information other than what was provided.

In a live one-on-one setting, participants conducted a free interview within a given time frame. Details can be found in (Fink et al., 2021a).

Setting B (RPs) In a triadic live setting, fellow students entered together and were then randomly assigned participatory roles (doctor, patient, observer). Patients were thus portrayed by fellow students, who received information directly before the simulation. They were told that acting was not necessary and the focus was on correctly delivering the provided medical information. Some information was organized into more general information for answering loosely formulated questions and details to be delivered only when a question directly targeting these specific details was asked. They were instructed to provide only the given information, and when they lacked the answer to a posed question, to indicate the missing information by responding with choices such as "I don't know," "I'm not sure," or "I don't remember." Participants then conducted a free anamnestic interview within a given time frame. The simulated patient was situated on a bed, usually sitting, in a room also containing students' desks and computers. (Participants in the observer role received a concise description and a protocol for observing the performance of the simulated doctor.)

Setting C (VPs) Anamnestic interviews were simulated via an interactive learning platform by means of question menus and a short videoclip prompted for each of the selected anamnestic questions. In each videoclip, participants heard a female voice repeating the selected question and then the patient answering the question. Patients were portrayed by professional actors who used scripted answers and were given directions for their nonverbal and paraverbal presentation in short videoclips. The camera setting was set to simulate a direct face-to-face interaction, filming the actor in medium close-up (MCU) directly looking into the camera. Moreover, a clinical setting was simulated (patient in or on a bed) and clearly visible in the setting. Anamnestic questions were visually clustered into categories. Videos containing the answers of the simulated patients were assessed by clicking on the respective question. Details can be found in (Fink et al., 2021b).

All studies were approved by the ethics committee of the Medical Faculty of LMU Munich. The comparability of simulations across the three settings was designed to be maximized. Across all settings, participants were given the same cases of patients presenting with dyspnea (case validation in (Fink et al., 2021a)) and asked to deliver a diagnosis after a simulated anamnestic interview. All settings contained briefings, including familiarization with content and technical aspects beforehand. All simulated interviews had a time limit and lasted 8 to 13 min. In all settings, participants filled out a questionnaire on perceived authenticity implemented in the online learning platform CASUS (Instruct, 2021) after they delivered their diagnosis of the simulated patient. For two (RPs, SPs) of the three settings, simulations included experimental variation of scaffolding via induced reflection on the diagnostic process (Fink et al., 2021b) adapted from Mamede et al., 2008, 2012, 2014) that was (a) absent, (b) prompted after the simulation (continuous simulation), or (c) implemented halfway through the interview (discontinuous simulation).

Measurement of perceived authenticity

We assessed the perceived authenticity of the experienced simulation in three dimensions with three subscales taken from validated questionnaires (Schubert et al., 2001; Seidel et al., 2011; Vorderer et al., 2004; Frank & Frank, 2015). All contained 5-point Likert-scaled items ranging from 1 (*disagree*) to 5 (*agree*):

Realness (3 items): "I consider the history-taking simulation as authentic.", "The simulation of the medical interview seemed like a real professional demand.", "The experience in the history-taking simulation resembled the experience of a real professional demand."

Involvement (4 items): "When I participated in history-taking, I focused strongly on the situation". "When I participated in history-taking, I intermittently forgot that I was taking part in a study", "When I participated in history-taking, I was immersed in the situation", "When I participated in history-taking, I was fully engaged."

Spatial Presence (3 items): "When I participated in history-taking, it seemed to me as if I was a real part of the simulated situation.", "When I participated in history-taking, I felt like I was physically present in the clinical environment.", "When I participated in history-taking, it seemed to me as if I could affect things - like in a real medical interview."

Presentation of the scales for each of the factors of perceived authenticity occurred consecutively. Assessment of internal consistency based on McDonald's ω indicated good internal consistency for each subscale: ω (Realness)=0.91, ω (Involvement)=0.82, ω (Spatial Presence)=0.81. As a basis for the ANOVAs, a mean score was calculated for each subscale.

Analysis

All analyses were performed with RStudio (RStudio Team, 2022) and JAMOVI Version 2.3.18 (The Jamovi project, 2022). All post hoc analyses were based on Tukey HSD to account for multiple comparisons. Alpha was set to 5% for all analyses.

Measurement invariance of construct across simulation settings

We employed confirmatory factor analysis to assess the measurement invariance of our hypothetical construct across the three simulation settings using lavaan (v 0.6.12, Rosseel, 2012). Model fit was evaluated with standard model fit indices, such as the confirmatory fit index (CFI) and χ^2 values. CFI values above 0.90 were considered to represent an acceptable fit and values above 0.95 good fit. Root mean square error of approximation (RMSEA) was used as a complementary fit index with values below 0.06 considered to represent a good fit. Measurement models were defined for all three constructs (Realness, Involvement, Spatial Presence) with both factor loadings and intercepts constrained to be equivalent across groups, thus testing for scalar invariance of the construct across simulation settings. We then compared the resulting fit with that of less constrained models (metric model, configural model) with factor loadings and item intercepts free to vary for each group to test for a significant increase in fit for the less constrained model and therefore the rejection of the scalar invariance assumption.

Perceived authenticity across settings

To compare authenticity across simulation settings, we employed a repeated-measures ANOVA, with perceived authenticity as a within-subjects factor with three levels (Realness, Involvement, and Spatial Presence) and simulation setting as a between-subjects factor with three levels (SPs, RPs, VPs), the dependent variable being the mean of each subscale. For comparability, observations of RP authenticity ratings were restricted to ratings from participants playing the role of the doctor.

Continuous versus discontinuous interview simulation

For the continuity of the simulation construct, we distinguished between conditions (RPs and VPs) where the simulated interview was interrupted for another activity (prompted reflection, *discontinuous*) and conditions where the simulation of the interview was uninterrupted (prompted reflection either followed the interview or was absent, *continuous*).

A repeated-measures ANOVA was computed to compare the effect of continuous versus discontinuous simulation on authenticity perception on observations of RPs and VPs with subscale as a within-subjects factor and continuity and setting as between-subjects factors.

Results

Measurement invariance of constructs across simulation settings

The results of our CFA supported the assumption of invariant construct measurement across the three simulation settings. Measurement models with scalar invariance could be established for all three scales: The fit of the scalar model was excellent, $\chi^2(124) = 86.00$, CFI=1.00, p=1.000, and the difference from the less constrained metric model was non-significant ($\Delta \chi^2 = 21.69$, Δ CFI=0.00, p=0.085). Hence, the items contributed to their respective latent construct to a similar degree across groups, and mean differences in the latent construct captured all mean differences in the shared variance of the items. These results (see Table 2 for a summary) validated the comparison of relationships and means within the constructs. There were substantial latent correlations between Realness and Involvement (r=0.371, p<0.001), Realness and Spatial Presence (r=0.883, p<0.001), and Involvement and Spatial Presence (r=0.420, p<0.001).

Invariance model	$\chi^2 (df)$	$\Delta \chi^2$	р	CFI	ΔCFI	RMSEA	ΔRMSEA	Decision
Configural	46.42 (96)			1.00		0.00		Accepted
Metric	64.31 (110)	17.89	0.212	1.00	0.00	0.00	0.00	Accepted
Scalar	86.00 (124)	21.69	0.085	1.00	0.00	0.00	0.00	Accepted

Table 2 Analysis of measurement invariance

CFI comparative fit index, RMSEA root mean square error of approximation



Fig. 1 Setting profiles in terms of Realness, Involvement, and Spatial Presence. Note. RPs included only participants who played the role of doctor

Perceived authenticity across settings

A repeated-measures ANOVA was computed to evaluate the effect of setting on perceived authenticity. The means and standard deviations for perceived authenticity are presented in Fig. 1.

An ANOVA comparing perceived authenticity on the three subscales across settings yielded significant main effects and a significant interaction. There was a significant medium-sized effect of subscale, F(2, 588) = 56.0, p < 0.001, $\eta^2 = 0.048$, and a significant medium-to-large-sized main effect of setting on authenticity perception, F(2, 297) = 27.2, p < 0.001, $\eta^2 = 0.10$. The interaction between setting and subscale was significant, F(4, 588) = 23.5, p < 0.001, $\eta^2 = 0.04$, indicating a small effect of setting on variation in authenticity ratings between subscales. As the interaction was disordinal, we abstained from interpreting the main effects. Post hoc pairwise comparisons of perceived authenticity between settings with a Tukey HSD adjustment indicated that perceived authenticity ratings in SP simulations were significantly higher than in RP simulations (p < 0.001) or VP simulations (p < 0.001). There was no significant difference in ratings of perceived authenticity between VP simulations and RP simulations (p = 0.095).

Post hoc pairwise comparisons of perceived authenticity between the subscales Realness, Involvement, and Spatial Presence with a Tukey HSD adjustment indicated that perceived authenticity ratings in SP simulations differed significantly between the three measures (all ps < 0.001).

Realness was highest for SPs and lowest for RPs. The differences between each of the settings were significant: SPs versus RPs (p < 0.001), SPs versus VPs (p = 0.006), RPs versus VPs (p < 0.001).

None of the settings differ significantly in Involvement.

Spatial Presence was significantly higher for SPs than RPs (p < 0.001) or VPs (p < 0.001). Ratings of perceived authenticity on RP and VP simulations did not differ significantly (p = 1.000).

Continuous versus discontinuous interview simulation

An ANOVA comparing perceived authenticity across continuous versus discontinuous simulations of the RP and VP simulations yielded the following results (see also Fig. 2): There was a significant main effect of continuity, F(1, 414)=4.179, p=0.042, $\eta^2=0.003$, indicating differences in authenticity ratings between continuous and discontinuous simulation conditions for observations of RPs and VPs.

The interactions between setting and continuity, F(1, 414) = 0.102, p = 0.750, and subscale and continuity, F(2, 828) = 1.418, p = 0.243, were nonsignificant.

Discussion

The present study was aimed at investigating differences in perceived authenticity across three forms of simulations. We were able to compare ratings of perceived authenticity under highly comparable conditions. We thus provide details on the respective profiles of these three settings in terms of Realness, Involvement, and Spatial Presence as facets of perceived authenticity.

Perceived authenticity differs between settings

The results of our comparison across SPs, RPs, and VPs confirmed our hypothesis that perceived authenticity differs across simulation settings and that different factors of perceived



Fig. 2 Effect of continuity of simulation

authenticity are unequally affected by their differences in how they modeled different aspects of the target situation. Our findings also demonstrate that assessing these factors individually may provide additional information about simulation settings in general.

In line with our hypothesis and previous research reporting high authenticity for SPs (e.g., Barrows, 1993) and results previously reported by (Fink et al., 2021b), we found that SPs were rated highest in Realness and Spatial Presence. The finding that Realness was higher for SPs than VPs or RPs was consistent with our assessment that SPs offer a detailed representation of the object (patient) as well as of clinically relevant subtasks (e.g., for evidence generation). As expected, Spatial Presence was higher for SPs as a live setting with high opportunities for action in comparison with VPs.

RPs were rated lower in Spatial Presence than SPs, and in contrast to our hypothesis, we found no difference in Spatial Presence between VPs and RPs, although these settings differ in terms of characteristics that potentially contribute to Spatial Presence: RPs offered more opportunity for action and entailing a direct spatial relationship with the simulated patient. VPs were assessed as offering easier separation between environmental cues that were consistent or inconsistent with the mental representation of the target situation, a factor that has been suggested to be implicated in Spatial Presence (Schubert et al., 2001).

VPs were associated with higher ratings of Realness than RPs, suggesting that judgments of Realness are based on representations of patient features (e.g., emotional displays or situation-specific behavior) over professionally relevant subtask representation.

Because one of the frequently discussed benefits of simulations includes the use of various scaffolding measures, we analyzed the effect of a measure that has received substantial attention: prompted reflection and might or might not be associated with discontinuous simulation. As expected, the analysis of the effect of discontinuity on ratings of perceived authenticity yielded a significant effect, supporting our hypothesis that interrupting a continuous interview simulation is associated with a reduction in perceived authenticity. As outlined, discontinuation of the simulated interview and its interjection through a reflection phase requires disengagement and reallocation of attention, which affects involvement, introduces stimuli conflicting with mental representation (Schubert et al., 2001), and requires reconstruction of said mental representation (Spatial Presence). Moreover, prompted reflection might not be consistent with learners' expectations of the target situation in professional practice (Realness).

Theoretical implications

RPs required the full subtasks of assessing which pieces of information may be diagnostically relevant and formulating a feasible question to elicit this information from the patient. Moreover, they comprised high interactivity and required flexibility in the context of the communication process. This was not the case for VPs. Here, the participants were offered a choice of possible questions and received complete information after selecting it with a click. Thus, the representation of the subtask of generating evidence was strongly reduced in the VP simulation. On the other hand, situation-specific patient properties were modeled in more detail in VPs, offering the representation through trained actors and including emotional representation in each videoclip.

We conclude that Realness judgments may be based on patients' representations, more than subtask representations of the diagnostic process. In the discussion generated by Hamstra et al. (2014), who related authenticity to task alignment rather than physical resemblance, we would incorporate our results in the following way: Perceptions of Realness may depend on the representation of objects' non-task-relevant aspects rather than detailed subtask representation

when human interaction is simulated. Behavioral displays of simulated patients in RPs may be viewed as being in disagreement with predictions of objects' behavior, and even if not directly task-relevant, this information may conflict with a coherent experience. Potential conflict between object prediction and representation may be particularly relevant when simulating human behavior. This is in line with the view that observing others entails action simulation and prediction, including higher cognition (Brass et al., 2000; Burke et al., 2010; Graf et al., 2007; Schuch & Tipper, 2007; van Schie et al., 2004; Wilson & Knoblich, 2005) or their mental states including emotion (Apperly et al., 2009; Premack & Woodruff, 1978).

We consider another explanation: In teaching practice, the use of RPs can be accompanied by low motivation to take part, unreadiness to accept the suggested environment as an alternative reality, and nonverbal and paraverbal cues to fellow students referencing the "as if" character of the simulation (e.g., laughter). Muckler (2017) discussed the suspension of disbelief as important for effective simulation. Such an unreadiness to suspend disbelief might complicate learners' mental representations of the target environment, thereby decreasing authenticity. This phenomenon should not simply be considered a confounding factor but instead needs to be addressed in teaching practice. However, in the RP simulation in this study, RPs took place in small groups of three or four students, with a precise task to deliver a correct diagnosis after the interview. We argue that this group setting should limit this potential influence. We also argue that lower motivation would not differ significantly between the three settings. Importantly, none of the simulation settings differed in Involvement ratings. We take this finding as evidence that each of these settings is feasible for engaging learners and offer an environment feasible for sufficient attentional allocation to provide opportunities for learning.

Based on our results, we also suggest that the assessment of a simulation's three-dimensional authenticity profiles yields relevant insights and allows for a better-targeted design in simulation development. Current and future developments in computer-based simulation include AI-driven approaches and bear influence on complexity, interactivity, adaptivity, adaptation, and surface representation (e.g., automatic image and avatar generation). Future systematic research on such characteristics' impact on multidimensional authenticity profiles could shed further light on these factors' effects on learner experience.

Practical implications

Perceptions of Realness and Spatial Presence are best induced with SP simulations. However, these simulations are associated with ongoing expenses and organizational effort (Weller et al., 2012; Taylor et al., 2019). VPs, as a method that offers many benefits (e.g., scalability, readiness for use, and adaptability to the learner; Cook & Triola, 2009), seem to induce higher perceptions of Realness than RPs and offer the benefit of standardization. This strengthens the argument for a proposed extension of the use of VPs to foster diagnostic competence (Cook & Triola, 2009; Berman et al., 2016) and illustrates its potentially advantageous profile in terms of perceived authenticity over the more traditional form of RPs. However, our results suggest that the simulation settings induced Involvement equally. Although Realness was judged lower for RPs, it may be preferable as a complement for fostering skill acquisition where more interactivity may be primal and may provide greater benefit (e.g., in communication skills training), and the learning goal includes the promotion of the patient-perspective and empathy (Bosse et al., 2012).

As scaffolding has been shown to be highly effective in facilitating the learning of complex skills (Rastle et al., 2021; Chernikova et al., 2020), one could argue that upholding perceived

authenticity may be secondary to the gain through instructional guidance that has been repeatedly demonstrated across different learning scenarios (Kirschner et al., 2006) and VPs in particular (Belland et al., 2017). However, suppression of competing or inconsistent stimuli and reallocation of attention may use up cognitive resources, which are limited (Miller, 1956; Chandler & Sweller, 1991). On the basis of our interpretation of these results, we would like to suggest implications for scaffolding design: These results favor scaffolds that do not require discontinuation of simulation, such as representational reduction (e.g., in Grossman et al., 2009) or enhancing the salience of relevant information, which has also been shown to impact authenticity (Chernikova et al., 2020). However, direct instructional guidance (Kirschner et al., 2006) could be represented as consistent with the simulated situation. To illustrate, encouragement to reflect on past or future actions (review by Mann et al., 2009) can be designed to be inconsistent or consistent with learners' mental representations. An example of prompting that is consistent with the simulated environment in VPs could be the simulation of a colleague or the patient asking questions, thus prompting reflection. Research that systematically investigates the effect of instructional scaffolds that are consistent with the simulated environment may shed further light on this issue. We conclude that scaffolding may compromise or reduce perceptions of authenticity in a simulation due to disruption of continuity and that further research on this point is needed.

Limitations

We would like to point out the following limitations: Although within each original study, participants were randomized into the experimental conditions such as participatory role and scaffolding, our analysis is based on a quasi-experimental design without randomized assignment of participants across settings. The original studies are associated with differences in the time of data collection, thus leading to unequal sampling from cohorts across settings.

However, participation in one setting over others was not based on setting preferences or aptitude. Furthermore, the sample was drawn from the same population of medical students in years 3–6 of their training. Participants received either financial or course credit as compensation. These different incentives across studies may have influenced motivation.

Further, we did not account for influences such as motivational factors or assess prior readiness to accept the simulated environment as an alternative reality (suspension of disbelief; Muckler, 2017), as our working definition of perceived authenticity was based on the simulation settings' aptitude to facilitate mental representation of the target situation, including factors such as learners' motivation.

In this study, we focus on authenticity as learners' subjective experience in the sense outlined above and examine its variation in relation to three types of simulations common in the medical education field. It is not within the scope of this study to formally assess and quantify the fidelity of these simulations in terms of task approximation. However, we aimed to discuss participants' perception of authenticity in relation to how well the simulation is representationally aligned with the target situation. To address the potential effects of discontinuity in more detail, including interactions with potential factors such as perspective or live versus media-based settings, a larger sample would be required.

Lastly, in contrast to a meta-analysis, our study draws on specific examples of the three simulation formats we seek to compare, limiting the generalizability of our results. The above-mentioned development of AI-driven approaches to simulation and its multitude of possibilities for implementation may also warrant a finer distinction between simulation scenarios that goes beyond our broad definition of a VP.

Conclusions

It seems worthwhile to assess different dimensions of perceived authenticity separately to evaluate a simulation in terms of a more detailed profile of perceived authenticity. The dimensions may be targeted separately and complement each other. In the example of implementation discussed here, VPs, which offer a variety of benefits, are associated with equal involvement and higher Realness perceptions than traditional live simulation through RPs. Analysis of authenticity profiles may be even more relevant with the development of AI-driven simulation, facilitating possibilities for (automated response generation) adaptation, adaptivity, and individualized interactivity with the use of surface representation such as avatars and automated image generation). In the context of anamnestic interview simulation, Realness judgments may be based on patient representation, not subtask representation. Discontinuation by scaffolding impairs perceived authenticity and should be considered a factor in simulation or scaffold designs. We suggest that future research address this gap in research on simulation-based learning.

Appendix

Model fit for a	a 1-factor	model of	authenticity:
-----------------	------------	----------	---------------

Test statistic	254.385				
Degrees of freedom	35				
P-value (Chi-square)	0.000				
Model Test Baseline Model:					
Test statistic	4348.411				
Degrees of freedom	45				
P-value	0.000				
User Model versus Baseline Model:					
Comparative Fit Index (CFI)	0.949				
Tucker-Lewis Index (TLI)	0.934				
Root Mean Square Error of A	oproximation:				

RMSEA 0.106

Covariances of Realness, Involvement and Spatial Presence

	Estimate	Std.Error	z-value	P(> z)	Std.lv	Std.all	
Realness							
Involvement	0.371	0.026	14.029	0.000	0.499	0.499	
Spatial presence	0.883	0.043	20.444	0.000	0.791	0.791	
Involvement							
Spatial presence	0.420	0.029	14.629	0.000	0.617	0.617	

Std.lv: latent variables standardized; Std.all: latent and observed variables standardized

Acknowledgements We would like to thank our colleagues Maximilian Fink, Hannah Gerstenkorn, Frank Fischer, and Matthias Siebeck for their contribution and Jane Zagorski for proofreading.

Funding Open Access funding enabled and organized by Projekt DEAL. The study received funding from the German Research Foundation (Deutsche Forschungsgemeinschaft, DFG) (Research Unit COSIMA, Project Number FOR2385).

Data Availability The data that support the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of interest The authors declare no competing interests.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

- Anderson, J. R., Reder, L. M., & Lebiere, C. (1996). Working memory: Activation limitations on retrieval. Cognitive Psychology, 30(3), 221–256.
- Apperly, I. A., Samson, D., & Humphreys, G. W. (2009). Studies of adults can inform accounts of theory of mind development. *Developmental Psychology*, 45(1), 190.
- Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1235–1243.
- Barrows, H. S. (1993). An overview of the uses of standardized patients for teaching and evaluating clinical skills. AAMC Academic Medicine, 68(6), 443–451.
- 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.
- Berman, N. B., Durning, S. J., Fischer, M. R., Huwendiek, S., & Triola, M. M. (2016). The role for virtual patients in the future of Medical Education. *Academic Medicine: Journal of the Association of American Medical Colleges*, 91(9), 1217–1222. https://doi.org/10.1097/ACM.000000000001146
- Bland, A. J., Topping, A., & Tobbell, J. (2014). Time to unravel the conceptual confusion of authenticity and fidelity and their contribution to learning within simulation-based nurse education. A discussion paper. *Nurse Education Today*, 34(7), 1112–1118.
- Bosse, H. M., Schultz, J. H., Nickel, M., Lutz, T., Möltner, A., Jünger, J., Huwendiek, S., & Nikendei, C. (2012). The effect of using standardized patients or peer role play on ratings of undergraduate communication training: A randomized controlled trial. *Patient Education and Counseling*, 87(3), 300–306. https://doi.org/10.1016/j.pec.2011.10.007

- Brass, M., Bekkering, H., Wohlschläger, A., & Prinz, W. (2000). Compatibility between observed and executed finger movements: Comparing symbolic, spatial, and imitative cues. *Brain and Cognition*, 44(2), 124–143.
- Burke, C. J., Tobler, P. N., Baddeley, M., & Schultz, W. (2010). Neural mechanisms of observational learning. Proceedings of the National Academy of Sciences of the United States of America, 107(32), 14431–14436. https://doi.org/10.1073/pnas.1003111107
- Chandler, P., & Sweller, J. (1991). Cognitive load theory and the format of instruction. *Cognition and Instruction*, 8(4), 293–332. https://doi.org/10.1207/s1532690xci0804_2
- Chernikova, O., Heitzmann, N., Stadler, M., Holzberger, D., Seidel, T., & Fischer, F. (2020). Simulationbased learning in higher education: A meta-analysis. *Review of Educational Research*, 90(4), 499–541.
- Codreanu, E., Sommerhoff, D., Huber, S., Ufer, S., & Seidel, T. (2020). Between authenticity and cognitive demand: Finding a balance in designing a video-based simulation in the context of mathematics teacher education. *Teaching and Teacher Education*, 95, 103146.
- Cook, D. A., & Triola, M. M. (2009). Virtual patients: A critical literature review and proposed next steps. *Medical Education*, 43(4), 303–311. https://doi.org/10.1111/j.1365-2923.2008.03286.x
- Cook, D. A., Brydges, R., Zendejas, B., Hamstra, S. J., & Hatala, R. (2013). Technology-enhanced simulation to assess health professionals: A systematic review of validity evidence, research methods, and reporting quality. *Academic Medicine*, 88(6), 872–883.
- D'Angelo, C., Rutstein, D., Harris, C., Bernard, R., Borokhovski, E., & Haertel, G. (2014). Simulations for STEM learning: Systematic review and meta-analysis. *Menlo Park: SRI International*, 5(23), 1–5.
- Davidsson, P., Verhagen, H. (2017). Types of Simulation. In: Edmonds, B., Meyer, R. (eds) Simulating Social Complexity: A handbook. Understanding Complex Systems. Springer, Cham. https://doi.org/10. 1007/978-3-319-66948-9_3
- de Jong, T. (2010). Instruction based on computer simulations. In R. E. Mayer & P. A. Alexander (Eds.), Handbook of research on learning and instruction (pp. 446–466). Routledge.
- Dutton, J. M., & Starbuck, W. H. (1971). Computer simulation of human behavior. Wiley.
- Ericsson, K. A., & Kintsch, W. (1995). Long-term working memory. Psychological Review, 102(2), 211– 245. https://doi.org/10.1037/0033-295X.102.2.211
- Fink, M. C., Reitmeier, V., Stadler, M., Siebeck, M., Fischer, F., & Fischer, M. R. (2021a). Assessment of diagnostic competences with standardized patients versus virtual patients: experimental study in the context of history taking. *Journal of Medical Internet Research*, 23(3), e21196. https://doi.org/10. 2196/21196
- Fink, M. C., Heitzmann, N., Siebeck, M., Fischer, F., & Fischer, M. R. (2021b). Learning to diagnose accurately through virtual patients: Do reflection phases have an added benefit? *BMC Medical Education*, 21, 523.
- Frank, B. (2015). Validierung. Presence messen in laborbasierter Forschung mit Mikrowelten: Entwicklung und erste Validierung eines Fragebogens zur Messung von Presence, 51–61. Springer Fachmedien Wiesbaden. https://doi.org/10.1007/978-3-658-08148-5
- Frankland, P. W., Josselyn, S. A., & Köhler, S. (2019). The neurobiological foundation of memory retrieval. *Nature Neuroscience*, 22(10), 1576–1585.
- Friedrich, M. J. (2002). Practice makes perfect: Risk-free medical training with patient simulators. Journal of the American Medical Association, 288(22), 2808–2812. https://doi.org/10.1001/jama.288.22.2808
- Gaba, D. M. (2004). The future vision of simulation in health care. *BMJ Quality & Safety, 13*(suppl 1), i2–i10.
- Godden, D. R., & Baddeley, A. D. (1975). Context-dependent memory in two natural environments: On land and underwater. *British Journal of Psychology*, 66(3), 325–331.
- Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *NeuroImage*, 36, T22–T32.
- Groot, F., Jonker, G., Rinia, M., Cate, T., & Hoff, R. G. (2020). Simulation at the frontier of the zone of proximal development: A test in acute care for inexperienced learners. *Academic Medicine*, 95(7), 1098–1105.
- Grossman, P. (Ed.). (2021). Teaching core practices in teacher education. Harvard Education.
- Grossman, P., Compton, C., Igra, D., Ronfeldt, M., Shahan, E., & Williamson, P. W. (2009). Teaching practice: A cross-professional perspective. *Teachers College Record*, 111(9), 2055–2100.
- Grossman, R., Heyne, K., & Salas, E. (2015). Game- and simulation-based approaches to training. In K. Kraiger, J. Passmore, N. R. dos Santos, & S. Malvezzi (Eds.), *The Wiley Blackwell handbook of the psychology of training, development, and performance improvement* (pp. 205–223). Wiley Blackwell. https://doi.org/10.1002/9781118736982.ch12

- Hamstra, S. J., Brydges, R., Hatala, R., Zendejas, B., & Cook, D. A. (2014). Reconsidering fidelity in simulation-based training. Academic Medicine, 89(3), 387–392. https://doi.org/10.1097/acm.00000 00000000130
- Hatfield, E., Cacioppo, J. T., & Rapson, R. L. (1993). Emotional contagion. Current Directions in Psychological Science, 2(3), 96–100. https://doi.org/10.1111/1467-8721.ep10770953
- Hegland, P. A., Aarlie, H., Strømme, H., & Jamtvedt, G. (2017). Simulation-based training for nurses: Systematic review and meta-analysis. *Nurse Education Today*, 54, 6–20.
- Heitzmann, N., Seidel, T., Opitz, A., Hetmanek, A., Wecker, C., Fischer, M., & Fischer, F. (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.
- Hofer, M. (2016). Presence und involvement (1st ed., pp. 978-973). Nomos.
- Instruct (2021). CASUS. Retrieved August 12, 2023, from https://www.instruct.eu/
- Jeffries, P. R. (2012). Simulation in nursing education: From conceptualization to evaluation. 2nd edition, National League for Nursing.
- Jennebach, J., Ahlers, O., Simonsohn, A., Adler, M., Özkaya, J., Raupach, T., & Fischer, M. R. (2022). Digital patient-centred learning in medical education: A national learning platform with virtual patients as part of the DigiPaL project. GMS Journal for Medical Education, 39(4), Doc47. https:// doi.org/10.3205/zma001568
- Joyner, B., & Young, L. (2006). Teaching medical students using role play: Twelve tips for successful role plays. *Medical Teacher*, 28(3), 225–229.
- Kaufman, D., & Ireland, A. (2016). Enhancing teacher education with simulations. *TechTrends*, 60, 260–267.
- 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.
- Kononowicz, A. A., Zary, N., Edelbring, S., Corral, J., & Hege, I. (2015). Virtual patients-what are we talking about? A framework to classify the meanings of the term in healthcare education. BMC Medical Education, 15(1), 1–7.
- Lavoie, P., Deschênes, M. F., Nolin, R., Bélisle, M., Garneau, A. B., Boyer, L., & Fernandez, N. (2020). Beyond technology: A scoping review of features that promote fidelity and authenticity in simulation-based health professional education. *Clinical Simulation in Nursing*, 42, 22–41.
- Lessiter, J., Freeman, J., Keogh, E., & Davidoff, J. (2001). A cross-media presence questionnaire: The ITC-Sense of Presence Inventory. *Presence: Teleoperators & Virtual Environments*, 10(3), 282–297.
- Maier, J., Eckert, C., & Clarkson, J. (2017). Model granularity in engineering design Concepts and framework. *Design Science*, 3, E1. https://doi.org/10.1017/dsj.2016.16
- Mamede, S., Schmidt, H. G., & Penaforte, J. C. (2008). Effects of reflective practice on the accuracy of medical diagnoses. *Medical Education*, 42(5), 468–475.
- Mamede, S., van Gog, T., Moura, A. S., de Faria, R. M., Peixoto, J. M., Rikers, R. M., & Schmidt, H. G. (2012). Reflection as a strategy to foster medical students' acquisition of diagnostic competence. *Medical Education*, 46(5), 464–472.
- Mamede, S., Van Gog, T., Sampaio, A. M., De Faria, R. M. D., Maria, J. P., & Schmidt, H. G. (2014). How can students' diagnostic competence benefit most from practice with clinical cases? The effects of structured reflection on future diagnosis of the same and novel diseases. *Academic Medicine*, 89(1), 121–127.
- Mann, K., Gordon, J., & MacLeod, A. (2009). Reflection and reflective practice in health professions education: a systematic review. Advances in Health Sciences Education, 14, 595–621.
- McAlpin, E., Levine, M., & Plass, J. L. (2023). Comparing two whole task patient simulations for two different dental education topics. *Learning and Instruction*, 83, 101690.
- Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, 63(2), 81.
- Muckler, V. C. (2017). Exploring suspension of disbelief during simulation-based learning. *Clinical Simulation in Nursing*, 13(1), 3–9.
- Peterson, L. R., & Peterson, M. J. (1959). Short-term retention of individual verbal items. Journal of Experimental Psychology, 58, 193–198. https://doi.org/10.1037/h0049234
- Pillemer, D. (2003). Directive functions of autobiographical memory: The guiding power of the specific episode. *Memory (Hove, England)*, 11(2), 193–202.
- Premack, D., & Woodruff, G. (1978). Does the chimpanzee have a theory of mind? Behavioral and Brain Sciences, 1(4), 515–526.

- Rastle, K., Lally, C., Davis, M. H., & Taylor, J. S. H. (2021). The dramatic impact of explicit instruction on learning to read in a new writing system. *Psychological Science*, 32(4), 471–484.
- Rosseel, Y. (2012). Lavaan: An R Package for Structural equation modeling. Journal of Statistical Software, 48(2), 1–36. https://doi.org/10.18637/jss.v048.i02
- RStudio Team (2022). *RStudio: Integrated Development Environment for R.* RStudio, PBC, Boston, MA. Retrieved 15 July, 2023. http://www.rstudio.com
- Salas, E., Wildman, J. L., & Piccolo, R. F. (2009). Using simulation-based training to enhance management education. Academy of Management Learning & Education, 8(4), 559–573.
- Schlaffke, L., Lissek, S., Lenz, M., Juckel, G., Schultz, T., Tegenthoff, M., Schmidt-Wilcke, T., & Brüne, M. (2015). Shared and nonshared neural networks of cognitive and affective theory-of-mind: A neuroimaging study using cartoon picture stories. *Human Brain Mapping*, 36(1), 29–39. https://doi. org/10.1002/hbm.22610
- Schubert, T., Regenbrecht, H., & Friedmann, F. (2000). Real and illusory interaction enhance presence in virtual environments. Paper presented at the 3rd International Workshop on Presence, University of Delft, The Netherlands.
- Schubert, T., Friedmann, F., & Regenbrecht, H. (2001). The experience of presence: Factor analytic insights. Presence: Teleoperators & Virtual Environments, 10(3), 266–281.
- Schuch, S., & Tipper, S. P. (2007). On observing another person's actions: Influences of observed inhibition and errors. *Perception & Psychophysics*, 69(5), 828–837.
- Seidel, T., Stürmer, K., Blomberg, G., Kobarg, M., & Schwindt, K. (2011). Teacher learning from analysis of videotaped classroom situations: Does it make a difference whether teachers observe their own teaching or that of others? *Teaching and Teacher Education*, 27(2), 259–267.
- Stachowiak, H. (1973). Allgemeine Modelltheorie. Springer.
- Stadler, M., Iliescu, D., & Greiff, S. (2021). Validly authentic: Some recommendations to researchers using simulations in psychological assessment [Editorial]. *European Journal of Psychological* Assessment, 37(6), 419–422. https://doi.org/10.1027/1015-5759/a000686
- Stegmann, K., Pilz, F., Siebeck, M., & Fischer, F. (2012). Vicarious learning during simulations: Is it more effective than hands-on training? *Medical Education*, 46(10), 1001–1008. https://doi.org/10. 1111/j.1365-2923.2012.04344.x
- Taylor, S., Haywood, M., & Shulruf, B. (2019). Comparison of effect between simulated patient clinical skill training and student role play on objective structured clinical examination performance outcomes for medical students in Australia. *Journal of Educational Evaluation for Health Professions*, 16, 3. https://doi.org/10.3352/jeehp.2019.16.3
- The jamovi project (2022). *jamovi*. (Version 2.3) [Computer Software]. Retrieved 01 August, 2023 from https://www.jamovi.org
- Tulving, E., & Pearlstone, Z. (1966). Availability versus accessibility of information in memory for words. *Journal of Verbal Learning and Verbal Behavior*, 5(4), 381–391.
- Tulving, E., & Thomson, D. M. (1973). Encoding specificity and retrieval processes in episodic memory. *Psychological Review*, 80(5), 352.
- van Schie, H. T., Mars, R. B., Coles, M. G., & Bekkering, H. (2004). Modulation of activity in medial frontal and motor cortices during error observation. *Nature Neuroscience*, 7(5), 549–554.
- Villaume, W. A., Berger, B. A., & Barker, B. N. (2006). Learning motivational interviewing: Scripting a virtual patient. American Journal of Pharmaceutical Education, 70(2), 33. https://doi.org/10.5688/aj700233
- Vorderer, P., Wirth, W., Gouveia, F. R., Biocca, F., Saari, T., Jäncke, F., Böcking, S., Schramm, H., Gysbers, A., Hartmann, T., Klimmt, C., Laarni, J., Ravaja, N., Sacau, A., Baumgartner, T., & Jäncke, P. (2004). *MEC Spatial Presence Questionnaire (MECSPQ): Short Documentation and Instructions for Application.* Report to the European Community, Project Presence: MEC (IST-2001-37661).
- Vu, N. V., & Barrows, H. S. (1994). Use of standardized patients in clinical assessments: Recent developments and measurement findings. *Educational Researcher*, 23(3), 23–30. https://doi.org/10.3102/ 0013189x023003023]
- Vygotsky, L. S., & Cole, M. (1978). Mind in society: Development of higher psychological processes. Harvard University Press.
- Weller, J. M., Nestel, D., Marshall, S. D., Brooks, P. M., & Conn, J. J. (2012). Simulation in clinical teaching and learning. *Medical Journal of Australia*, 196(9), 594–594.
- Wilson, M., & Knoblich, G. (2005). The case for motor involvement in perceiving conspecifics. *Psychological Bulletin*, 131(3), 460.
- Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psy*chology and Psychiatry, 17(2), 89–100.

Caroline Corves. Caroline.Corves@med.uni-muenchen.de

Current themes of research:

Simulation-based learning. Categorization. Diagnostic reasoning

Most relevant publications in the field of Psychology of Education:

No previous publication in Psychology of Education, closest connection: Graf, M., Reitzner, B., Corves, C., Casile, A., Giese, M., & Prinz, W. (2007). Predicting point-light actions in real-time. *Neuroimage*, 36, T22-T32

Matthias Stadler. Matthias.Stadler@med.uni-muenchen.de

Current themes of research:

Learning analytics. Simulation-based learning. Computer-based assessment

Most relevant publications in the field of Psychology of Education:

Stadler, M., Iliescu, D., & Greiff, S. (2021). Validly authentic: Some recommendations to researchers using simulations in psychological assessment. *European Journal of Psychological Assessment*, 37(6), 419–422. https://doi.org/10.1027/1015-5759/a000686

Martin R. Fischer. Martin.Fischer@med.uni-muenchen.de

Current themes of research:

Diagnostic reasoning. Simulation-based learning. Faculty and curriculum development

Most relevant publications in the field of Psychology of Education:

Fink, M.C., Heitzmann, N., Reitmeier, V., Siebeck, M., Fischer, F., & Fischer, M.R. (2023) Diagnosing virtual patients: the interplay between knowledge and diagnostic activities. Advances in Health Sciences Education, https://doi.org/10.1007/s10459-023-10211-4.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Authors and Affiliations

Caroline Corves^{1,2} · Matthias Stadler¹ · Martin R. Fischer¹

Caroline Corves Caroline.Corves@med.uni-muenchen.de

> Matthias Stadler Matthias.Stadler@med.uni-muenchen.de

Martin R. Fischer Martin.Fischer@med.uni-muenchen.de

- ¹ Institute of Medical Education, LMU University Hospital, LMU Munich, Pettenkoferstr. 8a, 80336 Munich, Germany
- ² Ludwig-Maximilians-Universität in Munich, Munich, Germany